USC-SIPI REPORT #418

An Introduction to Type-2 Fuzzy Logic Systems

by

Nilesh N. Karnik & Jerry M. Mendel

1998

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Viterbi School of Engineering
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Suite 400
Los Angeles, CA 90089-2564 U.S.A.

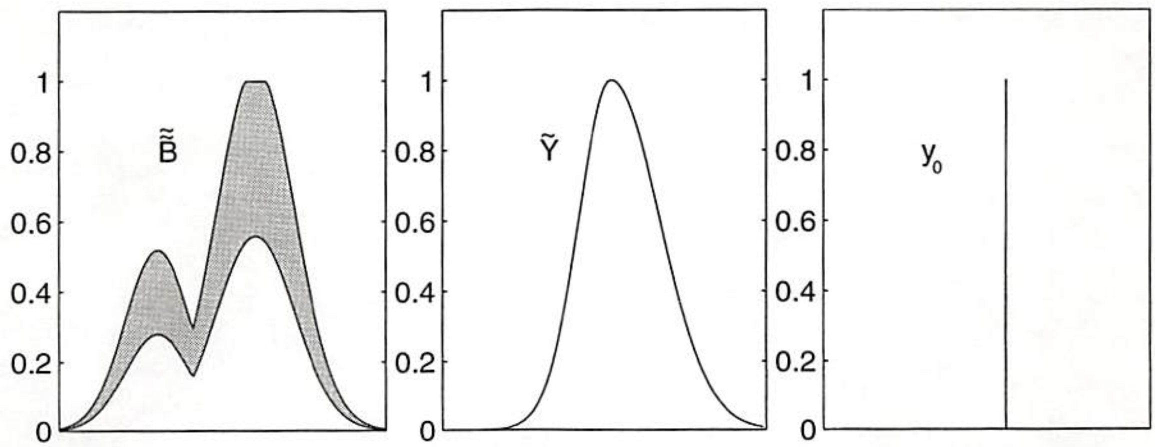
An Introduction to Type-2 Fuzzy Logic Systems

Nilesh N. Karnik & Jerry M. Mendel

Signal & Image Processing Institute
Dept. of Electrical Engineering - Systems
UNIVERSITY OF SOUTHERN CALIFORNIA
3740 McClintock Ave., EEB 400
Los Angeles, CA 90089-2564, USA

An Introduction to Type-2 Fuzzy Logic Systems

Nilesh N. Karnik & Jerry M. Mendel



Abstract

We introduce a new class of fuzzy logic systems (FLS), type-2 fuzzy logic systems, which make use of type-2 fuzzy sets for representing linguistic and/or numerical uncertainty. Type-2 fuzzy sets are fuzzy sets having fuzzy membership functions, i.e., the membership grade of each element of such a set is an ordinary (type-1) fuzzy set in $[0, 1]$. Type-2 sets are useful in circumstances where it is difficult to define the exact membership function for a fuzzy set.

Since the results in the existing type-2 fuzzy logic literature are not sufficient for our work, we start with the very basic operations of unions, intersections and complements of type-2 sets, and develop the results that we need in order to implement a type-2 FLS. In the process of this development, we study, in detail, set theoretic and algebraic operations for type-2 sets, properties of membership grades of type-2 sets, and type-2 relations and their compositions. We also examine, in great detail, the operations of "type-reduction" and defuzzification in a type-2 FLS. Type-reduction is a term that we have coined for "extended" versions of type-1 defuzzification methods. We provide results that considerably simplify the implementation of Gaussian and interval type-2 fuzzy logic systems. Whenever actual results are difficult to implement or generalize, we provide practical approximations.

We demonstrate the use of a type-2 FLS with the help of two examples : managing rules collected by means of a survey, and time-series prediction. In the survey example, we show how the linguistic uncertainty about membership functions of the FLS, as well as rule uncertainty from multiple experts, each of whom may give different answers to the same question, can be handled in the type-2 framework. In the time-series example, we show how information about the noise in the training data can be incorporated in a type-2 FLS to obtain bounds on the output and also better predictions.

Finally, we present our conclusions and some directions for future research.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. MIP-9419386. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the National Science Foundation.

Contents

List of Commonly Used Symbols

Set Notation

Symbol	Meaning
$\tilde{A}$	A type-1 fuzzy set
$\mu_{\tilde{A}}(x)$	The membership grade of x in type-1 fuzzy set $\tilde{A}$
$\int_x \mu_{\tilde{A}}(x)/x$	A type-1 fuzzy set, $\tilde{A}$, supported on a continuum
$\sum_x \mu_{\tilde{A}}(x)/x$	A type-1 fuzzy set, $\tilde{A}$, having a discrete support
$\tilde{\tilde{A}}$	A type-2 fuzzy set
$\tilde{\mu}_{\tilde{\tilde{A}}}(x)$	The membership grade of x in type-2 fuzzy set $\tilde{\tilde{A}}$
$C_{\tilde{A}}$	The centroid (a crisp number) of $\tilde{A}$
$\tilde{C}_{\tilde{\tilde{A}}}$	The centroid (a type-1 fuzzy set) of $\tilde{\tilde{A}}$

Operations on Fuzzy Sets

Symbol	Meaning
$\star, \mathcal{T}$	t-norm
$\vee$	maximum
$\wedge$	minimum
$\cup$	Union
$\cap$	Intersection
$\bar{(\cdot)}$	Complement of $(\cdot)$
$\sqcup$	Join
$\sqcap$	Meet
$\neg$	Negation
$\circ$	Sup-star composition
$\sum, +$	Algebraic sum or union of discrete quantities (as indicated)
$\int$	Integral or union on a continuum (as indicated)
$\times$	Algebraic or Cartesian product (as indicated)

Chapter 1

Introduction

Fuzzy logic systems are, as is well known, comprised of rules. Quite often, the knowledge that is used to construct these rules is *uncertain.* Such uncertainty leads to rules whose antecedents or consequents are uncertain, which translates into uncertain antecedent or consequent membership functions. Type-1 Fuzzy Logic Systems (FLS), whose membership functions are type-1 fuzzy sets, are unable to directly handle such rule uncertainties. We introduce a new class of fuzzy logic systems — type-2 fuzzy logic systems — in which the antecedent or consequent membership functions are type-2 fuzzy sets. Such sets are fuzzy sets whose membership grades themselves are type-1 fuzzy sets; they are very useful in circumstances where it is difficult to determine an exact membership function for a fuzzy set; hence, they are useful for incorporating uncertainties.

There is no prior work that we have been able to find on type-2 FLS's; hence, this work moves the world of fuzzy logic in a fundamentally new and important direction. What is this important direction and why is it important ? To make the answers as clear as possible, let us briefly digress to review some things that are, no doubt, familiar to the reader.

Probability theory is used to model random uncertainty, and within that theory we begin with a pdf, which embodies total information about random uncertainties. In most practical real-world applications, it is impossible to know or determine the pdf; so, we fall back on using the fact that a pdf is completely characterized by all of its moments. If the pdf is Gaussian, then, as is well known, two moments — the mean and variance — suffice to completely specify this pdf. For most pdfs, an infinite number of moments are required. Of course, it is not possible, in practice, to determine an infinite number of moments; so, instead, we compute as many moments as we believe are necessary to extract as much information as possible from the data. At the very least, we use two moments — the mean and variance; and, in some cases, we even use higher-than-second-order moments.

To just use the first-order moments would not be very useful, because random uncertainty requires an understanding of dispersion about the mean, and this information is provided by the variance. So, our accepted probabilistic modeling of random uncertainty focuses to a large extent on methods that use *at least* the first two moments of a pdf. This is, for example, why designs based on minimizing mean-squared errors are so popular.

Should we expect any less of a FLS for rule uncertainties ? To-date, we may view the output of a type-1 FLS, as analogous to the mean of a pdf. We may view computing the defuzzified output of a type-1 FLS as analogous to computing the mean of a pdf. Just as variance provides a measure of dispersion about the mean, and is almost always used to capture more about probabilistic uncertainty in practical statistical-based designs , FLS's also need some measure of dispersion to capture more about rule uncertainties than just a single number. Type-2 FL provides this measure of dispersion, and seems to be as fundamental to the design of systems that include linguistic and/or numerical uncertainties, that translate into rule uncertainties, as variance is to the mean.

Let us now familiarize ourselves with the concept of a type-2 fuzzy set.

1.1 The Concept of a Type-2 Fuzzy Set

The concept of a *type-2 fuzzy set* was introduced by Zadeh [30] as an extension of the concept of an ordinary fuzzy set (henceforth called a *type-1 fuzzy set*). A type-2 fuzzy set is characterized by a fuzzy membership function, i.e., the membership value (or membership grade) for each element of this set is a fuzzy set in $[0, 1]$, unlike a type-1 set where the membership grade is a crisp number in $[0, 1]$. Such sets can be used in situations where there is uncertainty about the membership grades themselves, e.g., an uncertainty in the shape of the membership function or in some of its parameters. Consider the transition from ordinary sets to fuzzy sets. When we cannot determine the membership of an element in a set as 0 or 1, we use fuzzy sets of type-1. Similarly, when the circumstances are so fuzzy that we have trouble determining the membership grade even as a crisp number in $[0, 1]$, we use fuzzy sets of type-2.

This does not mean that we need to have extraordinarily fuzzy circumstances to use type-2 sets. We can look at the situation from a different perspective. When something is uncertain (e.g., a measurement), we have trouble determining its exact value, and in this case, using type-1 sets, of course, makes more sense than using crisp sets. But then, even in the type-1 sets, we specify the membership functions exactly, which seems counter-intuitive. If we can not determine the exact value of an uncertain quantity, how can we determine its exact membership grade in a fuzzy set ? Of course, this criticism applies to type-2 sets as well, because even though the membership grade is fuzzy, we specify the membership function of the membership grade exactly, which again seems counter-intuitive. If we continue thinking along these lines, we can say that no finite-type fuzzy set can represent uncertainty "completely". So, ideally, we need to use a type-∞ fuzzy set to "completely" represent uncertainty ! Of course, we can not do this in practice, so we have to use some finite-type sets. So, type-1 fuzzy sets can be thought of as a *first-order approximation* to the uncertainty in real life. Our work with type-2 fuzzy sets tries to get at a *second-order approximation.* One may look at higher types too; but, as we go on to higher types, the complexity of the system increases rapidly. So, in this work we deal just with type-2 sets.

1.2 Examples of Type-2 Fuzzy Sets

Example 1.1 Consider the case of a fuzzy set characterized by a Gaussian membership function with mean m and a standard deviation that can take values in $[\sigma_1, \sigma_2]$, i.e.,

$$\mu(x) = e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}; \quad \sigma \in [\sigma_1, \sigma_2] \tag{1.1}$$

Corresponding to each value of σ, we will get a different membership curve (see Fig. 1.1). So, the membership grade of any particular x (except for $x = m$) can take any of a number of possible values depending upon the value of σ, i.e., *the membership grade is not a crisp number, it is a fuzzy set.* Figure 1.1 shows the domain of the fuzzy set associated with $x = 0.65$; however, the membership function associated with this fuzzy set is not shown in the figure. We return to this point in Section 1.4. □

Example 1.2 Consider the case of a fuzzy set with a Gaussian membership function having a fixed standard deviation σ, but an uncertain mean, taking values in $[m_1, m_2]$, i.e.,

$$\mu(x) = e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}; \quad m \in [m_1, m_2] \tag{1.2}$$

Again, $\mu(x)$ is a fuzzy set. Figure 1.2 shows an example of such a set. As in Fig. 1.1, it is not possible to deduce the membership function associated with the fuzzy membership of any x from Fig. 1.2. □

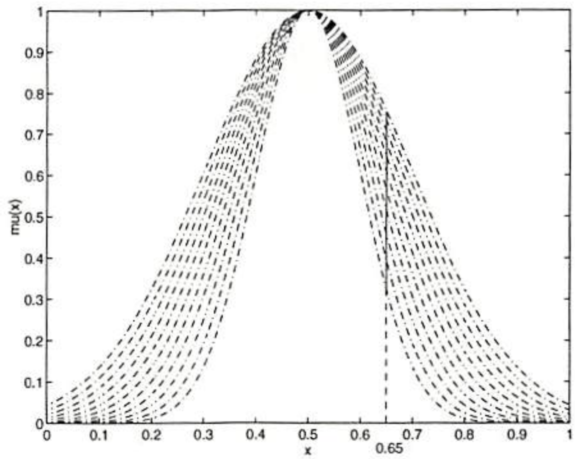


Figure 1.1: A type-2 fuzzy set representing a type-1 fuzzy set with uncertain standard deviation. The standard deviation is uncertain in the interval [0.1, 0.2]. The figure also shows the domain of the type-1 fuzzy set corresponding to x = 0.65; however, the membership grades in this type-1 fuzzy set are not shown.

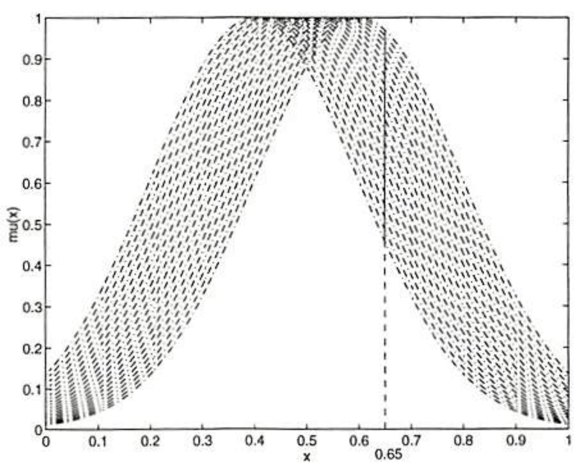


Figure 1.2: A type-2 fuzzy set representing a type-1 fuzzy set with uncertain mean. The mean is uncertain in the interval [0.4, 0.6]. The figure also shows the domain of the type-1 fuzzy set corresponding to x = 0.65; however, the membership grades in this type-1 fuzzy set are not shown.

In Examples 1.1 and 1.2, we considered fuzzy sets with Gaussian membership functions which had their standard deviations or means uncertain. Such sets can be used in situations where we want to use Gaussian fuzzy sets, but are not certain about their center or spread locations. If the situation is such, however, that we are uncertain even about the *shape* of the membership function (Gaussian / triangular / any other arbitrary shape), we can use a *Gaussian type-2 fuzzy set* defined as in Example 1.3.

Example 1.3 Consider a type-1 fuzzy set characterized by a Gaussian membership function (mean M and standard deviation σ_x), which gives one crisp membership $m(x)$ for each input $x \in X$, where

$$m(x) = e^{-\frac{1}{2}(\frac{x-M}{\sigma_x})^2} \tag{1.3}$$

This is depicted in Fig. 1.3. Now, imagine that this membership of x is a fuzzy set. Let us call the domain-elements of this set *primary memberships* of x (denoted by μ_1) and membership grades of these primary memberships *secondary memberships* of x [denoted by $\mu_2(x, \mu_1)$]. So, for a fixed x, we get a type-1 fuzzy set whose domain-elements are primary memberships of x and whose corresponding membership grades are secondary memberships of x. If we assume that the secondary memberships follow a Gaussian with mean $m(x)$ and standard deviation σ_m, as in Fig. 1.3, we can describe the secondary membership function for each x as

$$\mu_2(x, \mu_1) = e^{-\frac{1}{2}(\frac{\mu_1 - m(x)}{\sigma_m})^2} \tag{1.4}$$

where $\mu_1 \in [0, 1]$ and m is as in (1.3). Equations (1.3) and (1.4) can be combined as

$$\mu_2(x, \mu_1) = e^{-\frac{1}{2}(\frac{\mu_1 - e^{-\frac{1}{2}(\frac{x-M}{\sigma_x})^2}}{\sigma_m})^2} ; \tag{1.5}$$

where $\mu_1 \in [0, 1]$. Equation (1.5) stresses the fact that the secondary membership function can be viewed as a real function of two variables, x and μ_1. The membership grade for each x, $\mu(x)$, which represents all the primary memberships and their corresponding secondary memberships taken together, can be written as

$$\mu(x) = \int_{\mu_1 \in [0,1]} \mu_2(x, \mu_1)/\mu_1 \; ; \quad x \in X \tag{1.6}$$

where $\mu_2(x, \mu_1)$ is as in (1.5).

Observe that Example 1.3 is different than Examples 1.1 and 1.2 in that in Example 1.3 we are explicitly stating the secondary membership function. Actual values of the secondary membership grades were not defined in Examples 1.1 and 1.2. We return to Example 1.3 in Section 1.4. □

Now, let's see a situation in real life which needs to be described using type-2 fuzzy sets.

Example 1.4 Consider classes of people with *below average*, *average* and *above average* earnings. These sets, of course, are fuzzy. Now, if we ask someone what memberships s/he would have in these three fuzzy sets, most likely we are going to get an answer of the form "a *high* membership in *above average earnings* and low in others", rather than crisp numbers as memberships. This means that the membership grade is a fuzzy set or in other words, the aforementioned three fuzzy sets are of type-2 ! Observe that in this example, the person who is asked the question, knows her/his own income exactly, but the uncertainty in the membership grade arises due to the fact that s/he doesn't know the exact parameters of the membership functions for these 3 sets. (If Gaussian membership functions are used, this is analogous to ambiguity in mean and/or variance.)

In this example, "a *high* membership in *above average*" cannot be rephrased as "highly above average". In "highly above average", "highly" is a *hedge* on "above average", but after the application of this hedge, all we get is another type-1 fuzzy set, "highly above average". Now, if the

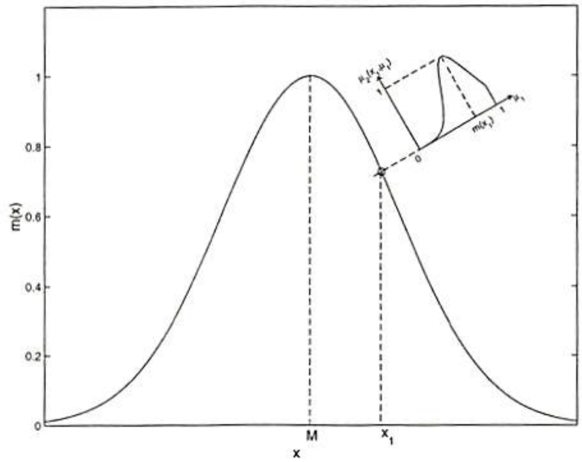


Figure 1.3: Figure for Example 1.3. The Gaussian $m(x)$ and membership grade corresponding to $x = x_1$ are shown. The membership grade is a Gaussian type-1 fuzzy set contained in $[0, 1]$ with mean $m(x_1)$.

same person were asked what would be her/his membership in the set "highly above average", s/he would probably say *medium*, which means the membership is again fuzzy and the set is still best described as a type-2 fuzzy set. If the person could not give a crisp membership for the "above average" set, the person would definitely not be able to give a crisp membership for the "highly above average" set [this new set is just a (possibly) non-linear transformation of the original one]. □

Let us contrast this example with that of a *non-singleton fuzzy logic system* [19]. Let's say we fix all the parameters of the membership functions characterizing the above 3 fuzzy sets and then ask someone how much membership a *rich* person would have in these three sets. Again, we are most probably going to get a fuzzy answer, but now the fuzziness is due to the uncertainty in the data (a *rich* person) and it can be modelled by using *non-singleton* systems. *A Non-Singleton system deals with the uncertainty in the input, whereas a type-2 system deals with the uncertainty in our knowledge about the system.* Hence, type-2 systems should be *robust* to rule uncertainties, whereas a non-singleton system is robust to measurement uncertainties.

From now on, we will use the membership terminology introduced in Example 1.3. *Membership grade* is a synonym for "degree of membership", which is a crisp number for type-1 sets, a type-1 set for type-2 sets, and in general, a type-k set for type-$(k+1)$ sets. In the case of type-2 sets, *primary memberships* are the domain-elements of a membership grade and *secondary memberships* are membership grades of primary memberships. For example, in (1.5) and (1.6), μ_1 indicates the primary memberships; $\mu_2(x, \mu_1)$ indicates the secondary memberships; and $\mu(x)$, which represents all the primary and secondary memberships taken together, indicates the membership grade of $x \in X$. Thus, for a type-1 set, $\mu(x)$ is short for $\mu_1(x)$ and $\mu_2(x, \mu_1) = 1$; and, for a type-2 set, $\mu(x)$ indicates the type-1 set $\int_{\mu_1} \mu_2(x, \mu_1)/\mu_1$.

A type-2 fuzzy set can also be thought of as a fuzzy valued function, which assigns to every $x \in X$, a type-1 fuzzy membership grade. In this sense, we will call X the *domain* of the type-2 fuzzy set.

1.3 Some Useful Type-2 Sets

Here we formally define three kinds of type-2 sets that we will talk about often in this report :

1. A *Gaussian type-2 set* is one in which the membership grade of every domain point is a Gaussian type-1 set contained in $[0, 1]$.

Example 1.3 shows an example of a Gaussian type-2 set. Note that it is not necessary for the principal membership function of a Gaussian type-2 set to also be a Gaussian, as is the case in Example 1.3. Figure 1.9 shows an example of a Gaussian type-2 set having a triangular principal membership function, using the 2-D pictorial representation described in Section 1.4.

2. An *interval type-2 set* is one in which the membership grade of every domain point is a crisp set whose domain is some interval contained in $[0, 1]$.

In Example 1.1, if we attach equal degree of uncertainty to every value of σ in $[\sigma_1, \sigma_2]$, i.e., if we let the standard deviation σ be a crisp set with domain $[\sigma_1, \sigma_2]$, we can set all the secondary memberships of the resulting type-2 set equal to 1. The membership grade corresponding to every x in this type-2 set, now, becomes a crisp set, and the type-2 set becomes an interval type-2 set (see Section 1.4).

Note that, although every membership grade of an interval type-2 set is a crisp set, the set itself is type-2, because the memberships are *sets* rather than crisp numbers. Interval type-2 sets are the simplest kind of type-2 sets to deal with, since all the secondary memberships are unity; and, we will often discuss them. We will refer to the membership grades of an interval type-2 set as "interval type-1 sets".

3. A *triangular type-2 set* is one in which the membership grade of every domain point is a triangular type-1 set contained in $[0, 1]$.

Unless otherwise specified, a "triangle" will always mean a "symmetrical triangle" in our work. The results for triangular type-2 sets are collected in [8].

1.4 Pictorial Representation

Now, let's try to represent a type-2 membership function pictorially. Observe that our earlier pictorial representations using 2-D plots (Figs. 1.1 and 1.2) did not indicate the numerical values of secondary memberships. All that one can see from those diagrams is just the set of primary memberships corresponding to each x. So, these representations do not contain all the information that we have. Recall the example of the Gaussian fuzzy set with uncertain mean. The 2-D diagram in Fig. 1.2 does not depend on the actual shape of the membership function for the fuzzy mean. It will remain the same as long as the support of the fuzzy set for the mean is $[m_1, m_2]$, which shows that these diagrams are not unique, i.e., we can get the same diagrammatic representation for two or more distinct situations. For example, Fig. 1.2 would remain unchanged if the fuzzy set for the mean followed a Gaussian membership curve or a triangular membership curve as long as the support is $[m_1, m_2]$. This indicates that the earlier pictorial representations are not "complete".

A type-2 membership function can be viewed as a function of two variables. For each input x and a primary membership μ_1, we get a secondary membership, which is a crisp number. Let's call this secondary membership μ_2. So, the membership function of a type-2 set can be represented as

$$\mu_2(x, \mu_1) \ : X \times [0, 1] \rightarrow [0, 1] \tag{1.7}$$

where X is the space of all inputs x. Pictorially, we can display this function as a 3-D diagram with x and μ_1 as independent variables and μ_2 as the dependent variable.

Recall Example 1.1 . Suppose that the degree of uncertainty that we attach to each value of σ in the range $[\sigma_1, \sigma_2]$ is the same; in other words, let the standard deviation σ be a *crisp* set with domain $[\sigma_1, \sigma_2]$. Since, each value of standard deviation is equally uncertain, we set all the secondary memberships of the resulting type-2 set equal to 1, i.e., the membership grade corresponding to each x is an interval in $[0, 1]$ (the resulting type-2 set is an interval type-2 set). Figure 1.4 (a) shows a 3-D representation of this type-2 set, assuming $\sigma_1 = 0.1$ and $\sigma_2 = 0.2$, and Fig. 1.4 (b) shows the membership grade for $x = 0.65$; the domain of this membership grade is indicated in

Fig. 1.1. Figure 1.5 (a) shows the 3-D diagram for Example 1.2, drawn by assuming that the mean m is a crisp set with domain $[m_1, m_2] = [0.4, 0.6]$. Figure 1.5 (b) shows the membership grade corresponding to $x = 0.65$ in this interval type-2 set. See Appendix A for examples which let the standard deviation of the Gaussian in Example 1.1 and the mean of the Gaussian in Example 1.2 be Gaussian type-1 sets.

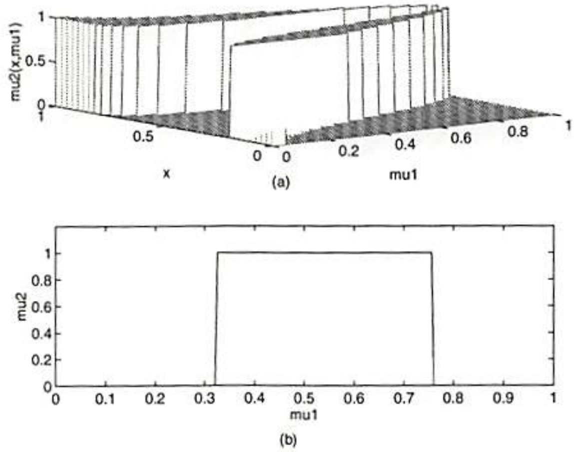


Figure 1.4: (a) Three dimensional representation of the type-2 set in Example 1.1, assuming that the standard deviation is a crisp set with domain $[\sigma_1, \sigma_2] = [0.1, 0.2]$. The membership grade for each x is a crisp set. (b) The membership grade corresponding to $x = 0.65$.

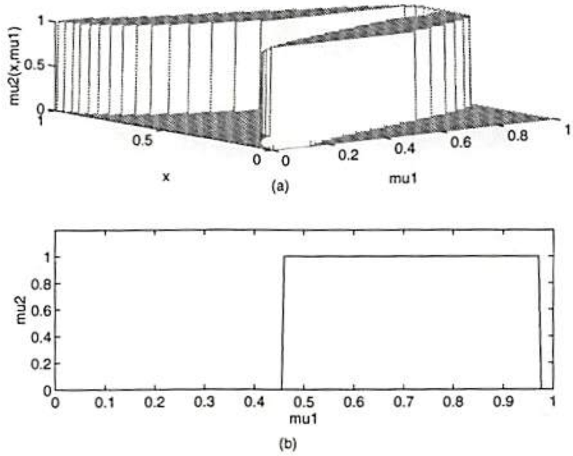


Figure 1.5: (a) Three dimensional representation of the type-2 set in Example 1.2, assuming that the mean is a crisp set with domain $[m_1, m_2] = [0.4, 0.6]$. The membership grade for each x is a crisp set. (b) The membership grade corresponding to $x = 0.65$.

Figure 1.6 (a) depicts a 3-D representation of (1.5) and Fig. 1.6 (b) depicts the fuzzy type-1 set $\mu(x)$ for an arbitrary value of x (obtained by taking a cross-section of Fig. 1.6 parallel to the $\mu_1 - \mu_2$ axes). $\mu(x)$ is a Gaussian, because we constructed it that way. Observe the similarity with a type-1 pictorial representation, where we display the membership function of a type-1 set as a 2-D picture (function of one variable, x).

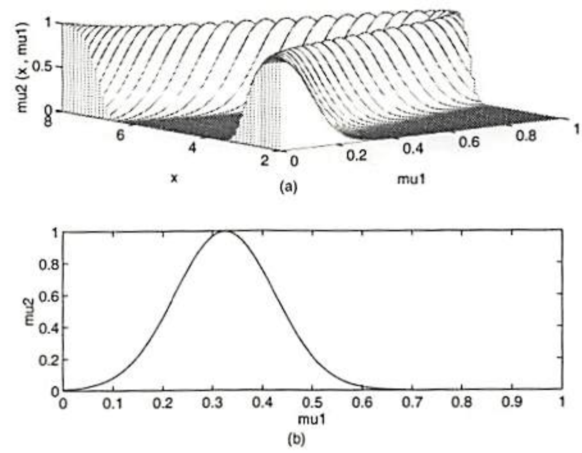


Figure 1.6: (a) Three dimensional representation of a Gaussian type-2 fuzzy set, having a Gaussian principal membership function. The membership grade for each x is Gaussian by construction. All these Gaussians have the same standard deviation. (b) The membership grade corresponding to $x = 6.5$.

Although the 3-D representation of a type-2 set conveys all the information that we have about the set, it is not very helpful to use these 3-D diagrams when we have to show more than one set on the same axes. Additionally, they can be quite complicated to construct. So, in spite of the aforementioned "incompleteness" of 2-D representations, we continue to use them in our analyses of type-2 sets. If there is a need to show the secondary membership functions explicitly, we will use a 3-D representation.

Figure 1.7 shows a 2-D representation of the Gaussian type-2 set depicted in Fig. 1.6 (a). We call the set of primary memberships that have secondary membership grades equal to 1, the *principal* membership function of the type-2 set (shown with a bold line in Fig. 1.7). From (1.4), we can see that $m(x)$ in (1.3) is the principal membership function. Since we are using Gaussian secondary membership functions for each input, only one primary membership has a secondary membership equal to 1. This seems reasonable, because the secondary membership function indicates the uncertainty in determining the membership grade for a particular input. We will often use secondary membership functions like Gaussians or triangles, which assign unity membership to only one point in their domain. Observe that this is not true for the crisp secondary membership functions shown in Figs. 1.4 and 1.5.

The concept of a principal membership function also illustrates the fact that a type-1 fuzzy set can be thought of as a special case of a type-2 fuzzy set. We can think of a type-1 fuzzy set as a type-2 fuzzy set whose membership grades are type-1 fuzzy singletons, having secondary membership equal to unity for only one primary membership and zero for all others, i.e., we can think of the principal membership function of a type-2 set as an embedded type-1 set. *Our fundamental design requirement in this work is that our type-2 system results reduce to type-1 results when we replace all the type-2 sets by their principal membership functions.*

Figure 1.8 (a) shows a 3-D representation of a Gaussian type-2 set having a *triangular* principal membership function, and Fig. 1.8 (b) shows the membership grade for $x = 6.5$. The 2-D representation of this set is depicted in Fig. 1.9. The difference between the two Gaussian type-2 sets, in Figs. 1.6 (a) and 1.8 (a), is seen more clearly in the 2-D representation.

The secondary membership functions for the Gaussian type-2 sets depicted in Figs. 1.6 and 1.8 have constant standard deviations, implying that the uncertainty in the membership grades remains constant for all x. Intuitively, however, it seems more appropriate that membership values

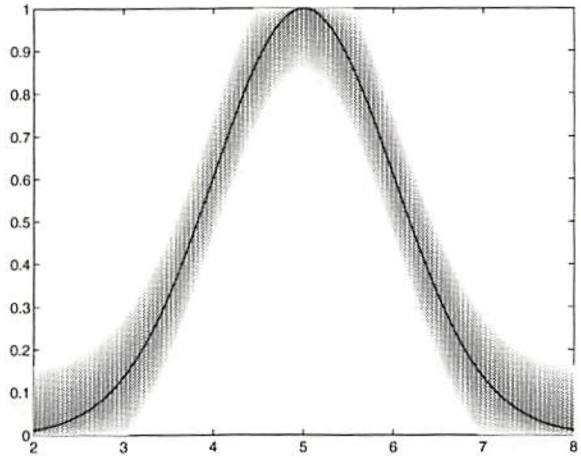


Figure 1.7: Two-dimensional representation of the Gaussian type-2 set depicted in Fig. 1.6 (a). The standard deviations of the secondary Gaussians are constant. The *principal* membership function, i.e., the set of primary memberships having secondary membership equal to 1, is indicated with a thick line. This principal membership function is a Gaussian because of the way the set is constructed. Intensity of the shading is approximately proportional to secondary membership grades. Darker areas indicate higher secondary memberships. The flat portion near the center and near the two ends, appears because primary memberships cannot be less than 0 or greater than 1 and so the Gaussians have to be "clipped".

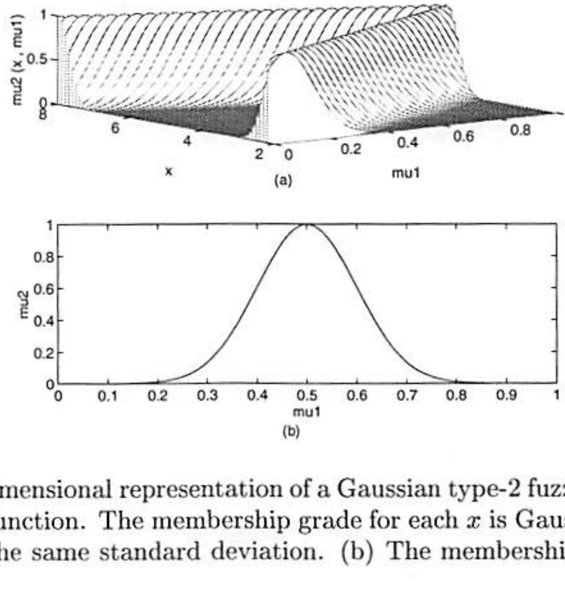


Figure 1.8: (a) Three dimensional representation of a Gaussian type-2 fuzzy set, having a triangular principal membership function. The membership grade for each x is Gaussian by construction. All these Gaussians have the same standard deviation. (b) The membership grade corresponding to $x = 6.5$.

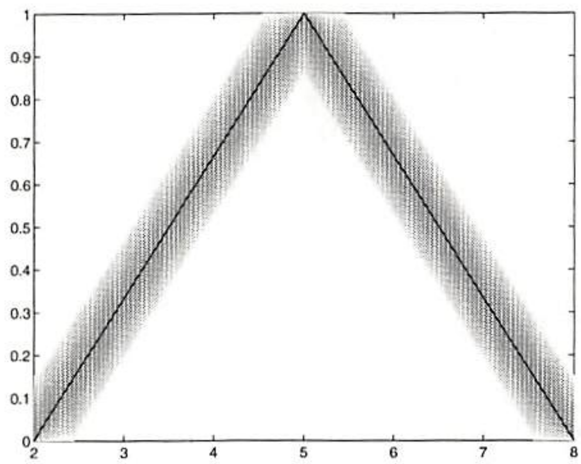


Figure 1.9: Two-dimensional representation of the Gaussian type-2 set depicted in Fig. 1.8 (a). The *principal* membership function is triangular. The standard deviations of the secondary Gaussians are constant.

near zero should have less uncertainty associated with them than membership values near 1. In other words, it seems more appropriate that the uncertainty in a membership value be expressible as some percentage of it. Such a type-2 set (Gaussian type-2 with Gaussian principal membership function) is depicted in Fig. 1.10. The secondary membership functions of this set have decreasing standard deviations, implying that the uncertainty in the membership grades decreases as x moves away from the mean of the principal membership function.

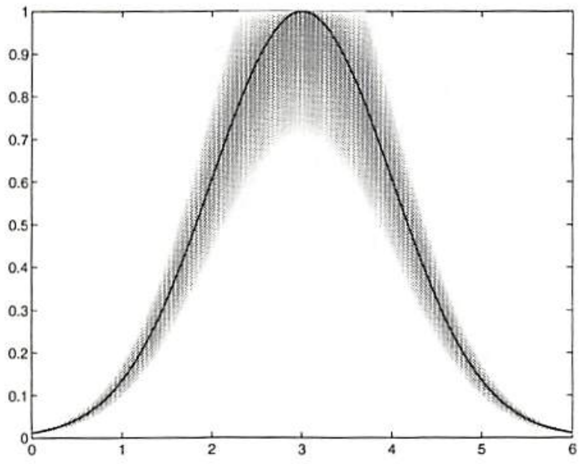


Figure 1.10: Two-dimensional representation of a Gaussian type-2 set, where standard deviations of the secondary Gaussians decrease by design, as x moves away from the mean of the principal membership function.

1.5 Applications of Type-2 Fuzzy Sets

Type-1 FLSs have been successfully used in widely different applications (see, for example, [5], [10], [11], [12], [13], [15], [19], [24], [25], [28], [16] and references therein). In this section, we give examples of some applications of fuzzy logic systems, which can be better modelled by type-2 sets.

Type-2 sets can be used to convey the uncertainties in membership functions of type-1 sets, due to the dependence of the membership functions on available linguistic and numerical information. Linguistic information (e.g., rules from experts), in general, does not give any information about the shapes of the membership functions. When membership functions are determined or tuned based on numerical data, the uncertainty in the numerical data, e.g., noise, translates into uncertainty in the membership functions. In all such cases, any available information about the linguistic/numerical uncertainty can be incorporated in the type-2 framework. Here, we describe three possible applications :

- When rules are collected by surveying experts, if we first determine the locations and spreads of fuzzy sets, associated with antecedent and consequent terms, based on the information gathered from surveys, it is very likely that we will get different answers from each survey. This leads to statistical uncertainties about locations and spreads of antecedent and consequent fuzzy sets. Such uncertainties can be incorporated into the descriptions of these sets using type-2 membership functions. In addition, usually, different experts give different answers to the same rule-question, which results in rules that have the same antecedents but different consequents. In such a case, it is also possible to represent the output of the FLS built from these rules as a fuzzy set rather than a crisp number. This can be achieved within the type-2 framework. We examine this application in Chapter 6.

- A fuzzy logic modulation classifier described in [28] centers type-1 Gaussian membership functions at constellation points in the in-phase/quadrature plane. In practice, the constellation points drift. This is analogous to the situation described in Fig. 1.2 [Gaussian with uncertain mean]; so, a type-2 formulation can capture this drift.

- All previous applications of FL to forecasting do not account for the noise in training data. In forecasting, since antecedents and consequents are the same variable, the uncertainty during training exists on both the antecedents and consequents. If we have information about the level of uncertainty, it can be used when we model antecedents and consequents as type-2 sets. We also examine this application in Chapter 6.

1.6 Existing Literature on Type-2 Sets

Here, we give a brief overview of the work on type-2 sets that has already been done by others. Zadeh [30] introduced the concept of a type-2 fuzzy set; however, the first paper that we have found that describes set theoretic operations on type-2 sets and properties of membership grades of type-2 sets, is by Mizumoto and Tanaka [17]. They examine type-2 fuzzy sets under the operations of algebraic product and algebraic sum in [18]. Nieminen [20] provides more detail about algebraic structure of type-2 sets.

The *join* and *meet* operations between fuzzy numbers under minimum t-norm have been discussed in the fuzzy arithmetic literature, under titles like "minimum and maximum of fuzzy numbers" or "fuzzification of minimum and maximum operations" (see, for example [5, 10]).

Dubois and Prade [4, 5] discuss fuzzy-valued logic and give a formula (without proof) for the composition of type-2 relations as an extension of the type-1 sup-star composition. This formula is the same as our extended sup-star composition of type-2 relations in Chapter 4. All of their work makes use of the minimum t-norm.

Literature on the applications of type-2 sets is scarce. Some examples are [1] and [29] for decision making, and [27] for solving fuzzy relational equations.

Interval type-2 sets are generally referred to as "interval valued fuzzy sets" [12, 30] in the literature. Some examples of the literature about interval valued fuzzy sets are [7, 23].

1.7 Outline

As is apparent from Section 1.6, not much work has been done in almost any area related to type-2 sets. So, in Chapter 2, we go back to the basics and discuss the set theoretic operations of unions, intersections and complements of type-2 sets and also algebraic operations on the membership grades of type-2 sets. We also introduce the concept of the centroid of a type-2 set in Chapter 2. In Chapter 3, we discuss properties of membership grades of type-2 sets in detail, and in Chapter 4, we describe type-2 fuzzy relations and operations between them. In Chapter 5, we discuss the implementation of type-2 FLS's; and, in Chapter 6, we present two examples (surveys and forecasting) that illustrate the use of a type-2 FLS. Finally, in Chapter 7, we present our conclusions.

Chapter 2

Operations on Type-2 Sets

In this chapter, we examine set theoretic operations on type-2 sets. We use the following notation. A type-1 fuzzy set P is denoted as $\tilde{\mathrm{P}}$. A type-2 fuzzy set A is denoted as $\tilde{\tilde{\mathrm{A}}}$. Consequently, if x_0 is an element of $\tilde{\tilde{\mathrm{A}}}$, the membership grade of x_0 in $\tilde{\tilde{\mathrm{A}}}$ is denoted as $\tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x_0)$. Recall that $\tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x_0)$ is itself a type-1 fuzzy set whose elements and their memberships are, respectively, the primary and secondary memberships of x_0.

2.1 Set Theoretic Operations

To begin, we recall some facts about type-1 sets. A fuzzy subset $\tilde{\mathrm{A}}$ of a set X is represented as follows :

$$\begin{aligned} \tilde{\mathrm{A}} &= \mu_{\tilde{\mathrm{A}}}(x_1)/x_1 + \mu_{\tilde{\mathrm{A}}}(x_2)/x_2 + \ldots + \mu_{\tilde{\mathrm{A}}}(x_n)/x_n \\ &= \sum_i \mu_{\tilde{\mathrm{A}}}(x_i)/x_i, \;\; x_i \in X \end{aligned} \tag{2.1}$$

where the sum represents *union*. If the support of $\tilde{\mathrm{A}}$ is a continuum, we write

$$\tilde{\mathrm{A}} = \int_X \mu_{\tilde{\mathrm{A}}}(x)/x \tag{2.2}$$

Suppose, we have 2 type-1 fuzzy sets $\tilde{\mathrm{F}}_1$ and $\tilde{\mathrm{F}}_2$ characterized by membership functions θ_1 and θ_2, as follows :

$$\tilde{\mathrm{F}}_1 = \sum_i \theta_1(y_i)/y_i \tag{2.3}$$

$$\tilde{\mathrm{F}}_2 = \sum_i \theta_2(y_i)/y_i \tag{2.4}$$

Using *max* t-conorm and *min* t-norm, the membership functions of the union, intersection and complement of these sets are given as [12] :

$$\mu_{\tilde{\mathrm{F}}_1 \cup \tilde{\mathrm{F}}_2}(y_i) = \max\{\theta_1(y_i), \theta_2(y_i)\} \;\; \forall i \tag{2.5}$$

$$\mu_{\tilde{\mathrm{F}}_1 \cap \tilde{\mathrm{F}}_2}(y_i) = \min\{\theta_1(y_i), \theta_2(y_i)\} \;\; \forall i \tag{2.6}$$

$$\mu_{\bar{\tilde{\mathrm{F}}}_1}(y_i) = 1 - \theta_1(y_i) \;\; \forall i \tag{2.7}$$

$$\mu_{\bar{\tilde{\mathrm{F}}}_2}(y_i) = 1 - \theta_2(y_i) \;\; \forall i \tag{2.8}$$

Since $\tilde{F}_1$ and $\tilde{F}_2$ are fuzzy sets of type-1, their membership grades $\theta_1(y_i)$ and $\theta_2(y_i)$ are crisp numbers and therefore, for each y_i, we can perform all the operations on the RHSs of Eqs. (2.5) - (2.8) in one step.

Now, suppose that $\tilde{\tilde{F}}_1$ and $\tilde{\tilde{F}}_2$ are type-2 fuzzy sets, so that the membership grades $\tilde{\theta}_1(y_i)$ and $\tilde{\theta}_2(y_i)$ are type-1 fuzzy sets. In order to compute the union, intersection and complement of $\tilde{\tilde{F}}_1$ and $\tilde{\tilde{F}}_2$, we need to extend the binary operations of *min* and *max*, and the unary operation of *negation* to fuzzy sets. We use Zadeh's *Extension Principle* for this purpose, which we state here for reference purposes.

The Extension Principle [5, 30] : Let X be a Cartesian product of universes, $X = X_1 \times X_2 \times \cdots \times X_r$, and $\tilde{A}_1$, $\tilde{A}_2$,...,$\tilde{A}_r$ be fuzzy sets in X_1, X_2, ...,X_r, respectively. Let f be a mapping from X to a universe Y such that $y = f(x_1, ..., x_r) \in Y$. Zadeh's Extension Principle allows us to induce from the r fuzzy sets $\tilde{A}_i$, a fuzzy set $\tilde{B}$ on Y, through f, such that

$$\begin{aligned} \mu_{\tilde{B}}(y) &= \sup_{x_1,...,x_r\,:y=f(x_1,...,x_r)} \min\{\mu_{\tilde{A}_1}(x_1), ..., \mu_{\tilde{A}_r}(x_r)\} \\ \mu_{\tilde{B}}(y) &= 0 \text{ if } f^{-1}(y) = \emptyset \end{aligned} \tag{2.9}$$

where $f^{-1}(y)$ is the inverse image of y under f. □

Equation (2.9) makes use of the *max* t-conorm. If for any primary membership in the union, we get more than one choice of secondary memberships, the effective secondary membership is taken to be the maximum of all these choices.

Equation (2.9) assumes that $x_1, \ldots, x_n$ are non-interactive or that there is no joint constraint on $x_1, \ldots, x_n$. For more discussion about this, see Appendix B.

Zadeh defined the Extension Principle using *min* t-norm and *max* t-conorm. The use of these operations is implicit in (2.10) and (2.9). There have been attempts to use other t-norms and t-conorms in place of *min* and *max*, respectively, e.g. [17], [5]. We will work mostly with *min* or *product* t-norm and *max* t-conorm.

The Extension Principle can be viewed as a composition of fuzzy relations [5]. Let $\tilde{R}$ be the Cartesian product $\tilde{A}_1 \times \cdots \times \tilde{A}_r$ defined as [5]

$$\tilde{A}_1 \times \cdots \times \tilde{A}_r = \int_{X_1\times\cdots\times X_r} \min\{\mu_{\tilde{A}_1}(x_1), ..., \mu_{\tilde{A}_r}(x_r)\}/(x_1, ..., x_r) \tag{2.10}$$

and let S be the ordinary relation defined by $\mu_S(x_1, ..., x_r, y) = 1$ iff $y = f(x_1, ..., x_r)$. Then, we have $B = f(\tilde{A}_1, ..., \tilde{A}_r) = \tilde{R} \circ S$, i.e., the Extension Principle appears as a particular case of the composition of fuzzy relations.

Finally, when we replace *min* in (2.9) by another t-norm, we are replacing the sup-min composition by the more general sup-$\star$ composition. □

Consider two fuzzy sets of type-2, $\tilde{\tilde{A}} \in X$ and $\tilde{\tilde{B}} \in X$. Let $\tilde{\mu}_{\tilde{\tilde{A}}}(x)$ and $\tilde{\mu}_{\tilde{\tilde{B}}}(x)$ be two fuzzy grades (fuzzy sets in $J \subseteq [0, 1]$) of these two sets, represented, for each $x \in X$, as

$$\begin{aligned} \tilde{\mu}_{\tilde{\tilde{A}}}(x) &= f_x(u_1)/u_1 + f_x(u_2)/u_2 + ... + f_x(u_m)/u_m \\ &= \sum_i f_x(u_i)/u_i, \quad ; \quad u_i \in J \end{aligned} \tag{2.11}$$

$$\begin{aligned} \tilde{\mu}_{\tilde{\tilde{B}}}(x) &= g_x(w_1)/w_1 + g_x(w_2)/w_2 + ... + g_x(w_n)/w_n \\ &= \sum_j g_x(w_j)/w_j, \quad ; \quad w_j \in J \end{aligned} \tag{2.12}$$

Observe that in (2.11) and (2.12), u_i and w_j are just dummy variables used to differentiate between the different primary memberships of x in $\tilde{\tilde{A}}$ and $\tilde{\tilde{B}}$, respectively.

Using (2.9), a real binary operation $*$ can be extended to the fuzzy grades, $\tilde{\mu}_{\tilde{\tilde{A}}}(x)$ and $\tilde{\mu}_{\tilde{\tilde{B}}}(x)$, as

$$\tilde{\mu}_{\tilde{\tilde{A}}}(x) * \tilde{\mu}_{\tilde{\tilde{B}}}(x) = \sum_{i,j} \Big(f_x(u_i) \wedge g_x(w_j)\Big)/(u_i * w_j) \tag{2.13}$$

where $\wedge$ indicates *min*. Note that, if we use a t-norm other than *min*, we replace the $\wedge$ in (2.13) with the chosen t-norm [see (B.5)].

Using the Extension Principle, the membership grades for union, intersection and negation of type-2 fuzzy sets $\tilde{\tilde{A}}$ and $\tilde{\tilde{B}}$ can be defined as follows [17] :

Union

$$\begin{aligned} \tilde{\tilde{A}} \cup \tilde{\tilde{B}} \Leftrightarrow \tilde{\mu}_{\tilde{\tilde{A}} \cup \tilde{\tilde{B}}}(x) &= \tilde{\mu}_{\tilde{\tilde{A}}}(x) \sqcup \tilde{\mu}_{\tilde{\tilde{B}}}(x) \quad ; \quad x \in X \\ &= \Big(\sum_i f_x(u_i)/u_i\Big) \sqcup \Big(\sum_j g_x(w_j)/w_j\Big) \\ &= \sum_{i,j} \Big(f_x(u_i) \wedge g_x(w_j)\Big)/(u_i \vee w_j) \end{aligned} \tag{2.14}$$

Intersection

$$\begin{aligned} \tilde{\tilde{A}} \cap \tilde{\tilde{B}} \Leftrightarrow \tilde{\mu}_{\tilde{\tilde{A}} \cap \tilde{\tilde{B}}}(x) &= \tilde{\mu}_{\tilde{\tilde{A}}}(x) \sqcap \tilde{\mu}_{\tilde{\tilde{B}}}(x) \quad ; \quad x \in X \\ &= \Big(\sum_i f_x(u_i)/u_i\Big) \sqcap \Big(\sum_j g_x(w_j)/w_j\Big) \\ &= \sum_{i,j} \Big(f_x(u_i) \wedge g_x(w_j)\Big)/(u_i \wedge w_j) \end{aligned} \tag{2.15}$$

Complement

$$\begin{aligned} \bar{\tilde{\tilde{A}}} \Leftrightarrow \tilde{\mu}_{\bar{\tilde{\tilde{A}}}}(x) &= \neg \tilde{\mu}_{\tilde{\tilde{A}}}(x) \quad ; \quad x \in X \\ &= \sum_i f_x(u_i)/(1 - u_i) \end{aligned} \tag{2.16}$$

where $\vee$ and $\wedge$ represent *max* and *min* respectively. As in [17], in the sequel, we refer to the operations $\sqcup$, $\sqcap$ and $\neg$ as *join*, *meet* and *negation*, respectively. In the continuous case, we use a notation similar to that in (2.2) and get expressions similar to those in (2.14), (2.15) and (2.16).

In order to compute the <u>union</u> (or <u>intersection</u>) of $\tilde{\tilde{A}}$ and $\tilde{\tilde{B}}$, we perform the <u>join</u> (or <u>meet</u>) operation between the membership grades of $\tilde{\tilde{A}}$ and $\tilde{\tilde{B}}$ at every domain point $x \in X$; and, in order to compute the <u>complement</u> of $\tilde{\tilde{A}}$ (or $\tilde{\tilde{B}}$), we perform the <u>negation</u> operation on the membership grade of $\tilde{\tilde{A}}$ (or $\tilde{\tilde{B}}$) at every $x \in X$.

It can be easily shown that these extended operations reduce to the original ones when we deal with type-1 sets. In case of type-1 sets, $f_x(u_i)$ $[g_x(w_j)]$ will have a value equal to 1 for only one of the indices, say i_1 (j_1); the rest of $f_x(u_i)$s and $g_x(w_j)$s will all be zero (since the membership grades are not fuzzy). Consider the *join* operation. When we find the minima between all the $f_x(u_i)$s and $g_x(w_j)$s, the only pair that will give a non-zero answer is $\{f_x(u_{i_1}),\, g_x(w_{j_1})\}$, and their minimum value will be equal to 1. All other minima will be 0. So, the union of the two sets will consist of only one element $u_{i_1} \vee w_{j_1}$ or $\max\{u_{i_1}, w_{j_1}\}$, which is what we would expect. The same applies to the *meet* operation. The *negation* is even easier to see. If u_{i_1} has a membership of 1 in $\tilde{\mu}_{\tilde{\tilde{A}}}(x)$ (the rest of the memberships being zero), $1 - u_{i_1}$ will have a membership of 1 in $\tilde{\mu}_{\bar{\tilde{\tilde{A}}}}(x)$.

Example 2.1 Consider two type-2 sets $\tilde{\tilde{\mathrm{A}}}$ and $\tilde{\tilde{\mathrm{B}}}$, and, for a particular element x, let the membership grades in these two sets be given as

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x) &= 0.5/0 + 0.7/0.1 \\
\tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}}(x) &= 0.3/0.4 + 0.9/0.8
\end{aligned}$$

Then, from (2.14), we have

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}\cup\tilde{\tilde{\mathrm{B}}}}(x) &= \tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x) \sqcup \tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}}(x) \\
&= (0.5/0 + 0.7/0.1) \sqcup (0.3/0.4 + 0.9/0.8) \\
&= \frac{0.5 \wedge 0.3}{0 \vee 0.4} + \frac{0.5 \wedge 0.9}{0 \vee 0.8} + \frac{0.7 \wedge 0.3}{0.1 \vee 0.4} + \frac{0.7 \wedge 0.9}{0.1 \vee 0.8} \\
&= 0.3/0.4 + 0.5/0.8 + 0.3/0.4 + 0.7/0.8 \\
&= \max\{0.3, 0.3\}/0.4 + \max\{0.5, 0.7\}/0.8 \\
&= 0.3/0.4 + 0.7/0.8
\end{aligned}$$

Additionally, from (2.15), we have

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}\cap\tilde{\tilde{\mathrm{B}}}}(x) &= \tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x) \sqcap \tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}}(x) \\
&= (0.5/0 + 0.7/0.1) \sqcap (0.3/0.4 + 0.9/0.8) \\
&= \frac{0.5 \wedge 0.3}{0 \wedge 0.4} + \frac{0.5 \wedge 0.9}{0 \wedge 0.8} + \frac{0.7 \wedge 0.3}{0.1 \wedge 0.4} + \frac{0.7 \wedge 0.9}{0.1 \wedge 0.8} \\
&= 0.3/0 + 0.5/0 + 0.3/0.1 + 0.7/0.1 \\
&= \max\{0.3, 0.5\}/0 + \max\{0.3, 0.7\}/0.1 \\
&= 0.5/0 + 0.7/0.1
\end{aligned}$$

Finally, from (2.16), we have

$$\begin{aligned}
\tilde{\mu}_{\overline{\tilde{\tilde{\mathrm{A}}}}}(x) &= \neg\tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x) \\
&= 0.5/(1-0) + 0.7/(1-0.1) \\
&= 0.5/1 + 0.7/0.9
\end{aligned}$$

□

Algebraic product is another popular t-norm operation, especially in engineering applications [15]. The union and intersection of type-2 fuzzy sets under *product* t-norm and *max* t-conorm can be defined as follows :

Union

$$\begin{aligned}
\tilde{\tilde{\mathrm{A}}} \cup \tilde{\tilde{\mathrm{B}}} \Leftrightarrow \tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}\cup\tilde{\tilde{\mathrm{B}}}}(x) &= \tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x) \sqcup \tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}}(x) \\
&= \Big(\sum_i f_x(u_i)/u_i\Big) \sqcup \Big(\sum_j g_x(w_j)/w_j\Big) \\
&= \sum_{i,j} \Big(f_x(u_i) g_x(w_j)\Big)/(u_i \vee w_j) \qquad (2.17)
\end{aligned}$$

Intersection

$$\tilde{\tilde{\mathrm{A}}} \cap \tilde{\tilde{\mathrm{B}}} \Leftrightarrow \tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}\cap\tilde{\tilde{\mathrm{B}}}}(x) = \tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x) \sqcap \tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}}(x)$$

$$= \Big(\sum_i f_x(u_i)/u_i\Big) \sqcap \Big(\sum_j g_x(w_j)/w_j\Big)$$
$$= \sum_{i,j} \Big(f_x(u_i) g_x(w_j)\Big)/(u_i w_j) \quad (2.18)$$

Observe that, we use the same symbols for *join* and *meet* operations, as we used in case of the *min* t-norm. The definition of complement does not change.

Next, we take a closer look at the operations of *join, meet* and *negation*, under both *min* and *product* t-norms.

2.2 A Closer Look at Type-2 Set Theoretic Operations

From Section 2.1, we see that the membership grade of any point in the union or intersection of two type-2 fuzzy sets is obtained by the *join* or *meet* of the membership grades of that point, respectively. Now, we look more closely at these two operations. Most of the discussion below concerns the *join* or *meet* of two sets at a time; however, we also state generalized versions of our results when more than two sets are involved in the *join* or *meet* operations.

We will generally deal with *real fuzzy sets*, i.e., fuzzy subsets of the real line, which are *convex* and *normal*. Such sets are also known as *fuzzy numbers* [5, 10]; therefore, sometimes we will use the terms fuzzy sets and fuzzy numbers interchangeably.

2.2.1 Join and Meet under Minimum t-norm

As Theorem 2.1 below illustrates, *join* and *meet* operations with *min* t-norm give particularly simple results, if the participating type-1 sets are convex and normal.

In part (a) of Theorem 2.1, we talk about type-1 fuzzy sets $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$ having membership functions f and g. To connect this with our earlier discussion, we can think of two type-2 fuzzy sets $\tilde{\tilde{\mathrm{A}}}$ and $\tilde{\tilde{\mathrm{B}}}$ as in (2.11) and (2.12). Then, for an arbitrary input x_0, if we rename $\tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x_0)$ as $\tilde{\mathrm{F}}$ and $\tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}}(x_0)$ as $\tilde{\mathrm{G}}$, and also drop the subscript x_0 on the membership functions f_{x_0} and g_{x_0}, we can apply Theorem 2.1 to compute $\tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}\cup\tilde{\tilde{\mathrm{B}}}}(x_0)$ and $\tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}\cap\tilde{\tilde{\mathrm{B}}}}(x_0)$. Part (b) of Theorem 2.1 generalizes the results in part (a) to the *join/meet* of more than two sets.

Theorem 2.1 *(a) Suppose that we have two convex, normal, type-1 real fuzzy sets $\tilde{F}$ and $\tilde{G}$ characterized by membership functions f and g, respectively. Let $v_0 \in \Re$ and $v_1 \in \Re$ be such that $v_0 \leq v_1$ and $f(v_0) = g(v_1) = 1$. Then the membership functions of the join and meet of $\tilde{F}$ and $\tilde{G}$, using max t-conorm and min t-norm, can be expressed as*

$$\mu_{\tilde{\mathrm{F}}\sqcup\tilde{\mathrm{G}}}(\theta) = \begin{cases} f(\theta) \wedge g(\theta) & ; \quad \theta < v_0 \\ g(\theta) & ; \quad v_0 \leq \theta \leq v_1 \\ f(\theta) \vee g(\theta) & ; \quad \theta > v_1 \end{cases} \quad (2.19)$$

and

$$\mu_{\tilde{\mathrm{F}}\sqcap\tilde{\mathrm{G}}}(\theta) = \begin{cases} f(\theta) \vee g(\theta) & ; \quad \theta < v_0 \\ f(\theta) & ; \quad v_0 \leq \theta \leq v_1 \\ f(\theta) \wedge g(\theta) & ; \quad \theta > v_1 \end{cases} \quad (2.20)$$

(b) Suppose that we have n convex, normal, type-1 real fuzzy sets $\tilde{F}_1, \ldots, \tilde{F}_n$ characterized by membership functions $f_1, \ldots, f_n$, respectively. Let $v_1, v_2, \ldots, v_n$ be real numbers such that $v_1 \leq v_2 \leq \cdots \leq v_n$ and $f_1(v_1) = f_2(v_2) = \cdots = f_n(v_n) = 1$. Then, using max t-conorm and min t-norm,

$$\mu_{\sqcup_{i=1}^n \tilde{\mathrm{F}}_i}(\theta) = \begin{cases} \bigwedge_{i=1}^n f_i(\theta) & ; \quad \theta < v_1 \\ \bigwedge_{i=k+1}^n f_i(\theta) & ; \quad v_k \leq \theta \leq v_{k+1} \quad ; \quad 1 \leq k \leq n-1 \\ \bigvee_{i=1}^n f_i(\theta) & ; \quad \theta > v_n \end{cases} \quad (2.21)$$

and

$$\mu_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta) = \begin{cases} \bigvee_{i=1}^{n} f_i(\theta) & ; \quad \theta < v_1 \\ \bigwedge_{i=1}^{k} f_i(\theta) & ; \quad v_k \le \theta \le v_{k+1} \quad ; \quad 1 \le k \le n-1 \\ \bigwedge_{i=1}^{n} f_i(\theta) & ; \quad \theta > v_n \end{cases} \tag{2.22}$$

□

See Appendix C.1 for the proof of Theorem 2.1.
NOTE : Dubois and Prade present the same result given in part (a) of Theorem 2.1 in a different context and a different manner in [3]. They present it in the context of fuzzification of *max* and *min* operations. Though their method of proof is very similar to ours, they prove the result for a special case, where f and g have at most three points of intersection and one needs to keep track of the points of intersection of f and g to use their theorem. We reprove this theorem in a general setting in Appendix C.1. We believe that our statements of $\mu_{\tilde{F}\sqcup\tilde{G}}(\theta)$ and $\mu_{\tilde{F}\sqcap\tilde{G}}(\theta)$ are more amenable to computer implementations than those of Dubois and Prade. Generalization to more than two sets [part (b) of Theorem 2.1] is also difficult in case of Dubois and Prade's result.

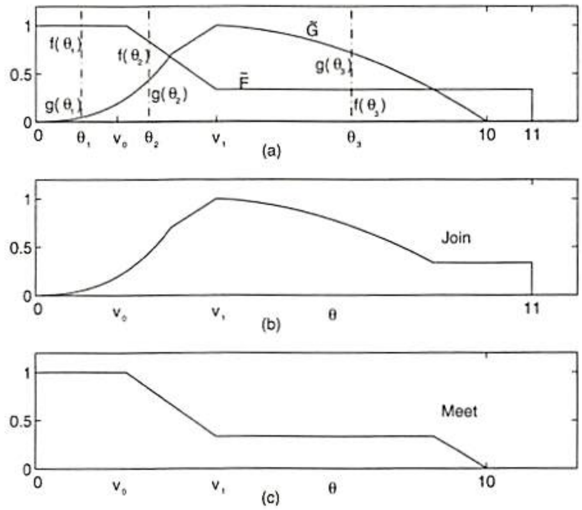


Figure 2.1: An example of two general membership functions, f and g, that satisfy the requirements of Theorem 2.1, part (a). Observe that for set $\tilde{F}$, any of the points at which f attains its maximum value of unity may be chosen as v_0. We arbitrarily chose $v_0 = 1.8$. (a) The three possibilities : $\theta_1 < v_0$, $v_0 \le \theta_2 < v_1$, $\theta_3 > v_1$. (b) Result of the *join* operation. (c) Result of the *meet* operation. The t-norm used is *min*.

Figures 2.1 and 2.2 show examples of application of Theorem 2.1. As a consequence of Theorem 2.1, we have the following important result :

Corollary 2.1 *(a) If $f(\theta)$ is the membership function of a convex, normal type-1 real fuzzy set $\tilde{F}$, and if $\tilde{G}$ is another type-1 set with membership function $f(\theta - k)$, where k is a positive constant, then, $\tilde{F} \sqcup \tilde{G} = \tilde{G}$ and $\tilde{F} \sqcap \tilde{G} = \tilde{F}$.*
(b) If we have n convex, normal, type-1 fuzzy sets $\tilde{F}_1, \ldots, \tilde{F}_n$ characterized by membership functions $f_1, \ldots, f_n$, respectively, such that $f_i(\theta) = f_1(\theta - k_i)$, and $0 = k_1 \le k_2 \cdots \le k_n$; then $\sqcup_{i=1}^{n} \tilde{F}_i = \tilde{F}_n$ and $\sqcap_{i=1}^{n} \tilde{F}_i = \tilde{F}_1$. □

See Appendix C.3 for the proof of this corollary. Figures 2.3 and 2.4 illustrate Corollary 2.1.

It can be observed, by applying Theorem 2.1, that, if two type-1 fuzzy sets are such that their membership functions do not touch, then Corollary 2.1 also holds true, i.e., $\tilde{F} \sqcup \tilde{G} = \tilde{G}$

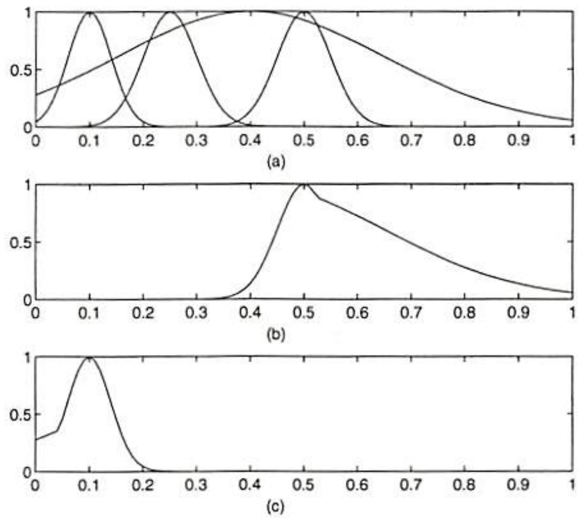


Figure 2.2: An illustration of Theorem 2.1, part (b), for Gaussians. (a) Participating Gaussians; (b) *join*; and (c) *meet*.

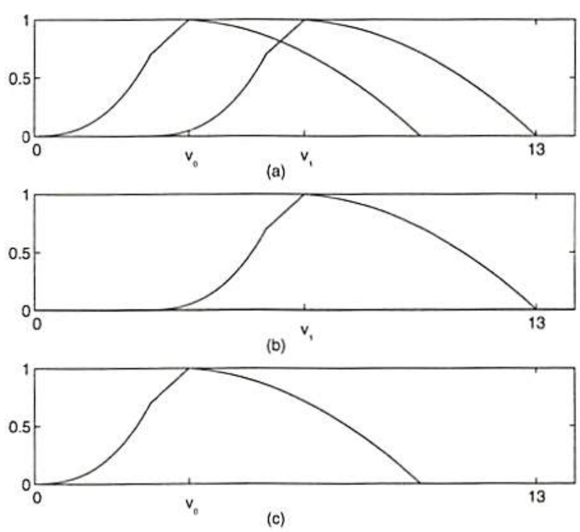


Figure 2.3: An illustration of Corollary 2.1. (a) Convex type-1 sets $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$. The membership functions of $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$ are shifted versions of each other. (b) $\tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}} = \tilde{\mathrm{G}}$. (c) $\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}} = \tilde{\mathrm{F}}$.

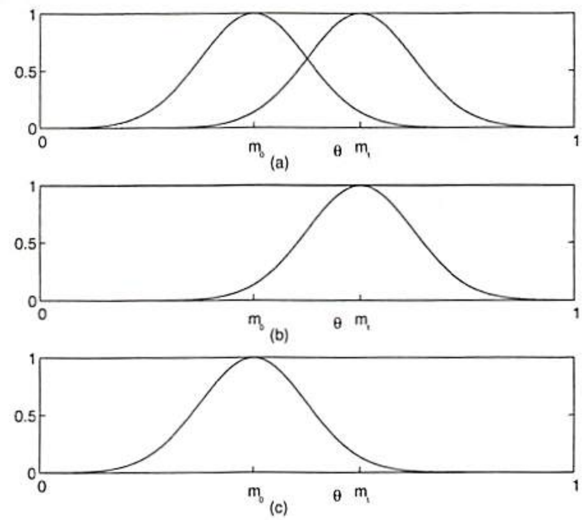


Figure 2.4: An illustration of Corollary 2.1 for the Gaussian case. (a) Participating Gaussians; (b) *join* is the Gaussian with larger mean; and (c) *meet* is the Gaussian with smaller mean.

and $\tilde{F} \sqcap \tilde{G} = \tilde{F}$. In the case of membership functions which extend infinitely (e.g., Gaussians), if the two membership functions (their centers) are far away from each other, then Corollary 2.1 approximately holds true, i.e., $\tilde{F} \sqcup \tilde{G} \approx \tilde{G}$ and $\tilde{F} \sqcap \tilde{G} \approx \tilde{F}$. Examples of this can be observed while working out the union and intersection of the Gaussian type-2 sets depicted in Figs. 2.7 (a), as explained later in this section.

Kaufmann and Gupta [10] give a result that is a bit more general than Corollary 2.1. They give this result in terms of α-cuts as follows : Consider two fuzzy numbers $\tilde{A}$ and $\tilde{B}$, such that

$$\begin{aligned} A_\alpha &= [a_1^{(\alpha)}, a_2^{(\alpha)}] \text{ and} \\ B_\alpha &= [b_1^{(\alpha)}, b_2^{(\alpha)}] \end{aligned}$$

where α-cuts, A_α and B_α are crisp sets

$$A_\alpha = \{x | \mu_{\tilde{A}}(x) \geq \alpha\}, \;\; \alpha \in [0,1]$$
$$B_\alpha = \{x | \mu_{\tilde{B}}(x) \geq \alpha\}, \;\; \alpha \in [0,1]$$

If $\forall \alpha \in [0,1]$, $a_1^{(\alpha)} \leq b_1^{(\alpha)}$ and $a_2^{(\alpha)} \leq b_2^{(\alpha)}$, then $\tilde{A} \leq \tilde{B}$.

In our work, we frequently deal with normalized Gaussian membership functions. Corollary 2.1 takes a particularly simple form when the sets involved are Gaussians. For two Gaussians having the *same standard deviation*, the result of the join operation between them is the Gaussian with the larger mean, and the result of the meet operation is the Gaussian with the smaller mean. Gaussians having *different standard deviations* cannot be expressed as shifted versions of each other and hence Corollary 2.1 does not apply to them. Theorem 2.1 can, of course, be used in this case. Figures 2.2 and 2.5 show examples of the *join* and *meet* operations between Gaussians having different standard deviations, under the *min* t-norm.

Similar results hold for triangular, trapezoidal, or, for that matter, any other convex membership function.

From the definition of the *negation* operation, it follows that :

Theorem 2.2 *If a type-1 fuzzy set* $\tilde{F}$ *has a membership function* $f(\theta)$ $(\theta \in \Re)$, $\neg\tilde{F}$ *has a membership function* $f(1-\theta)$ $(\theta \in \Re)$. □

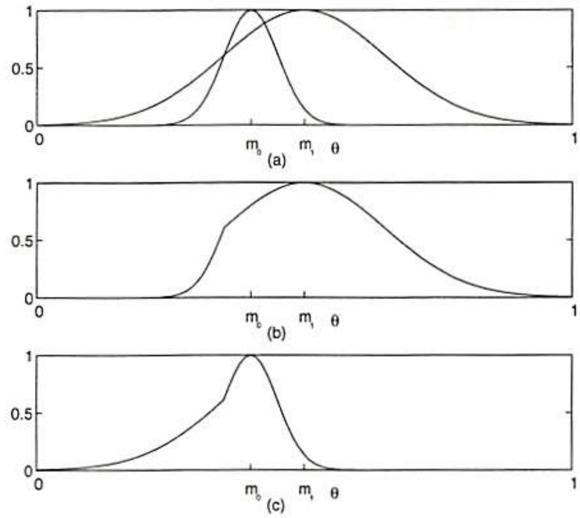


Figure 2.5: *Join* and *meet* operations between Gaussians under *min* t-norm. (a) Participating Gaussians; (b) *join*; and (c) *meet*.

See Appendix C.4 for the proof. Figure 2.6 shows an example of the *negation* operation.

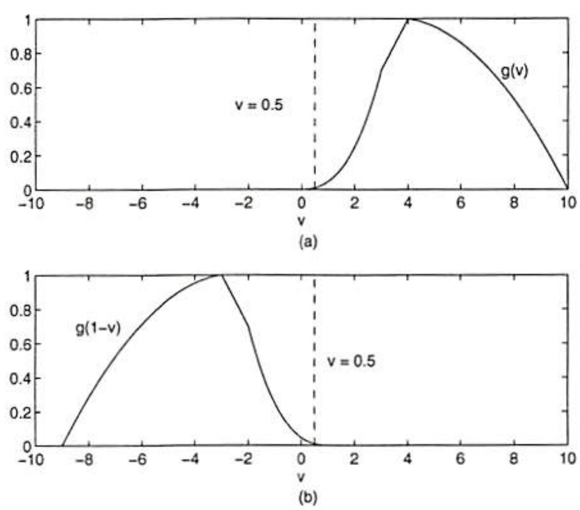


Figure 2.6: An illustration of the *negation* operation. (a) Type-1 fuzzy set $\tilde{G}$, (b) $\neg\tilde{G}$.

So, we perform *join*, *meet* and *negation* operations on membership grades of type-2 sets while finding unions, intersections and complements of type-2 sets. Having seen how the results of these individual operations look, let's see how the overall type-2 set looks as a result of these operations. Figures 2.7 and 2.8 show examples of union, intersection and complement of Gaussian fuzzy sets using the 2-D representation introduced in Section 1.4 (see Fig. 1.10). In Fig. 2.7 (a), if we draw a vertical line at any x, we get the membership grades of x in the two participating Gaussian type-2 sets. These membership grades are, of course, themselves Gaussian fuzzy sets confined to the interval $[0, 1]$. To these two type-1 sets, we apply Theorem 2.1 and get the results for union and intersection depicted in Figs. (b) and (c) respectively. Of course, while applying the theorem, we should be careful to see that $\tilde{F}$ and $\tilde{G}$ do satisfy all its requirements. Similarly, in Fig. 2.8 (a), we project upwards from x to obtain it's membership grade and then apply Theorem 2.2 to it.

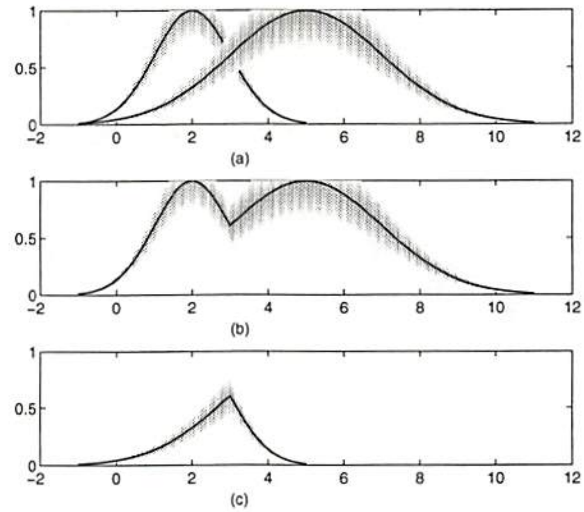


Figure 2.7: Union and Intersection of Gaussian type-2 fuzzy sets using the 2-D pictorial representation introduced in Fig. 1.10. (a) Participating sets; (b) union ; and (c) intersection.

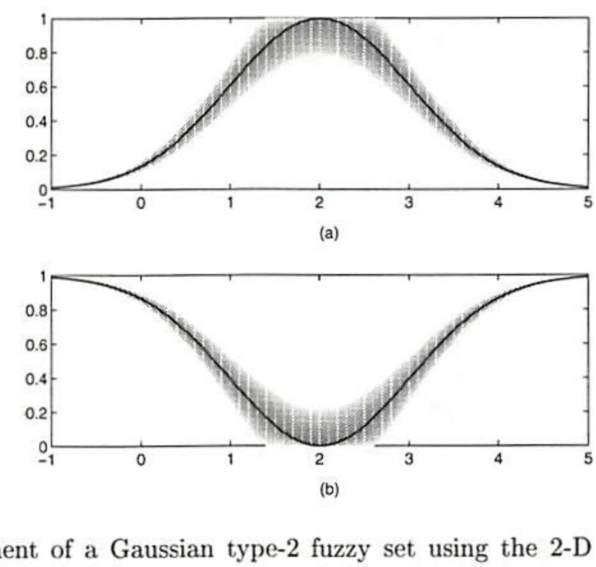


Figure 2.8: Complement of a Gaussian type-2 fuzzy set using the 2-D pictorial representation introduced in Fig. 1.10. (a) Gaussian type-2 set; (b)complement.

Observe that, in Figs. 2.7 and 2.8, if we look just at the *principal* membership functions, we can see that *the principal membership function of the result of an operation (union, intersection or complement) can be obtained by performing that operation on the principal membership functions of the participating type-2 sets.* So, if we replace all the type-2 sets by type-1 sets, which have the principal membership functions of the type-2 sets as their type-1 membership functions, all our results remain valid. This demonstrates the fact that all our type-2 operations collapse to the correct type-1 operations.

Theorem 2.1 considers the *join* and *meet* operations under *min* t-norm. Now, we examine these operations under the *product* t-norm, the t-conorm being *max*. This case was not considered by Dubois and Prade in [3].

2.2.2 Join under Product t-norm

The *join* operation with *product* t-norm gives a result very similar to that in Theorem 2.1. Consider the two convex normal type-1 fuzzy sets $\tilde{F}$ and $\tilde{G}$ used in Theorem 2.1. The membership function of the *join* of $\tilde{F}$ and $\tilde{G}$ using the *max* t-conorm and *product* t-norm, can be expressed as

$$\mu_{\tilde{F}\sqcup\tilde{G}}(\theta) = \begin{cases} f(\theta)g(\theta) & ; \quad \theta < v_0 \\ g(\theta) & ; \quad v_0 \leq \theta \leq v_1 \\ f(\theta) \vee g(\theta) & ; \quad \theta > v_1 \end{cases} \tag{2.23}$$

Figure 2.9 (b) shows an example of this operation. Comparing Figs. 2.9 (b) and 2.1 (b), we see that results of the two *join* operations (with *min* as well as *product* t-norm are exactly the same. This can be explained as follows. From (2.19) and (2.23), we see that the two results can differ only for $\theta < v_0$. In this range, the *min* t-norm gives $f(\theta) \wedge g(\theta)$ and *product* t-norm gives $f(\theta)g(\theta)$. In the example we have chosen, $f(\theta) = 1$ for $\theta < v_0$; therefore, for our example, these two results turn out the same. Generalization to more than two sets is also very similar to that in (2.21). It can be obtained by replacing the minima in (2.21) with products as follows :

$$\mu_{\sqcup_{i=1}^{n}\tilde{F}_i}(\theta) = \begin{cases} \prod_{i=1}^{n} f_i(\theta) & ; \quad \theta < v_1 \\ \prod_{i=k+1}^{n} f_i(\theta) & ; \quad v_k \leq \theta \leq v_{k+1} \quad ; \quad 1 \leq k \leq n-1 \\ \bigvee_{i=1}^{n} f_i(\theta) & ; \quad \theta > v_n \end{cases} \tag{2.24}$$

where we take $\prod_{i=n}^{n} f_i(\theta)$ to mean f_n. See Appendix C.5 for the proofs of (2.23) and (2.24).

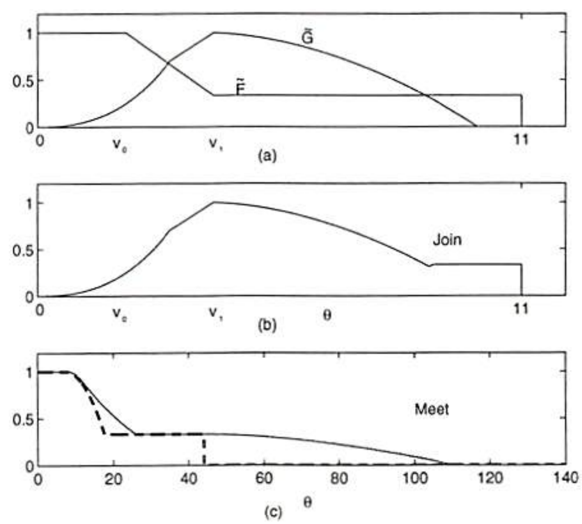


Figure 2.9: *Join* and *meet* operations under *product* t-norm. (a) Participating type-1 fuzzy sets; (b) *join* ; (c) *meet* : the actual result is shown with the thin solid line and the approximation in (2.41) with the thick dashed line.

Observe that (2.23) is very similar to (2.19). In fact, the information in both these pairs of equations can be conveyed as follows : For two type-1 sets $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$ described in part (a) of Theorem 2.1,

$$\mu_{\tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}}}(\theta) = \begin{cases} f(\theta) \star g(\theta) & ; \quad \theta < v_0 \\ g(\theta) & ; \quad v_0 \leq \theta \leq v_1 \\ f(\theta) \vee g(\theta) & ; \quad \theta > v_1 \end{cases} \tag{2.25}$$

where $\star$ denotes the t-norm operation, which corresponds to *min* in (2.19) and *product* in (2.23).

Similarly comparing (2.21) and (2.24), for n type-1 sets $\tilde{\mathrm{F}}_1, \ldots, \tilde{\mathrm{F}}_n$ described in part (b) of Theorem 2.1,

$$\mu_{\sqcup_{i=1}^n \tilde{\mathrm{F}}_i}(\theta) = \begin{cases} \mathcal{T}_{i=1}^n f_i(\theta) & ; \quad \theta < v_1 \\ \mathcal{T}_{i=k+1}^n f_i(\theta) & ; \quad v_k \leq \theta \leq v_{k+1} \quad ; \quad 1 \leq k \leq n-1 \\ \bigvee_{i=1}^n f_i(\theta) & ; \quad \theta > v_n \end{cases} \tag{2.26}$$

where $\mathcal{T}$ indicates the t-norm used, min or product.

We state this result formally as :

Theorem 2.3 *(a) Suppose that we have two convex, normal, type-1 real fuzzy sets $\tilde{F}$ and $\tilde{G}$ characterized by membership functions f and g, respectively. Let $v_0 \in \Re$ and $v_1 \in \Re$ be such that $v_0 \leq v_1$ and $f(v_0) = g(v_1) = 1$. Then the membership functions of the join of $\tilde{F}$ and $\tilde{G}$, using max t-conorm, can be expressed as*

$$\mu_{\tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}}}(\theta) = \begin{cases} f(\theta) \star g(\theta) & ; \quad \theta < v_0 \\ g(\theta) & ; \quad v_0 \leq \theta \leq v_1 \\ f(\theta) \vee g(\theta) & ; \quad \theta > v_1 \end{cases} \tag{2.27}$$

where $\star$ denotes the t-norm operation used, min or product.

(b) Suppose that we have n convex, normal, type-1 real fuzzy sets $\tilde{F}_1, \ldots, \tilde{F}_n$ characterized by membership functions $f_1, \ldots, f_n$, respectively. Let $v_1, v_2, \ldots, v_n$ be real numbers such that $v_1 \leq v_2 \leq \cdots \leq v_n$ and $f_1(v_1) = f_2(v_2) = \cdots = f_n(v_n) = 1$. Then, the membership function of $\sqcup_{i=1}^n \tilde{F}_i$ using max t-conorm, can be expressed as

$$\mu_{\sqcup_{i=1}^n \tilde{\mathrm{F}}_i}(\theta) = \begin{cases} \mathcal{T}_{i=1}^n f_i(\theta) & ; \quad \theta < v_1 \\ \mathcal{T}_{i=k+1}^n f_i(\theta) & ; \quad v_k \leq \theta \leq v_{k+1} \quad ; \quad 1 \leq k \leq n-1 \\ \bigvee_{i=1}^n f_i(\theta) & ; \quad \theta > v_n \end{cases} \tag{2.28}$$

where $\mathcal{T}$ indicates the t-norm used, min or product. □

Figures 2.10, 2.11 (b) and 2.12 (b) show results of *join* operations on Gaussians under *product* t-norm. Note that Corollary 2.1 is not valid under *product* t-norm.

Example 2.2 In this example, we illustrate the use of Theorem 2.3, when $F_1, \ldots, F_n$ are interval type-1 sets. (We drop the tilde, since the sets are crisp.) Let the domains of $F_1, \ldots, F_n$ be the intervals $[l_1, r_1], \ldots, [l_n, r_n]$, respectively. The membership functions for these interval sets can be expressed as

$$\mu_{\tilde{\mathrm{F}}_i}(\theta) = \begin{cases} 1 & ; \quad \theta \in [l_i, r_i] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{2.29}$$

Without loss of generality, let us also assume that $l_1 \leq \cdots \leq l_n$. Since all the memberships in interval type-1 sets are unity, for each F_i ($i = 1, \ldots, n$), any domain point can be chosen as v_i (see Theorem 2.3). Let us choose $v_i = l_i$ for $i = 1, \ldots, n$. To use (2.28), observe, from (2.29), that

1. $f_i(\theta) = 0$ for $\theta < v_1 = l_1$;
2. when $\theta \in [v_k, v_{k+1})$, $f_i(\theta) = 0$ for $i = k+1, \ldots, n$;

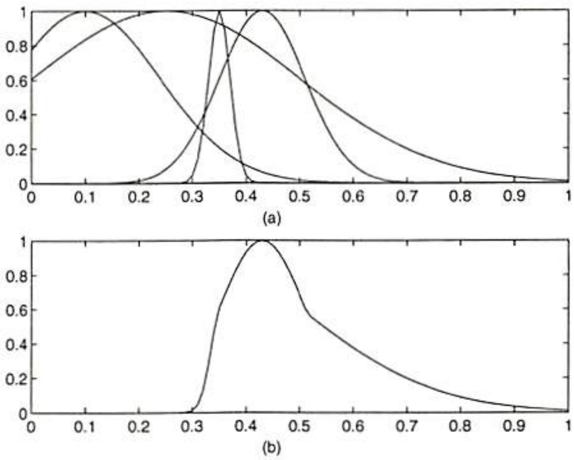


Figure 2.10: An illustration of (2.24) for the Gaussian case. (a) Participating Gaussians; (b) *join* under product t-norm.

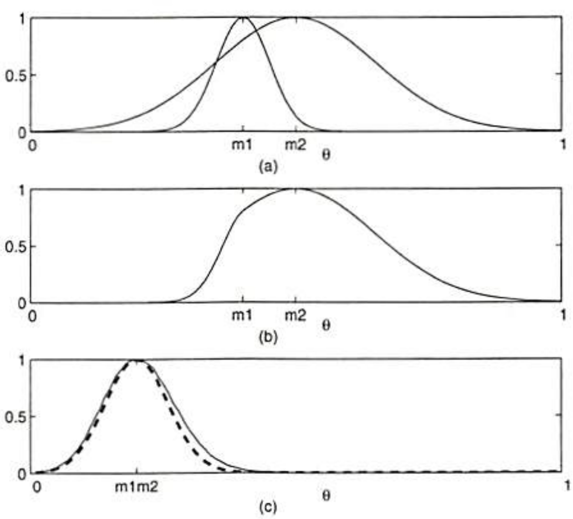


Figure 2.11: *Join* and *meet* operations between Gaussians under *product* t-norm. (a) Participating Gaussians; (b) *join*; and (c) *meet* : the thin solid line depicts the actual result and the thick dashed line shows the approximation in (2.46). Compare these results with those in Fig. 2.5 obtained using the *min* t-norm.

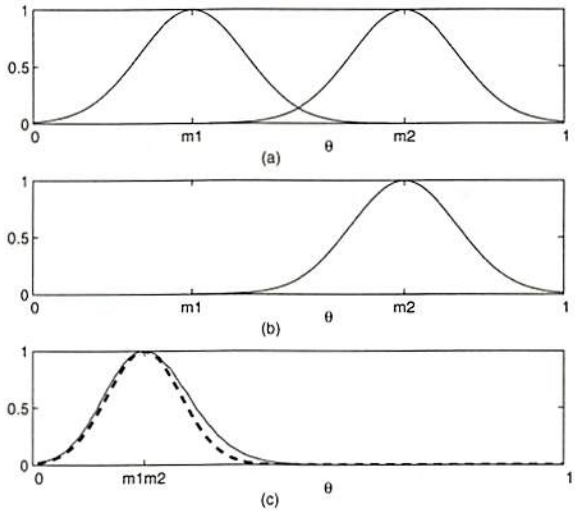


Figure 2.12: *Join* and *meet* operations between Gaussians, having *the same standard deviation*, under *product* t-norm. (a) Participating Gaussians; (b) *join*; and (c) *meet* : the thin solid line depicts the actual result and the thick dashed line shows the approximation in (2.46).

3. for $v_n = l_n \leq \theta \leq \vee_{i=1}^{n} r_i$, $\vee_{i=1}^{n} f_i(\theta) = 1$; and,
4. for $\theta > \vee_{i=1}^{n} r_i$, $\vee_{i=1}^{n} f_i(\theta) = 0$.

Using these observations with (2.28), we can see that the $\sqcup_{i=1}^{n} \tilde{F}_i$ is an interval type-1 set with domain $[l_n, \vee_{i=1}^{n} r_i]$. More generally, $\sqcup_{i=1}^{n} \tilde{F}_i$ is an interval type-1 set with domain $[\vee_{i=1}^{n} l_i, \vee_{i=1}^{n} r_i]$ (recall that we assumed that $l_1 \leq \cdots \leq l_n$).

Note that the result in this example is true for *any* t-norm, since the membership grades in interval type-1 sets are just 0 or 1. □

2.2.3 Meet under Product t-norm

The *meet* operation between $\tilde{F}$ and $\tilde{G}$ (convex, normal, type-1 fuzzy sets used in Theorem 2.1), under the *product* t-norm can be represented as

$$\tilde{F} \sqcap \tilde{G} = \int_{v \in \Re} \int_{w \in \Re} [f(v)g(w)]/(vw) \tag{2.30}$$

Observe that this equation involves the product of primary memberships v and w rather than a *min* or *max* operation between them; hence, the analysis of the *meet* operation under *product* t-norm is quite different than that of *join* or *meet* operations previously discussed.

Equation (2.30) simplifies considerably when $\tilde{F}$ and $\tilde{G}$ are interval type-1 sets, as we show with an example next. (Recall, from Chapter 1, that interval type-1 sets are crisp sets whose domains are intervals on the real line.)

Example 2.3 Let F and G be two interval type-1 sets with domains $[l_f, r_f]$ and $[l_g, r_g]$, respectively. Using (2.30), the *meet* between F and G, under product t-norm, can be obtained as

$$F \sqcap G = \int_{v \in F} \int_{w \in G} (1 \times 1)/(vw) \tag{2.31}$$

Observe, from (2.31), that

- each term in $F \sqcap G$ is equal to the product vw for some $v \in F$ and $w \in G$, the smallest term being $l_f l_g$ and the largest $r_f r_g$; and,
- since both F and G have continuous domains, $F \sqcap G$ also has a continuous domain;

consequently, $F \sqcap G$ is an interval type-1 set with domain $[l_f l_g, r_f r_g]$, i.e.,

$$F \sqcap G = \int_{v \in [l_f l_g, r_f r_g]} 1/v \tag{2.32}$$

In a similar manner, the *meet*, $\sqcap_{i=1}^n F_i$, of n interval type-1 sets $F_1, \ldots, F_n$, having domains $[l_1, r_1], \ldots, [l_n, r_n]$, respectively, is an interval set with domain $\left[\prod_{i=1}^n l_i, \prod_{i=1}^n r_i\right]$.

[10] gives a similar result while discussing multiplication of fuzzy numbers (see Section 2.4 for algebraic operations on fuzzy sets). □

If the sets involved in the *meet* operation are not interval type-1 sets, generally a direct application of (2.30) does not give such a nice result. We, then, analyze this operation as follows. If θ is an element of $\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}}$, then the membership grade of θ can be found by finding all the pairs $\{v, w\}$ such that $v \in \tilde{\mathrm{F}}$, $w \in \tilde{\mathrm{G}}$ and $vw = \theta$; multiplying the membership grades of v and w in each pair; and then finding the maximum of these products of membership grades. The possible admissible $\{v, w\}$ pairs whose product is θ are $\{v, \theta/v\}$ ($v \in \Re$, $v \neq 0$) for $\theta \neq 0$ and $\{v, 0\}$ or $\{0, w\}$, where $v, w \in \Re$ for $\theta = 0$. We find the products of membership grades of v and w from each such pair and take the maximum of all these products as the membership grade of θ, i.e.,

$$\begin{aligned} \mu_{\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}}}(\theta) &= \sup_{v \in \Re, v \neq 0} f(v) g\left(\frac{\theta}{v}\right) ; \; \theta \in \Re, \; \theta \neq 0 \\ \mu_{\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}}}(0) &= [\sup_{v \in \Re} f(v) g(0)] \vee [\sup_{w \in \Re} f(0) g(w)] \end{aligned} \tag{2.33}$$

Observe that

$$\begin{aligned} \sup_{v \in \Re} f(v) g(0) &= g(0) \sup_{v \in \Re} f(v) \\ &= g(0) \times 1 \\ &= g(0) \end{aligned} \tag{2.34}$$

and similarly,

$$\sup_{w \in \Re} f(0) g(w) = f(0) \, ; \tag{2.35}$$

therefore, summarizing the above discussion, we have that for two convex, normal, type-1 fuzzy sets $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$ (satisfying conditions of Theorem 2.1), the membership function of the *meet* under *product* t-norm can be expressed as

$$\begin{aligned} \mu_{\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}}}(\theta) &= \sup_{v \in \Re, v \neq 0} f(v) g\left(\frac{\theta}{v}\right) \; ; \; \theta \neq 0 \\ \mu_{\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}}}(0) &= f(0) \vee g(0) \end{aligned} \tag{2.36}$$

If we substitute $\theta/v = w$ in (2.36), we get a similar expression in terms of $f(\theta/w) g(w)$. Since the *meet* operation is commutative (see Chapter 3), we get the same result whether we substitute $\theta/w = v$ or $\theta/v = w$.

As is apparent from (2.36), the result is very much dependent on functions f and g and does not easily generalize like the *join* and *meet* operations considered earlier, and generally, it is very difficult to obtain a closed form expression for the result of the *meet* operation [which is why we have not stated (2.36) as a theorem]. Even if both the fuzzy sets involved have Gaussian membership functions, it is difficult to obtain a nice closed form expression for the result of the *meet* operation. See Appendix C.6 for more discussion about *meet* between Gaussians under product t-norm.

Figure 2.9 (c) shows an example of this operation. To determine the membership of a particular point θ in $\tilde{F} \sqcap \tilde{G}$, we find all the pairs $\{v, w\}$ such that $v \in \Re$, $w \in \Re$ and $vw = \theta$; and multiply the memberships of each pair. The membership grade of θ is given by the supremum of the set of all these products. For example, if $\theta = 20$, all the pairs $\{v, w\}$ that give 20 as their product are v and $\frac{20}{v}$ ($v \in \Re, v \neq 0$). So, the membership grade of 20 is given by the supremum of the set of all the products $f(v)g(\frac{20}{v})$ ($v \in \Re, v \neq 0$). Figure 2.13 shows how $f(v)g(\frac{20}{v})$ looks for $\tilde{F}$ and $\tilde{G}$ depicted in Fig. 2.9 (a). Clearly, it is no easy matter to represent (2.36) visually.

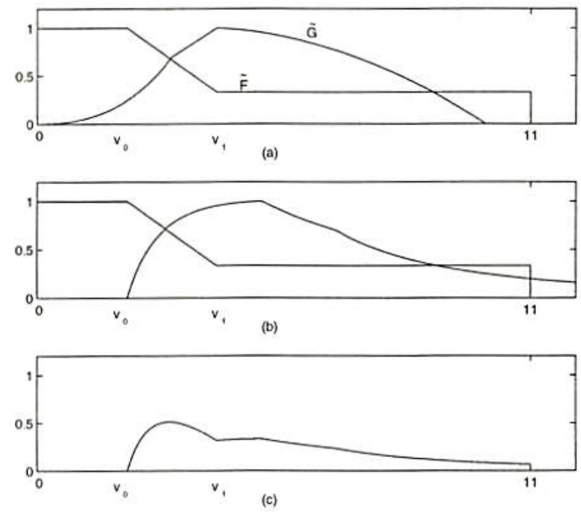


Figure 2.13: An example showing how $f(v)g(\frac{20}{v})$ looks for the curves we considered in the proof of Theorem 2.1. (a) The membership functions f and g of type-1 sets $\tilde{F}$ and $\tilde{G}$, respectively. (b) $f(v)$, which is the same as that in Fig. (a) and $g(\frac{20}{v})$. (c) The product $f(v)g(\frac{20}{v})$.

One situation when the result of the *meet* operation in (2.36) simplifies considerably is when either one of $\tilde{F}$ or $\tilde{G}$ is a fuzzy singleton. For example, assume that $\tilde{F}$ is a fuzzy singleton, such that $f(v_0) = 1$ and $f(v) = 0$ for $v \neq v_0$. Now, $f(v)g(\frac{\theta}{v})$ is non-zero only at $v = v_0$, implying that $\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) = g(\frac{\theta}{v_0})$. Similarly, if $\tilde{G}$ is a fuzzy singleton, such that $g(v_1) = 1$ and $g(w) = 0$ for $w \neq v_1$, $f(v)g(\frac{\theta}{v})$ is non-zero only at $\frac{\theta}{v} = v_1$, i.e., only when $v = \frac{\theta}{v_1}$, implying that $\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) = f(\frac{\theta}{v_1})$.

Because the *meet* under *product* t-norm will be heavily used by us in the sequel, we seek approximations to it that will make it practical.

2.3 Approximations for Meet under Product t-norm

The *meet* operation (analogous to the t-norm in the type-1 case) is the most heavily used operation in a type-2 fuzzy logic system (FLS) [see Chapter 5]. Consequently, saving computational effort in the *meet* operation means making the overall type-2 FLS operation considerably faster. In this section, we discuss some approximations to the *meet* under product t-norm that will help us make the *meet* calculations more efficient. Subsection 2.3.1 discusses an ad hoc approximation that can be used with any normal membership functions. Subsection 2.3.2 discusses a Gaussian approximation for Gaussian membership functions. A triangular approximation for symmetrical triangular membership functions is discussed in [8]. Both the Gaussian and the triangular approximations have the following two very desirable properties : 1) they can be computed very easily from the membership functions of the type-1 sets involved in the *meet* operation; and 2) they are easily generalizable to the *meet* of more than two fuzzy sets.

2.3.1 First Approximation

As explained earlier, if one of the two fuzzy sets is a fuzzy singleton, the *meet* operation simplifies a lot, e.g., if $\tilde{F}$ is a singleton, having membership equal to 1 at v_0 and zero at all other points, as explained at the end of Section 2.2.3, the result of the *meet* operation is (assuming $v_0 \neq 0$)

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) = g\left(\frac{\theta}{v_0}\right) \; ; \theta \in \Re \tag{2.37}$$

If $v_0 = 0$, $\tilde{F} = 1/0$. Then, from (2.30), we have

$$\begin{aligned} \tilde{F} \sqcap \tilde{G} &= \int_{w\in\Re} [f(0)g(w)]/0 \\ &= \int_{w\in\Re} g(w)/0 \\ &= [\sup_{w} g(w)]/0 \\ &= 1/0 \end{aligned} \tag{2.38}$$

where we have made use of the facts that the integrals in (2.30) denote union, the t-conorm used is maximum, and $\tilde{G}$ is normal.

Similarly, if $\tilde{G}$ is a singleton, having membership equal to 1 at v_1 and zero at all other points, the result of the *meet* operation is (assuming $v_1 \neq 0$)

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) = f\left(\frac{\theta}{v_1}\right) \; ; \theta \in \Re \tag{2.39}$$

This motivates the following ad hoc approximation for *meet* under *product* t-norm,

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) \approx f\left(\frac{\theta}{v_1}\right) \vee g\left(\frac{\theta}{v_0}\right) \; ; \theta \in \Re \tag{2.40}$$

This expression does not take into account the possibility that $\tilde{F}$ or $\tilde{G}$ may have more than one point with membership grade equal to 1 in their support (an example of such a fuzzy set is F in Fig. 2.1). To account for this case, we generalize (2.40) as follows :

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) \approx \int_{v\in V} f\left(\frac{\theta}{v}\right) \vee \int_{w\in W} g\left(\frac{\theta}{w}\right) \tag{2.41}$$

where V is the crisp (non-fuzzy) set of all points having a membership grade equal to 1 in $\tilde{F}$ and W is the crisp set of all points having a membership grade equal to 1 in $\tilde{G}$. Figure 2.9 (c) shows the above approximation along with the actual result. We do not claim that this approximation is optimal in any sense; however, it looks intuitively reasonable and is much easier to compute than the actual result [particularly because, as mentioned earlier, we often deal with type-1 fuzzy sets (membership grades of type-2 sets) that have only one point at which the secondary membership reaches 1, so that we can use Eq. (2.40)]. Additionally, as we show next, it collapses to the correct type-1 result if fuzzy memberships are replaced by appropriate crisp memberships, i.e., if we replace the type-2 sets by appropriate type-1 sets, the results remain valid. (Given a type-2 set, an "appropriate" type-1 set is one which has a membership function equal to the principal membership function of the type-2 set.)

Recall, that

$$\tilde{\mu}_{\tilde{\tilde{A}}\cap\tilde{\tilde{B}}}(x_0) = \tilde{\mu}_{\tilde{\tilde{A}}}(x_0) \sqcap \tilde{\mu}_{\tilde{\tilde{B}}}(x_0) \tag{2.42}$$

In our analysis, we denote $\tilde{\mu}_{\tilde{\tilde{A}}}(x_0)$ and $\tilde{\mu}_{\tilde{\tilde{B}}}(x_0)$ by $\tilde{F}$ and $\tilde{G}$, respectively. We have assumed that the membership functions of $\tilde{F}$ and $\tilde{G}$, namely f and g, are such that $f(v_0) = g(v_1) = 1$. If all the type-2 sets are replaced by the appropriate type-1 sets, $\tilde{F}$ and $\tilde{G}$ reduce to singletons $1/v_0$ and $1/v_1$, respectively; so that $\tilde{F} \sqcap \tilde{G} = 1/(v_0 v_1) = v_0 v_1$.

Now let's see how the approximation in (2.40) reduces to this result.

$$\begin{aligned} f(\theta) &= \begin{cases} 1 \; ; & \theta = v_0 \\ 0 \; ; & \text{otherwise} \end{cases} \\ \Rightarrow f(\frac{\theta}{v_1}) &= \begin{cases} 1 \; ; & \theta = v_0 v_1 \\ 0 \; ; & \text{otherwise} \end{cases} \end{aligned} \tag{2.43}$$

and

$$\begin{aligned} g(\theta) &= \begin{cases} 1 \; ; & \theta = v_1 \\ 0 \; ; & \text{otherwise} \end{cases} \\ \Rightarrow g(\frac{\theta}{v_0}) &= \begin{cases} 1 \; ; & \theta = v_0 v_1 \\ 0 \; ; & \text{otherwise} \end{cases} \end{aligned} \tag{2.44}$$

From (2.40), (2.43) and (2.44), we can see that when both $\tilde{F}$ and $\tilde{G}$ are singletons, the result of the first approximation is equal to $v_0 v_1$, which is the true result of the *meet*.

In Theorem 2.1 and our discussion about the *product* t-norm, we have considered *join* and *meet* operations between general fuzzy sets, which have the real line as their support; however, when dealing with type-2 sets, we use these operations between fuzzy membership grades, which are type-1 fuzzy sets supported in $[0, 1]$; hence, results of all the *join* and *meet* operations, for both *min* as well as *product* t-norms are again type-1 fuzzy sets supported in $[0, 1]$. Additionally we will frequently be interested in Gaussian membership functions.

Example 2.4 Let's see how the above approximation for *meet* looks in the case of Gaussian fuzzy sets. Since Gaussians reach unity height at only a single point, we can use (2.40). If f and g are Gaussians with means m_f and m_g, and standard deviations σ_f and σ_g respectively, then from (2.40), we have

$$\begin{aligned} \mu_{\tilde{F}\sqcap\tilde{G}}(\theta) &\approx e^{-\frac{1}{2}\left(\frac{\frac{\theta}{m_g} - m_f}{\sigma_f}\right)^2} \vee e^{-\frac{1}{2}\left(\frac{\frac{\theta}{m_f} - m_g}{\sigma_g}\right)^2} \; ; \theta \in [0, 1] \\ &= e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{m_g \sigma_f}\right)^2} \vee e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{m_f \sigma_g}\right)^2} \; ; \theta \in [0, 1] \end{aligned} \tag{2.45}$$

On the RHS of (2.45), we are comparing two Gaussians with equal means. Obviously, their maximum will equal the Gaussian with the larger value of $m_g \sigma_f$ or $m_f \sigma_g$; therefore,

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) \approx e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{\max\{m_f \sigma_g, m_g \sigma_f\}}\right)^2} \; ; \theta \in [0, 1] \tag{2.46}$$

So, *the approximation of* meet *between two Gaussians is a Gaussian, whose mean is equal to the product of the means of the two participating Gaussians.* Figure 2.11 (c) depicts an example of *meet* (approximation with the actual result) of Gaussians under *product* t-norm and *max* t-conorm. Figure 2.12 (c) depicts a similar result for Gaussians with equal standard deviations.

Let's see how this generalizes to more than two Gaussians at a time. Consider the meet between three Gaussians $\tilde{F}_1$, $\tilde{F}_2$, $\tilde{F}_3$ with means m_1, m_2, m_3 and standard deviations σ_1, σ_2,

σ_3. Since the *meet* operation is associative under *product* t-norm and *max* t-conorm (as we show later on in Chapter 3), we can first find the *meet* between $\tilde{F}_1$ and $\tilde{F}_2$ and then find the *meet* of the resulting function with $\tilde{F}_3$. Using (2.46), we see that the approximation of $\tilde{F}_1 \sqcap \tilde{F}_2$ is a Gaussian, say $\tilde{F}_{12}$ with mean $m_{12} = m_1 m_2$ and standard deviation $\sigma_{12} = \max\{m_1\sigma_2, m_2\sigma_1\}$. Using (2.46) again, we see that the approximation of $\tilde{F}_{12} \sqcap \tilde{F}_3$ is again a Gaussian, say $\tilde{F}_{123}$ with mean $m_{123} = m_{12}m_3 = m_1 m_2 m_3$ and standard deviation

$$\sigma_{123} = \max\{m_{12}\sigma_3, m_3\sigma_{12}\} = \max\{m_1 m_2 \sigma_3, m_1 m_3 \sigma_2, m_2 m_3 \sigma_1\} \tag{2.47}$$

The generalization of this result is straightforward. If there are n Gaussian fuzzy sets $\tilde{F}_1, \tilde{F}_2, \cdots, \tilde{F}_n$ with means $m_1, m_2, \cdots, m_n$ and standard deviations $\sigma_1, \sigma_2, \cdots, \sigma_n$, respectively, then repeated application of (2.40) yields

$$\mu_{\tilde{F}_1 \sqcap \tilde{F}_2 \sqcap \cdots \sqcap \tilde{F}_n}(\theta) \approx e^{-\frac{1}{2}\left(\frac{\theta - m_1 m_2 \cdots m_n}{\bar{\sigma}}\right)^2} \tag{2.48}$$

where

$$\bar{\sigma} = \max\left\{\sigma_1 \prod_{i\,;\, i\neq 1} m_i, \sigma_2 \prod_{i\,;\, i\neq 2} m_i, \cdots, \sigma_j \prod_{i; i\neq j} m_i, \cdots, \sigma_n \prod_{i; i\neq n} m_i\right\};\ i = 1, 2, \cdots, n \tag{2.49}$$

Figure 2.14 shows examples of *meet* of more than two Gaussians and compares the actual results with the approximations.

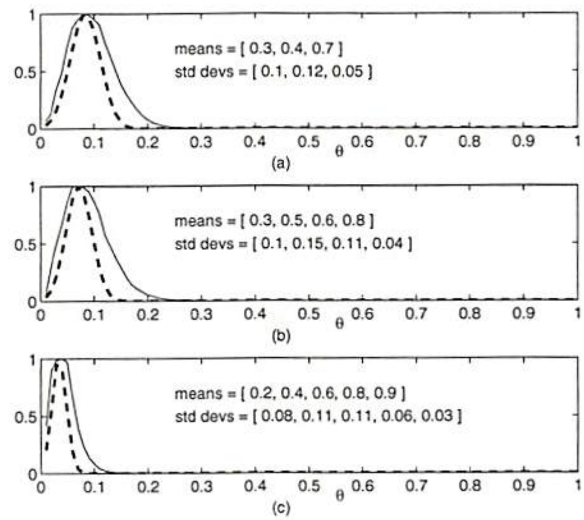


Figure 2.14: Examples of *meet* of more than two Gaussians at a time for *product* t-norm. The approximation in (2.48) is shown with the thick dashed line. The thin solid line shows the actual result, which was calculated numerically.

Note that the Gaussians are contained in $[0, 1]$ and may, therefore, be clipped (see Appendix C.8.2). In this example, we did not consider effects of clipping; however, we consider them in Appendix C.8.3 while calculating a lower bound for the Gaussian approximation derived in Section 2.3.2. The process of finding a lower bound on the Gaussian approximation is very similar to computing our first approximation to the *meet* between Gaussians. □

2.3.2 A Gaussian Approximation

The approximation in (2.46) was motivated by a general membership function, not necessarily Gaussian [see (2.40) and (2.41)]. If we focus on Gaussian fuzzy sets, we can come up with a better approximation for *meet* under the product t-norm. Observe that the *meet* operation is performed between membership grades of type-2 sets; therefore in the following, we require that the *secondary* membership functions of the type-2 sets involved be Gaussians. *Their principal membership functions, however, can have any shape (e.g., triangular, Gaussian, trapezoidal).*

Consider the case when $f(v)$ and $g(w)$ are Gaussians with support $[0,1]$ with means m_f, m_g and standard deviations σ_f, σ_g, respectively. Then,

$$\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}} = \int_v \int_w e^{-\frac{1}{2}\left(\frac{v-m_f}{\sigma_f}\right)^2} e^{-\frac{1}{2}\left(\frac{w-m_g}{\sigma_g}\right)^2} / (vw) \tag{2.50}$$

Recall that the integral in the above equation denotes union in the continuum. If θ is an element of $\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}}$, then the membership grade of θ can be found by : finding all the pairs $\{v, w\}$ such that $v \in \tilde{\mathrm{F}}$, $w \in \tilde{\mathrm{G}}$ and $vw = \theta$; multiplying the membership grades of v and w in each pair; and then finding the maximum of these products of membership grades. That is,

$$\mu_{\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}}}(\theta) = \sup \left\{ e^{-\frac{1}{2}\left(\frac{v-m_f}{\sigma_f}\right)^2} e^{-\frac{1}{2}\left(\frac{w-m_g}{\sigma_g}\right)^2} ; vw = \theta; v \in \tilde{\mathrm{F}}; w \in \tilde{\mathrm{G}} \right\} \tag{2.51}$$

Given any v (assuming $v \neq 0$), the constraint $vw = \theta$ gives us $w = \theta/v$. Further, since $w \in [0,1]$, it follows that $\theta/v \leq 1$ or $v \geq \theta$. So, given any $\theta \in (0,1]$, the acceptable $\{v, w\}$ pairs that can give θ as the result of the product operation are $\{(v, \frac{\theta}{v}); 0 < \theta \leq v \leq 1\}$; therefore, from (2.51), we have

$$\begin{aligned} \mu_{\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}}}(\theta) &= \sup_{v \in [\theta,1]} e^{-\frac{1}{2}\left[\left(\frac{v-m_f}{\sigma_f}\right)^2 + \left(\frac{\frac{\theta}{v}-m_g}{\sigma_g}\right)^2\right]}, \ \theta \neq 0 \\ &= \sup_{v \in [\theta,1]} e^{-\frac{1}{2}\left[\left(\frac{v-m_f}{\sigma_f}\right)^2 + \left(\frac{\theta-v m_g}{v\sigma_g}\right)^2\right]}, \ \theta \neq 0 \end{aligned} \tag{2.52}$$

When $\theta = 0$, either $v = 0$ and w is any number in $[0,1]$, or $w = 0$ and v is any number in $[0,1]$; therefore, from (2.51), we have

$$\begin{aligned} \mu_{\tilde{\mathrm{F}} \sqcap \tilde{\mathrm{G}}}(0) &= \sup_{w \in [0,1]} e^{-\frac{1}{2}\left(\frac{m_f}{\sigma_f}\right)^2} e^{-\frac{1}{2}\left(\frac{w-m_g}{\sigma_g}\right)^2} \vee \sup_{v \in [0,1]} e^{-\frac{1}{2}\left(\frac{v-m_f}{\sigma_f}\right)^2} e^{-\frac{1}{2}\left(\frac{m_g}{\sigma_g}\right)^2} \\ &= e^{-\frac{1}{2}\left(\frac{m_f}{\sigma_f}\right)^2} \vee e^{-\frac{1}{2}\left(\frac{m_g}{\sigma_g}\right)^2} \end{aligned} \tag{2.53}$$

Solving the optimization problem in (2.52), in general, is quite complicated and does not lead to a closed-form expression (see Appendix C.6). Also, since the final result is non-Gaussian, it can not be easily generalized to the case of the *meet* of more than two Gaussians at a time; therefore, we now try to find a Gaussian approximation to this result.

The supremum in (2.52) can be obtained by minimizing the exponent on the RHS of (2.52). Let us call the exponent $J(v)$. So, we want to minimize

$$J(v) = \left(\frac{v - m_f}{\sigma_f}\right)^2 + \left(\frac{\theta - m_g v}{\sigma_g v}\right)^2, \ \theta \neq 0 \tag{2.54}$$

with the constraint $v \in [\theta, 1]$. Since the second term on the RHS of (2.54) has v in its denominator, J is non-convex and is difficult to minimize. The actual function resulting from the minimization of J is non-Gaussian. In order to find a Gaussian approximation for $\tilde{F} \sqcap \tilde{G}$, we simplify the problem a bit.

Equation (2.50) can be interpreted as follows. Each element v of set $\tilde{F}$ multiplies every element w of set $\tilde{G}$, and, at the same time, the membership grade of v in $\tilde{F}$ multiplies the membership grade of w in $\tilde{G}$. So, given a particular element v_1 of $\tilde{F}$, what we get as a result of these multiplications is a scaled version of the membership function of $\tilde{G}$ (scaled along both the axes : along the independent axis by v_1 and along the dependent axis by $e^{-\frac{1}{2}(\frac{v_1 - m_f}{\sigma_f})^2}$). This process is repeated for every element of $\tilde{F}$ and finally, the meet of $\tilde{F}$ and $\tilde{G}$ is given by the envelope of all the above scaled Gaussians. The expression for the membership of an element θ in $\tilde{F} \sqcap \tilde{G}$ is given by (2.52) and (2.53).

In order to simplify the problem, we replace the v in the denominator of the second term on the RHS of (2.54) by a constant k. By solving this simplified optimization problem, we get an approximation to $\tilde{F} \sqcap \tilde{G}$. Let's call it $\tilde{E}$, so that

$$\begin{aligned} \mu_{\tilde{E}}(\theta) &= \sup_{v \in [\theta,1]} e^{-\frac{1}{2}\left(\frac{v-m_f}{\sigma_f}\right)^2} e^{-\frac{1}{2}\left(\frac{\theta - m_g v}{k\sigma_g}\right)^2} \\ &= \sup_{v \in [\theta,1]} e^{-\frac{1}{2}\left[\left(\frac{v-m_f}{\sigma_f}\right)^2 + \left(\frac{\theta - m_g v}{k\sigma_g}\right)^2\right]} \end{aligned} \tag{2.55}$$

Observe that the only difference between (2.55) and (2.52) is that the standard deviation of the second Gaussian in (2.55) is a constant ($k\sigma_g$), whereas that in (2.52) is proportional to v ($v\sigma_g$).

To see the dependence of $\mu_{\tilde{E}}(\theta)$ on k, let

$$H(v,k) = \left(\frac{v - m_f}{\sigma_f}\right)^2 + \left(\frac{\theta - m_g v}{k\sigma_g}\right)^2 \quad ; \quad k \in (0,1] \tag{2.56}$$

Obviously,

$$\begin{array}{lcccccl} & H(v,\epsilon) & \geq & H(v,k) & \geq & H(v,1) & ; \ 0 < \epsilon \leq k \leq 1 \\ \Rightarrow & \inf_{v\in[\theta,1]} H(v,\epsilon) & \geq & \inf_{v\in[\theta,1]} H(v,k) & \geq & \inf_{v\in[\theta,1]} H(v,1) & ; \ 0 < \epsilon \leq k \leq 1 \\ \Rightarrow & \mu_{\tilde{E}}(\theta)\Big|_{k=\epsilon} & \leq & \mu_{\tilde{E}}(\theta) & \leq & \mu_{\tilde{E}}(\theta)\Big|_{k=1} & ; \ 0 < \epsilon \leq k \leq 1 \end{array} \tag{2.57}$$

Observe that as $\epsilon \to 0$, $H(v,\epsilon) \to \infty$ and $\mu_{\tilde{E}}(\theta) \to 0$. Since k in (2.55) replaces v in the actual problem (which varies between 0 and 1), it is apparent from (2.57), that $\lim_{k\to 0} \mu_{\tilde{E}}(\theta)$ gives a lower bound on $\mu_{\tilde{F}\sqcap\tilde{G}}(\theta)$ and $\mu_{\tilde{E}}(\theta)\Big|_{k=1}$ gives an upper bound on $\mu_{\tilde{F}\sqcap\tilde{G}}(\theta)$.

Recall that our earlier approximation in (2.46), which was a Gaussian with standard deviation equal to $\max\{m_f\sigma_g, m_g\sigma_f\}$, was motivated by considering the case where one of the Gaussians was a singleton, i.e., one of the Gaussians has a zero standard deviation, which is analogous to assuming that $k = 0$ in (2.55). Therefore, from the discussion above, it is clear that this earlier approximation acts as a lower bound on $\mu_{\tilde{F}\sqcap\tilde{G}}(\theta)$. [Figure 2.15 (f) depicts a result that seems to contradict this statement; however, it is obtained because only half of one of the participating Gaussians is contained in [0, 1]. See the discussion at the end of this Section.]

From this discussion, it seems conceivable that choosing some $\mu_{\tilde{E}}(\theta)$ between the upper and the lower bounds, i.e., substituting some value of $k \in (0,1)$ into (2.55), should enable us to obtain a good approximation to $\mu_{\tilde{F}\sqcap\tilde{G}}(\theta)$. Our criterion for choosing k is that the approximation should be commutative (i.e., if we switch $\tilde{F}$ and $\tilde{G}$, we should still get the same result), because the true result is commutative. Considering both these factors, we choose $k = m_f$. Refer to Appendix C.6 for details of the solution of (2.55) and the choice of k. We state the result here.

If $\tilde{F}$ and $\tilde{G}$ are two Gaussian type-1 fuzzy sets in $[0,1]$ having means m_f and m_g and standard deviations σ_f and σ_g, then the membership function for the *meet* of $\tilde{F}$ and $\tilde{G}$ can be approximated as

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) \approx e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{\sqrt{m_g^2\sigma_f^2 + m_f^2\sigma_g^2}}\right)^2} \tag{2.58}$$

Generalization to the case of more than two Gaussians is straightforward. Assume that $\tilde{L}$ is a Gaussian fuzzy set with mean m_l and standard deviation σ_l. Using the associative property, we have $\tilde{F}\sqcap\tilde{G}\sqcap\tilde{L} = (\tilde{F}\sqcap\tilde{G})\sqcap\tilde{L}$; hence, using (2.58), we have

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) \approx e^{-\frac{1}{2}\left(\frac{\theta - m_{fg}}{\sigma_{fg}}\right)^2} \tag{2.59}$$

where $m_{fg} = m_f m_g$ and $\sigma_{fg} = \sqrt{m_g^2\sigma_f^2 + m_f^2\sigma_g^2}$. Using (2.58) again, we have

$$\begin{aligned} \mu_{\tilde{F}\sqcap\tilde{G}\sqcap\tilde{L}}(\theta) &\approx e^{-\frac{1}{2}\left(\frac{\theta - m_{fg}m_l}{\sqrt{m_l^2\sigma_{fg}^2 + m_{fg}^2\sigma_l^2}}\right)^2} \\ &= e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g m_l}{\sqrt{m_l^2 m_g^2\sigma_f^2 + m_l^2 m_f^2\sigma_g^2 + m_f^2 m_g^2\sigma_l^2}}\right)^2} \end{aligned} \tag{2.60}$$

If there are n Gaussian fuzzy sets $\tilde{F}_1, \tilde{F}_2, \cdots, \tilde{F}_n$ with means $m_1, m_2, \cdots, m_n$ and standard deviations $\sigma_1, \sigma_2, \cdots, \sigma_n$, respectively, then repeated application of (2.58) yields

$$\mu_{\tilde{F}_1\sqcap\tilde{F}_2\sqcap\cdots\sqcap\tilde{F}_n}(\theta) \approx e^{-\frac{1}{2}\left(\frac{\theta - m_1 m_2\cdots m_n}{\bar{\sigma}}\right)^2} \tag{2.61}$$

where

$$\bar{\sigma} = \sqrt{\sigma_1^2 \prod_{i\,;\,i\neq 1} m_i^2 + \sigma_2^2 \prod_{i\,;\,i\neq 2} m_i^2 + \cdots + \sigma_j^2 \prod_{i;i\neq j} m_i^2 + \cdots + \sigma_n^2 \prod_{i;i\neq n} m_i^2}\,;\ i = 1, 2, \cdots, n \tag{2.62}$$

Comparing (2.61) with (2.48), we see that the two approximations have the same mean. Only the standard deviations are different. All the results and approximations that we have developed will finally be used for operations between membership grades of type-2 sets.

Recall that we require that all our results remain valid if we replace all the type-2 sets by corresponding type-1 sets (i.e., type-1 sets having the principal membership functions of the type-2 sets as their membership functions). In case of Gaussian type-2 sets, replacing type-2 sets by corresponding type-1 sets is analogous to reducing the standard deviations of all the secondary membership functions to zero. [Observe that a Gaussian with a zero standard deviation is like an impulse function. As we reduce the standard deviation of a Gaussian, keeping its mean constant, it grows narrower and narrower. The height of the Gaussian at the mean remains unchanged though. In the limit, as the standard deviation reduces to zero, only the mean of the Gaussian has a non-zero membership, which is equal to 1. Mathematically, if we have a Gaussian with mean m and standard deviation σ, $\lim_{\sigma\to 0}(\frac{\theta - m}{\sigma}) \to \infty$ if $\theta \neq m$ and $\lim_{\sigma\to 0}(\frac{\theta - m}{\sigma}) = 0$ if $\theta = m$. So, in the limit $\exp\{-\frac{1}{2}(\frac{\theta - m}{\sigma})^2\}$ is equal to 1 if $\theta = m$ and is equal to 0 otherwise.] All that remains of a membership grade after doing this is a crisp number in $[0,1]$ equal to the center of the Gaussian in the type-2 case; and the *meet* operation reduces to the product of all these crisp numbers. The actual type-2 result for the *meet* between Gaussians (Appendix C.6) as well as both our approximations [(2.61) and (2.48)] obey the same result. If we reduce the standard deviations of all the Gaussians involved to zero, the result of the *meet* operation is equal to that in the type-1 case.

Figures 2.15 (a) - (f) show some examples of this approximation. In general, if the Gaussians have ± 2 (or more) standard deviations contained within $[0, 1]$, the results look quite good. In Fig. 2.15 (f), one of the Gaussians is centered at 1, so only half of this Gaussian (the part lying to the left of the mean) is contained in $[0, 1]$. Consequently, the result of the *meet* is much more "non-Gaussian" than earlier cases, i.e., the difference between our Gaussian approximation and the actual curve is larger than that in the other examples. Also, observe that the first approximation in (2.48) does not act as a lower bound on the result of the *meet* in this case.

Appendix C.6 derives bounds on the error between the Gaussian approximation and the result of the actual *meet* operation, by finding upper and lower bounds which contain both the approximation as well as the actual function.

2.3.3 A Triangular Approximation

See [8] for a triangular *meet* approximation similar to the Gaussian approximation derived in Section 2.3.2.

2.4 Algebraic Operations on Fuzzy Numbers

As already mentioned, convex and normal type-1 fuzzy subsets of the real line are also known as *fuzzy numbers* [5, 10]. Algebraic operations like addition and multiplication between fuzzy numbers can be defined using the Extension Principle, just as we defined the t-norm and t-conorm (i.e., *meet* and *join*) operations (see, for example, [5, 10]). The two operations of most interest to us are multiplication and addition. We will use the results of this section in Chapter 5.

2.4.1 Multiplication of Type-1 Fuzzy Numbers

The product of two fuzzy numbers $\tilde{F} = \int f(v)/v$ and $\tilde{G} = \int g(w)/w$ is defined as

$$\tilde{F} \times \tilde{G} = \int_v \int_w [f(v) \star g(w)]/[v \times w] \tag{2.63}$$

where $\star$ indicates the t-norm used.

Observe, from (2.30) and (2.63), that under product t-norm, the product of $\tilde{F}$ and $\tilde{G}$ is the same as the *meet* of $\tilde{F}$ and $\tilde{G}$, i.e., $\tilde{F} \times \tilde{G} = \tilde{F} \sqcap \tilde{G}$; so, all our earlier discussion about *meet* under product t-norm applies to multiplication of fuzzy numbers under product t-norm. We do not discuss multiplication under minimum t-norm, rather we focus on the addition of fuzzy numbers.

2.4.2 Addition of Type-1 Fuzzy Numbers

The addition of two fuzzy numbers $\tilde{F} = \int f(v)/v$ and $\tilde{G} = \int g(w)/w$ is defined as

$$\tilde{F} + \tilde{G} = \int_v \int_w [f(v) \star g(w)]/[v + w] \tag{2.64}$$

When $\tilde{F}$ and $\tilde{G}$ are interval type-1 sets, (2.64) simplifies considerably, as we show next.

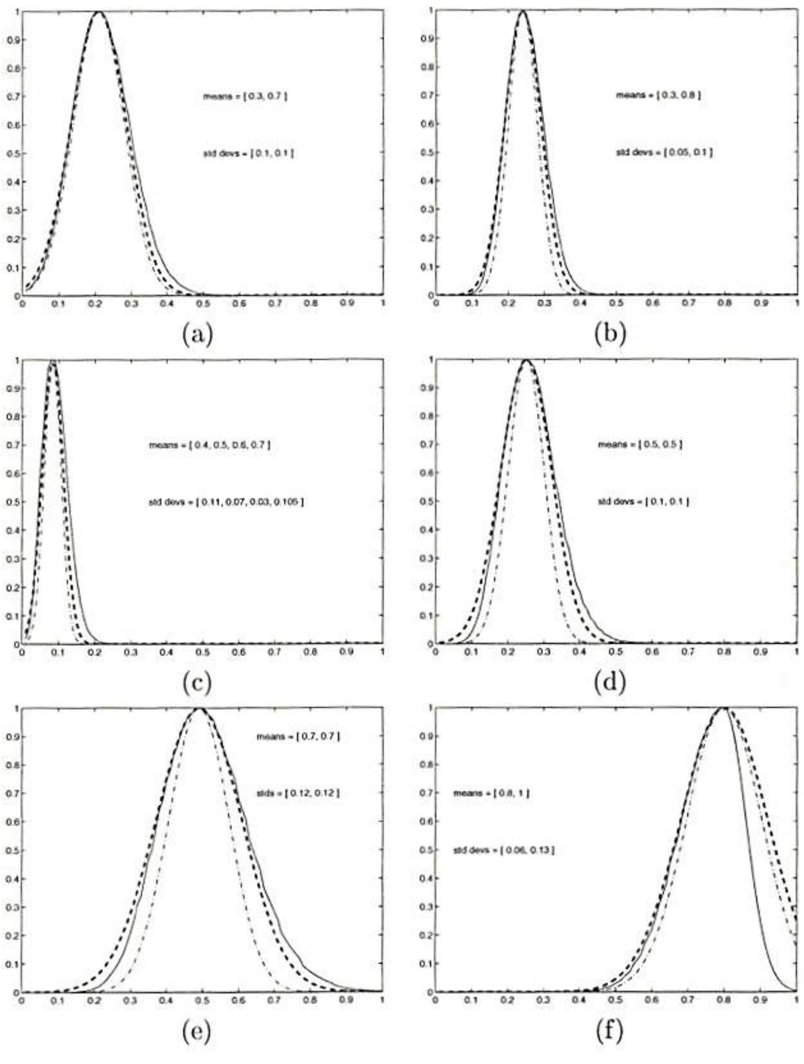


Figure 2.15: Actual and approximate results of the *meet* operation between Gaussians under product t-norm. The thin solid line shows the actual result computed numerically. The thin dash-dotted line shows the first approximation in (2.48). The approximation in (2.61) is shown by the thick dashed line. Means and standard deviations of the Gaussians are as indicated in the figure. In Figs. (d) and (e), the two Gaussians are coincident (the same curve). The first approximation does very poorly in this case. In Fig. (f), observe the difference between the approximation and the actual curve on the RHS of the mean. This is due to the fact that one of the Gaussians is centered at 1, i.e., only half of it lies in $[0, 1]$. This clipping effect is discussed in Appendix C.8.2.

2.4.2.1 Addition of Interval Type-1 Numbers

Let F and G be two interval type-1 sets with domains $[l_f, r_f]$ and $[l_g, r_g]$, respectively. Using (2.64), the algebraic sum of F and G, can be obtained as

$$F + G = \int_{v \in F} \int_{w \in G} (1 \star 1)/(v + w) \tag{2.65}$$

Observe, from (2.65), that

- each term in $F + G$ is equal to the sum $(v + w)$ for some $v \in F$ and $w \in G$, the smallest term being $(l_f + l_g)$ and the largest $(r_f + r_g)$; and,
- since both F and G have continuous domains, $F + G$ also has a continuous domain;

consequently, $F + G$ is an interval type-1 set with domain $[l_f + l_g, r_f + r_g]$, i.e.,

$$F + G = \int_{v \in [l_f + l_g, r_f + r_g]} 1/v \tag{2.66}$$

Similarly, the algebraic sum of n interval type-1 numbers $F_1, \ldots, F_n$, having domains $[l_1, r_1], \ldots, [l_n, r_n]$, respectively, is an interval type-1 set with domain $[\sum_{i=1}^n l_i, \sum_{i=1}^n r_i]$. See [10] for a similar result.

See Theorem D.1 for an expression for an affine combination of interval type-1 sets.

NOTE : Observe, from (2.65) that while performing algebraic operations on interval type-1 sets, the choice of t-norm does not matter, since all the memberships involved are unity.

2.4.2.2 Addition of Gaussian Type-1 Numbers

Theorem 2.4 *Given n type-1 Gaussian fuzzy numbers $\tilde{F}_1, \ldots, \tilde{F}_n$, with means $m_1, m_2, \ldots, m_n$ and standard deviations $\sigma_1, \sigma_2, \ldots, \sigma_n$, their affine combination $\sum_{i=1}^n \alpha_i \tilde{F}_i + \beta$, where α_i $(i = 1, \ldots, n)$ and β are crisp constants, is also a Gaussian fuzzy number with mean $\sum_{i=1}^n \alpha_i m_i + \beta$, and standard deviation Σ', where*

$$\Sigma' = \begin{cases} \sqrt{\sum_{i=1}^n \alpha_i^2 \sigma_i^2} & , \textit{ if product t-norm is used} \\ \sum_{i=1}^n |\alpha_i \sigma_i| & , \textit{ if minimum t-norm is used} \end{cases} \tag{2.67}$$

□

See Appendix C.9 for the proof of Theorem 2.4.

2.4.2.3 Addition of Triangular Type-1 Numbers

For comparable results about triangular fuzzy numbers, see [8].

2.5 Centroid of a Type-2 Set

In this section, we extend the concept of the centroid of a fuzzy set from type-1 to type-2. We will use the results of this section in Chapter 5, when we introduce the new operation of type-reduction.

The centroid of a type-1 set $\tilde{A}$ whose domain is discrete with N points or is discretized into N points, is given as

$$C_{\tilde{A}} = \frac{\sum_{i=1}^N x_i \mu_{\tilde{A}}(x_i)}{\sum_{i=1}^N \mu_{\tilde{A}}(x_i)} \tag{2.68}$$

Similarly, the centroid of a type-2 set $\tilde{\tilde{A}}$, whose domain is discretized into N points (see Chapter 1 for the definition of the domain of a type-2 set), can be defined using the Extension Principle [see Appendix B, especially (B.13)] as follows. If we let $\tilde{D}_i = \tilde{\mu}_{\tilde{\tilde{A}}}(x_i)$, then

$$\tilde{C}_{\tilde{\tilde{A}}} = \int_{\theta_1} \cdots \int_{\theta_N} [\mu_{\tilde{D}_1}(\theta_1) \star \cdots \star \mu_{\tilde{D}_N}(\theta_N)] \bigg/ \frac{\sum_{i=1}^{N} x_i \theta_i}{\sum_{i=1}^{N} \theta_i} \tag{2.69}$$

where $\theta_i \in \tilde{D}_i$.

Equation (2.69) can be described in words as follows. Each point x_i of $\tilde{\tilde{A}}$ has a type-1 fuzzy membership grade, $\tilde{D}_i = \tilde{\mu}_{\tilde{\tilde{A}}}(x_i)$, associated with it. To find the centroid, we consider every possible combination $\{\theta_1, \ldots, \theta_N\}$ such that $\theta_i \in \tilde{D}_i$. For every such combination, we perform the type-1 centroid calculation in (2.68) by using θ_i's in place of $\tilde{\mu}_{\tilde{A}}(x_i)$'s; and, to each point in the centroid, we assign a membership grade equal to the t-norm of the membership grades of the θ_is in the $\tilde{D}_i$s. If more than one combination of θ_i's gives us the same point in the centroid, we keep the one with the largest membership grade. If we let $\sum_{i=1}^{N} x_i\theta_i \big/ \sum_{i=1}^{N} \theta_i = x$, then (2.69) can also be written as

$$\tilde{C}_{\tilde{\tilde{A}}} = \int_x \sup_{\{\theta_1,\ldots,\theta_N\}} [\mu_{\tilde{D}_1}(\theta_1) \star \cdots \star \mu_{\tilde{D}_N}(\theta_N)] \big/ x \tag{2.70}$$

where $\{\theta_1, \ldots, \theta_N\}$ are such that $\sum_{i=1}^{N} x_i\theta_i \big/ \sum_{i=1}^{N} \theta_i = x$. We will illustrate the calculation of $\tilde{C}_{\tilde{\tilde{A}}}$ below, in Example 2.5. First, however, we provide some general insights into (2.69).

A type-2 set $\tilde{\tilde{A}}$ can be thought of as a collection of type-1 sets, which we call type-1 sets *embedded* in $\tilde{\tilde{A}}$.

Definition : A type-1 set $\tilde{A}$ embedded in a type-2 set $\tilde{\tilde{A}}$ is a type-1 set for which : (1) $x \in \tilde{\tilde{A}} \Leftrightarrow x \in \tilde{A}$, and (2) $\mu_{\tilde{A}}(x) \in \tilde{\mu}_{\tilde{\tilde{A}}}(x)\ \forall x \in \tilde{\tilde{A}}$.

In the 2-D representation of a type-2 set, an embedded type-1 set is one whose membership function lies inside the shaded region. Figure 2.16 shows an example of a type-1 set embedded in a type-2 set. For the type-2 set $\tilde{\tilde{A}}$ in (2.69), every combination $\{\theta_1, \ldots, \theta_N\}$ such that $\theta_i \in \tilde{D}_i$, corresponds to the membership function of an embedded type-1 set. The centroid of $\tilde{\tilde{A}}$, $\tilde{C}_{\tilde{\tilde{A}}}$, can be thought of as a type-1 set whose elements are the centroids of all the embedded type-1 sets in $\tilde{\tilde{A}}$. The membership grade of an embedded set centroid in $\tilde{C}_{\tilde{\tilde{A}}}$ is calculated as the t-norm of all the secondary memberships corresponding to $\{\theta_1, \ldots, \theta_N\}$ that make up that embedded set. When $\tilde{\tilde{A}}$ collapses to an embedded type-1 set $\tilde{A}$, which corresponds to the combination $\{\theta'_1, \ldots, \theta'_N\}$, each $\tilde{D}_i$ reduces to a fuzzy singleton, such that $\mu_{\tilde{D}_i}(\theta'_i) = 1$ and $\mu_{\tilde{D}_i}(\theta_i) = 0$ if $\theta_i \neq \theta'_i$; therefore, we get $\mathcal{T}_{i=1}^{N} \mu_{\tilde{D}_i}(\theta'_i) = 1$, and for all other $\{\theta_1, \ldots, \theta_N\}$ combinations $\mathcal{T}_{i=1}^{N} \mu_{\tilde{D}_i}(\theta_i) = 0$. Consequently, $\tilde{C}_{\tilde{\tilde{A}}}$ reduces to the crisp number $C_{\tilde{A}}$, the centroid of $\tilde{A}$.

Observe that if the domain of $\tilde{\tilde{A}}$ and/or $\tilde{\mu}_{\tilde{\tilde{A}}}(x)$ ($x \in \tilde{\tilde{A}}$) is continuous, the domain of $\tilde{C}_{\tilde{\tilde{A}}}$ is also continuous. The number of all the embedded type-1 sets in $\tilde{\tilde{A}}$, in this case, is uncountable; therefore, the domains of $\tilde{\tilde{A}}$ and each $\tilde{\mu}_{\tilde{\tilde{A}}}(x)$ ($x \in \tilde{\tilde{A}}$) have to be discretized for the calculation of $\tilde{C}_{\tilde{\tilde{A}}}$. Observe, from (2.69), that if the domain of each $\tilde{D}_l$ is discretized into M points, the number of possible $\{\theta_1, \ldots, \theta_N\}$ combinations is M^N, which can be very large even for small M and N. If, however, the membership functions of $\tilde{D}_l$'s have a regular structure (e.g., Gaussian, triangular, interval), we can approximate the centroid without having to do all the calculations. See Example 2.5, Sections 2.5.2.1 and 2.5.2.2, and [8] for more details. In the case of an interval type-2 set, even the actual centroid can be obtained relatively easily by using the computational procedure described in Appendix D.

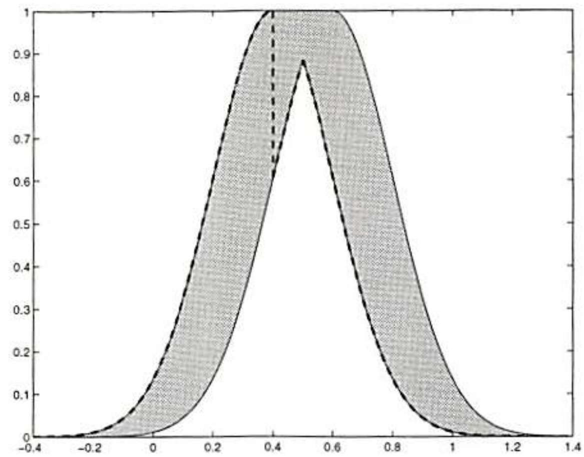


Figure 2.16: Example of a type-1 set, shown with the thick dashed line, embedded in a type-2 set.

Example 2.5 In this example, we show the centroid calculation for a type-2 set that results from a type-1 set with only location uncertainty, e.g., see Example 1.2. We focus on the special case of Gaussian membership functions with uncertain means, such that every value of the mean is equally uncertain. In this case, we set all the secondary memberships equal to 1, to indicate that the level of uncertainty associated with every primary membership is the same, so that the resulting set is an interval type-2 set.

Figure 2.17 (a) shows a type-2 set $\tilde{\tilde{A}}$ resulting from a Gaussian type-1 set with mean uniformly uncertain in the interval $[m_1, m_2]$. In the figure, $m_1 = 0.45$, $m_2 = 0.55$ and the standard deviation, $\sigma = 0.2$. All the secondary memberships are equal to 1.

Observe, from (2.69), that :

1. all the secondary memberships are equal to 1, so the membership of each point in the centroid is also equal to 1, i.e., $\mu_{\tilde{D}_1}(\theta_1) \star \cdots \star \mu_{\tilde{D}_N}(\theta_N) = 1$; hence, the centroid is a crisp set;

2. the mean varies on a continuous domain $[m_1, m_2]$, so the crisp set corresponding to the centroid will also have a continuous domain; and,

3. each Gaussian centered at $m \in [m_1, m_2]$ is an embedded set in $\tilde{\tilde{A}}$, so the centroid of each such Gaussian $\Big($ i.e., each $m \in [m_1, m_2]\Big)$ will be an element of the centroid.

From these three observations, we see that the centroid of $\tilde{\tilde{A}}$ is some interval, $[c_l, c_r]$, which contains $[m_1, m_2]$. Now, we have to find the end-points of this interval. To do this, we show how to compute the left end-point, c_l. Since the set is symmetrical, the calculation of c_r will be similar.

It is easy to verify that the left end-point c_l is the centroid of the embedded type-1 set which assigns the highest possible memberships to all the points to the left of its centroid and lowest possible memberships to all the points to the right of its centroid (see the computational procedure in Appendix D.1 for more discussion). Any change in this membership function will always cause its centroid to move towards the right, implying that the centroid of this embedded type-1 set is equal to c_l. An example of such an embedded type-1 set is shown by the thick dashed line in Fig. 2.17 (b).

Though we do not know the exact value of c_l, we can make an estimate by considering the embedded type-1 set shown in Fig. 2.17 (c). This type-1 set is formed by assigning the highest possible memberships to the points to the left of m_1 and the lowest possible memberships to the points to the right of m_1. The membership function of this set looks like a Gaussian with a small

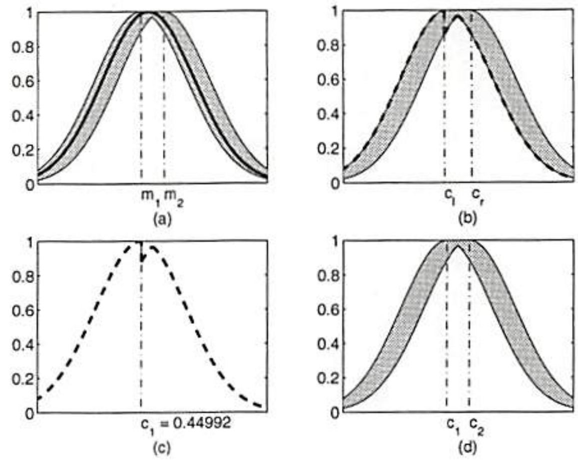


Figure 2.17: Figures for Example 2.5. (a) An interval type-2 set $\tilde{\tilde{A}}$ resulting from a Gaussian type-1 set with standard deviation equal to 0.2 and mean uniformly uncertain in the interval $[m_1, m_2] = [0.45, 0.55]$. The thick line shows an embedded Gaussian type-1 set. (b) The embedded type-1 set whose centroid equals c_l is shown with a thick dashed line. (c) The type-1 set formed by assigning highest possible memberships to the points to the left of m_1 and lowest possible memberships to the points to the right of m_2. The centroid of this set is $c_1 = 0.44992 \approx c_l$. (d) The centroid of $\tilde{\tilde{A}}$ is the interval $[c_l, c_r] \approx [c_1, c_2] \approx [m_1, m_2]$.

portion missing, and its centroid (which was calculated numerically) is equal to $c_1 = 0.44992$, which is a little bit to the left of m_1. The fact that c_1 is just slightly less than m_1 shows that the area to the left of m_1 in the type-1 set in Fig. 2.17 (c) is just slightly more than the area to its right; therefore, c_l will also be just slightly less than m_1.

Similarly, the embedded type-1 set constructed by assigning highest possible memberships to all the points to the right of m_2 and lowest possible memberships to the points to the left of m_2, has a centroid $c_2 = 0.55008$, which is slightly larger than m_2; hence, we conclude that c_r will be just a little bit larger than m_2. We can, therefore, say that $c_l \approx c_1$ and $c_r \approx c_2$. Figure 2.17 (d) shows $\tilde{\tilde{A}}$ with its centroid , which is a crisp set with domain $[c_l, c_r] \approx [c_1, c_2] \approx [m_1, m_2]$.

It can be shown that, if $(m_2 - m_1)$ is small compared to the standard deviation (σ) of $\tilde{A}$, then $[c_l, c_r] \approx [m_1, m_2]$ (see Appendix C.10 for the proof).

If we increase $(m_2 - m_1)$, keeping σ the same, the difference between the approximation and the true centroid (computed using the computational procedure in Appendix D.1) increases, e.g., for $\sigma = 0.2$, if $\{m_1, m_2\} = \{0.4, 0.5\}$, $\{c_1, c_2\} = \{0.39855, 0.60145\}$, and, if $\{m_1, m_2\} = \{0.3.0.7\}$, $\{c_1, c_2\} = \{0.28146, 0.71854\}$. We, therefore, recommend using the computational procedure described in Appendix D.1 to obtain the centroid, if $(m_2 - m_1)$ is not small compared to σ. □

See Appendix D for a computational procedure to compute the centroid of a general interval type-2 set. We next describe a problem that arises when one attempts to compute the centroid of a type-2 set having a continuous domain using product t-norm.

2.5.1 Centroid Calculation Using the Product t-norm

Calculation of the centroid, using product t-norm, of a type-2 set which has a continuous domain and not all of whose secondary memberships are unity, gives us an unexpected result. In this section, we concentrate on type-2 sets having a continuous domain whose secondary membership functions are such that, for any domain point, only one primary membership has a secondary membership equal to one, e.g., Gaussian or triangular type-2 sets. We first describe the problem and then discuss its cause and remedy.

Problem

In the discussion associated with (2.69), we assumed that the domain of $\tilde{\tilde{A}}$ is discretized into N points. The true centroid of $\tilde{\tilde{A}}$ (assuming $\tilde{\tilde{A}}$ has a continuous domain) is the limit of $\tilde{C}_{\tilde{\tilde{A}}}$ in (2.69) as $N \to \infty$. When we use the product t-norm $\lim_{N\to\infty} \mathcal{T}_{i=1}^{N} \mu_{\tilde{D}_i}(\theta_i) = \lim_{N\to\infty} \prod_{i=1}^{N} \mu_{\tilde{D}_i}(\theta_i)$.

Let $\tilde{B}$ be an embedded type-1 set in $\tilde{\tilde{A}}$. The centroid of $\tilde{B}$ is computed as

$$C_{\tilde{B}} = \frac{\sum_{i=1}^{N} x_i \mu_{\tilde{B}}(x_i)}{\sum_{i=1}^{N} \mu_{\tilde{B}}(x_i)} \tag{2.71}$$

and the membership of $C_{\tilde{B}}$ in $\tilde{C}_{\tilde{\tilde{A}}}$ [denoted as $\mu_{\tilde{C}}(C_{\tilde{B}})$] is

$$\mu_{\tilde{C}}(C_{\tilde{B}}) = \prod_{i=1}^{N} \mu_{\tilde{D}_i}(\theta_i) \tag{2.72}$$

where $\{\theta_1, \ldots, \theta_N\}$ are the primary memberships that make up the type-1 set $\tilde{B}$.

Let $\tilde{A}$ denote the principal membership function of $\tilde{\tilde{A}}$. Obviously, $\mu_{\tilde{C}}(C_{\tilde{A}}) = 1$.

Consider the case where the secondary membership functions are like Gaussians or triangles (having only one point with unity membership). We make two observations :

1. $\lim_{N\to\infty} \mu_{\tilde{C}}(C_{\tilde{B}})$ is non-zero only if $\tilde{B}$ differs from $\tilde{A}$ in *at most a finite number of points.* For all other embedded sets $\tilde{B}$, the product of an infinite number of quantities less than one will cause $\mu_{\tilde{C}}(C_{\tilde{B}})$ to go to zero as $N \to \infty$.

2. For any embedded set $\tilde{B}$, whose membership function differs from that of $\tilde{A}$ in only a finite number of points (i.e., when $\mu_{\tilde{B}}(x) \neq \mu_{\tilde{A}}(x)$, for only a finite number of points x), $C_{\tilde{B}} = C_{\tilde{A}}$. This can be explained as follows :

   The (true) centroid of $\tilde{B}$ is defined as

   $$C_{\tilde{B}} = \frac{\int_x x \mu_{\tilde{B}}(x) dx}{\int_x \mu_{\tilde{B}}(x) dx} \tag{2.73}$$

   where $x \in \tilde{B}$. Since $\tilde{A}$ and $\tilde{B}$ share the same domain (both are embedded sets in $\tilde{\tilde{A}}$), $x \in \tilde{A} \Leftrightarrow x \in \tilde{B}$; and since $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$ differ only in a finite number of points, $\int_x x\mu_{\tilde{B}}(x)dx = \int_x x\mu_{\tilde{A}}(x)dx$ and $\int_x \mu_{\tilde{B}}(x)dx = \int_x \mu_{\tilde{A}}(x)dx$; therefore, $C_{\tilde{B}} = C_{\tilde{A}}$.

From these two observations, we can see that the only point having non-zero membership in $\tilde{C}_{\tilde{\tilde{A}}}$ is equal to $C_{\tilde{A}}$; and its membership grade is equal to the supremum of the membership grades of all the embedded type-1 sets which have the same centroid, which is equal to 1 [since $\mu_{\tilde{C}}(C_{\tilde{A}}) = 1$]. In other words, $\tilde{C}_{\tilde{\tilde{A}}} = 1/C_{\tilde{A}} = C_{\tilde{A}}$, i.e., the centroid of $\tilde{\tilde{A}}$ will be equal to a crisp number ..... the centroid of its principal membership function !

Cause

The above problem occurs because, under the product t-norm, $\lim_{N\to\infty} \mathcal{T}_{i=1}^{N} \mu_{\tilde{D}_i}(\theta_i) = \lim_{N\to\infty} \prod_{i=1}^{N} \mu_{\tilde{D}_i}(\theta_i) = 0$, unless only a finite number of $\mu_{\tilde{D}_i}(\theta_i)$'s are less than 1. The minimum t-norm does not cause such a problem.

Remedy

One obvious way to deal with the problem explained above is to *not* use product t-norm for centroid calculation. *From now on, we will always use the minimum t-norm to calculate the centroid of a type-2 set having a continuous domain.*

2.5.2 Approximations to Centroids of Certain Type-2 Sets

In this section we develop approximations to the Centroids of Gaussian and interval type-2 sets.

2.5.2.1 Centroid of a Gaussian Type-2 Set

We first prove a general result and then use it to find the centroid of a Gaussian type-2 set.
Weighted Average of Gaussian Type-1 Sets : Consider the weighted average

$$y(z_1, \ldots, z_M, w_1, \ldots, w_M) = \frac{\sum_{l=1}^{M} w_l z_l}{\sum_{l=1}^{M} w_l} \tag{2.74}$$

where $z_l \in \Re$ and $w_l \in [0,1]$ for $l = 1, \ldots, M$. If each z_l is replaced by a type-1 fuzzy set $\tilde{\mathrm{Z}}_l \subset \Re$ and each w_l is replaced by a type-1 fuzzy set $\tilde{\mathrm{W}}_l \subset [0,1]$, then the extension of (2.74) gives

$$\tilde{\mathrm{Y}}(\tilde{\mathrm{Z}}_1, \ldots, \tilde{\mathrm{Z}}_M, \tilde{\mathrm{W}}_1, \ldots, \tilde{\mathrm{W}}_M) = \int_{z_1} \cdots \int_{z_M} \int_{w_1} \cdots \int_{w_M} \mathcal{T}_{l=1}^{M} \mu_{\tilde{\mathrm{Z}}_l}(z_l) \star \mathcal{T}_{l=1}^{M} \mu_{\tilde{\mathrm{W}}_l}(w_l) \Bigg/ \frac{\sum_{l=1}^{M} w_l z_l}{\sum_{l=1}^{M} w_l} \tag{2.75}$$

where $\mathcal{T}$ and $\star$ both indicate the t-norm used ... product or minimum, $w_l \in \tilde{\mathrm{W}}_l$ and $z_l \in \tilde{\mathrm{Z}}_l$ for $l = 1, \ldots, M$.

Theorem 2.5 *If each $\tilde{Z}_l$ is a Gaussian type-1 set, with mean m_l and standard deviation σ_l, and if each $\tilde{W}_l$ is also a Gaussian type-1 set with mean h_l and standard deviation Δ_l, then $\tilde{Y}$ is approximately a Gaussian type-1 set, with mean $\mathcal{M}$ and standard deviation Σ, where*

$$\mathcal{M} = \frac{\sum_{l=1}^{M} h_l m_l}{\sum_{l=1}^{M} h_l} \tag{2.76}$$

and

$$\Sigma = \begin{cases} \dfrac{\sqrt{\sum_{l=1}^{M} \left[(h_l \sigma_l)^2 + [m_l - \mathcal{M}]^2 \Delta_l^2\right]}}{\sum_{l=1}^{M} h_l} & , \text{ if product } t\text{-norm is used} \\ \dfrac{\sum_{l=1}^{M} [(h_l \sigma_l) + |m_l - \mathcal{M}| \Delta_l]}{\sum_{l=1}^{M} h_l} & , \text{ if minimum } t\text{-norm is used} \end{cases} \tag{2.77}$$

provided that

$$\frac{k \sum_{l=1}^{M} \Delta_l}{\sum_{l=1}^{M} h_l} \ll 1 , \tag{2.78}$$

where k is the number of standard deviations of a Gaussian considered significant (generally, $k = 2$ or 3). The Gaussian approximation improves as $k\left(\sum_{l=1}^{M} \Delta_l \Big/ \sum_{l=1}^{M} h_l\right)$ grows smaller, and the result is exact when $\sum_{l=1}^{M} \Delta_l = 0$, i.e., when $\Delta_l = 0$ for $l = 1, \ldots, M$. □

See Appendix C.11 for the proof. A sufficient condition that satisfies (2.78) is that the Gaussian $\tilde{\mathrm{W}}_l$'s are narrow, i.e., $k\Delta_l / h_l \ll 1$ for $l = 1, \ldots, M$. Observe, however, that there is no condition on the standard deviations of the $\tilde{Z}_l$'s; consequently, when all the $\tilde{W}_l$'s are crisp numbers, the theorem gives an exact result. See the comments at the end of Appendix C.11 for bounds on the domain of $\tilde{Y}$.

Recall that we will use only minimum t-norm for the centroid calculation of a type-2 set with a continuous domain. (If the domain is discrete, however, product t-norm may be used.) From Theorem 2.5, we get the following result for the centroid of a Gaussian type-2 set.

Corollary 2.2 *The centroid of a Gaussian type-2 set $\tilde{\tilde{A}}$ is approximately a Gaussian type-1 set with mean $\mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}})$ [Eq. (2.79)] and standard deviation $\Sigma(\tilde{C}_{\tilde{\tilde{A}}})$ [Eq. (2.80)], if the standard deviations of the secondary memberships are small compared to their means, i.e., if (2.81) is satisfied.*

Proof : Observe that the x_i's in (2.69), which are crisp numbers, correspond to the z_l's in (2.75), the $\tilde{D}_i$'s in (2.69) $[\tilde{D}_i = \tilde{\mu}_{\tilde{\tilde{A}}}(x_i)]$ correspond to the $\tilde{W}_l$'s in (2.75), and the sum in (2.69) goes from 1 to N instead of from 1 to M. If we denote the mean and the standard deviation of $\tilde{\mu}_{\tilde{\tilde{A}}}(x_i)$ as $m(x_i)$ and $\sigma(x_i)$, respectively, then using Theorem 2.5, $\tilde{C}_{\tilde{\tilde{A}}}$ is approximately a Gaussian type-1 set with mean $\mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}})$ and standard deviation $\Sigma(\tilde{C}_{\tilde{\tilde{A}}})$, where

$$\mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}}) = \frac{\sum_{i=1}^{N} x_i m(x_i)}{\sum_{i=1}^{N} m(x_i)} \tag{2.79}$$

and

$$\Sigma(\tilde{C}_{\tilde{\tilde{A}}}) = \frac{\sum_{i=1}^{N} \left|x_i - \mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}})\right| \sigma(x_i)}{\sum_{i=1}^{N} m(x_i)} \tag{2.80}$$

provided that

$$k\frac{\sum_{i=1}^{N} \sigma(x_i)}{\sum_{i=1}^{N} m(x_i)} \ll 1 \tag{2.81}$$

where k has the same meaning as in Theorem 2.5. Equation (2.81) is satisfied if standard deviations of the secondary membership functions are small compared to their means. □

Comment 1 : See Fig. 1.10 for an example of a type-2 set, which can be made to satisfy condition (2.81) easily. In this set the standard deviation of every membership grade is proportional to its mean. If we set the constant of proportionality to a small value, (2.81) can be satisfied. See Example 2.8 for an expression for the centroid of such a type-2 set.

Comment 2 : Because the membership grade of each $x \in \tilde{\tilde{A}}$ is a Gaussian type-1 set, the primary membership which has a secondary membership equal to unity is $m(x)$; and, since the principal membership function is the set of those primary memberships for which the secondary memberships are equal to 1, $m(x)$ *for $x \in \tilde{\tilde{A}}$ is the same as the principal membership function of $\tilde{\tilde{A}}$*. Observe, therefore, from (2.79), that the mean of the approximate centroid, $\mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}})$, corresponds to the centroid of the principal membership function of $\tilde{\tilde{A}}$.

Example 2.6 Consider the centroid calculation of a type-2 set [see (2.69)]. If the type-2 set is discretized into N points $(x_1, \ldots, x_N)$ and if the membership grade of every x_i is discretized into M points, the total number of possible $\{\theta_1, \ldots, \theta_N\}$ combinations is M^N. This number can be very large even for modest values of M and N, e.g., if $N = 10$ and $M = 5$, the number of possible combinations is $9,765,625$, i.e., about 10 million ! And for each of these combinations, we have to compute the weighted average $\sum_{i=1}^{N} x_i\theta_i \big/ \sum_{i=1} \theta_i$. On the other hand, if (2.81) is satisfied, all we have to do to compute the centroid is compute two weighted averages, one for the mean of the centroid [(2.79)] and one for its standard deviation [(2.80)]. This example demonstrates the significance of our Gaussian approximation results. □

Example 2.7 Now, we demonstrate the use of Corollary 2.2 with an example. Consider a Gaussian type-2 set with a discrete domain consisting of only 3 points, $x_1 = 1$, $x_2 = 3$ and $x_3 = 5$ [see Fig. 2.18 (a)]. Suppose that $m(x_1) = 0.1$, $m(x_2) = 0.8$ and $m(x_3) = 0.6$. We consider three cases :

1. If $\sigma(x_i) = 0.05m(x_i)$ for $i = 1, 2, 3$, the membership grades of x_1, x_2 and x_3 are shown in Fig. 2.18 (b), (c) and (d), respectively; and, the true centroid and the approximation in Corollary 2.2 are as shown in Fig. 2.19 (a). In this case, when $k = 2$, $k\Big[\sum_i \sigma(x_i)\Big]\Big/\Big[\sum_i m(x_i)\Big] = 0.1$.

2. If $\sigma(x_1) = 0.3m(x_1)$, $\sigma(x_2) = 0.1m(x_2)$ and $\sigma(x_3) = 0.2m(x_3)$, the membership grades of x_1, x_2 and x_3 are shown in Fig. 2.18 (e), (f) and (g), respectively; and, the true centroid and the approximation in Corollary 2.2 are as shown in Fig. 2.19 (b). In this case, when $k = 2$, $k\Big[\sum_i \sigma(x_i)\Big]\Big/\Big[\sum_i m(x_i)\Big] = 0.3066$.

3. If $\sigma(x_i) = 0.5m(x_i)$ for $i = 1, 2, 3$, the membership grades of x_1, x_2 and x_3 are shown in Fig. 2.18 (h), (i) and (j), respectively; and, the true centroid and the approximation in Corollary 2.2 are as shown in Fig. 2.19 (c). In this case, when $k = 2$, $k\Big[\sum_i \sigma(x_i)\Big]\Big/\Big[\sum_i m(x_i)\Big] = 1$.

When computing the true centroids, only primary membership values between $m(x_i) \pm 2\sigma(x_i)$ were considered. Observe that though the domain of the type-2 set is discrete, that of its centroid is continuous, because the membership grades of x_1, x_2 and x_3 have continuous domains.

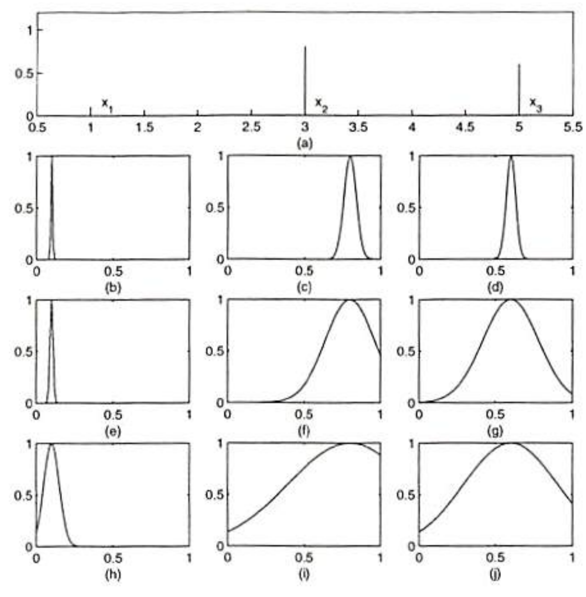


Figure 2.18: Figures for Example 2.7. The domain of the discrete Gaussian type-2 set having 3 points, $x_1 = 1$, $x_2 = 3$ and $x_3 = 5$, is depicted in (a). The membership grades of x_1, x_2 and x_3 for case 1 are depicted in (b), (c) and (d), respectively ; the membership grades for case 2 are depicted in (e), (f) and (g); and, those for case 3 are depicted in (h), (i) and (j). Each of the figures (b) to (j) show plots of primary versus secondary memberships. In each case, $m(x_1) = 0.1$, $m(x_2) = 0.8$ and $m(x_3) = 0.6$. For case 1, $\sigma(x_i) = 0.05m(x_i)$ for $i = 1, 2, 3$; for case 2, $\sigma(x_1) = 0.3m(x_1)$, $\sigma(x_2) = 0.1m(x_2)$ and $\sigma(x_3) = 0.2m(x_3)$; and, for case 3, $\sigma(x_i) = 0.5m(x_i)$ for $i = 1, 2, 3$.

Observe that the approximation in the first two cases is much closer to the true centroid than that in the third case; however, though a smaller value for $\Big[\sum_i \sigma(x_i)\Big]\Big/\Big[\sum_i m(x_i)\Big]$ will generally

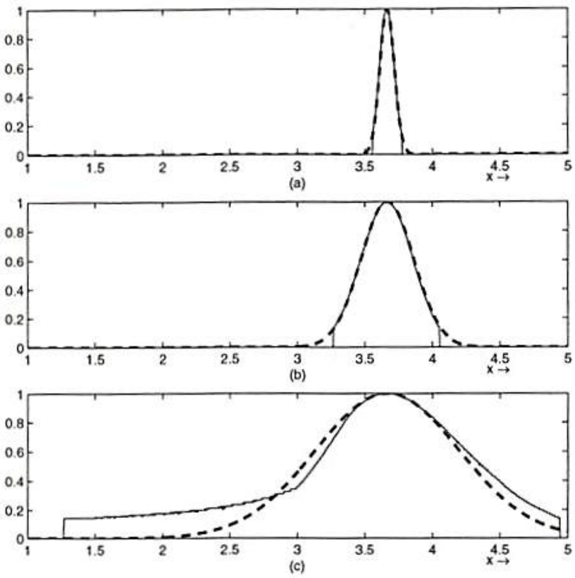


Figure 2.19: Figures for Example 2.7. Centroids of the Gaussian type-2 set depicted in Fig. 2.18 for the three choices of $\sigma(x_i)$ $(i = 1, 2, 3)$. (a) $\sigma(x_i) = 0.05m(x_i)$ for $i = 1, 2, 3$. (b) $\sigma(x_1) = 0.3m(x_1)$, $\sigma(x_2) = 0.1m(x_2)$ and $\sigma(x_3) = 0.2m(x_3)$. (c) $\sigma(x_i) = 0.5m(x_i)$ for $i = 1, 2, 3$. When computing the true centroids only primary membership values between $m(x_i) \pm 2\sigma(x_i)$ are considered.

give a better approximation, it is not at all easy to predict how close the actual centroid of a given Gaussian type-2 set will be to its approximation. The same can be said about the approximation in Theorem 2.5, which allows the domain points x_i's to be replaced by fuzzy sets. □

Example 2.8 Consider a Gaussian type-2 set $\tilde{\tilde{A}} \subset X$. Let the principal membership function of $\tilde{\tilde{A}}$ be a Gaussian type-1 set with mean M and standard deviation Σ; and, let the standard deviation of each secondary membership function of $\tilde{\tilde{A}}$ be proportional to the mean of that secondary membership function. Figure 1.10 shows an example of such a Gaussian type-2 set. In this example, we obtain an expression for the centroid of $\tilde{\tilde{A}}$, using Corollary 2.2.

Recall, from comment 2 at the end of Corollary 2.2, that $m(x)$ *for* $x \in X$ *is the same as the principal membership function of* $\tilde{\tilde{A}}$. The membership grade of every $x \in X$ in $\tilde{\tilde{A}}$ can, therefore, be described as

$$\tilde{\mu}_{\tilde{\tilde{A}}}(x) = G\left(m(x), \sigma(x)\right) \quad ; \quad x \in X \tag{2.82}$$

where

$$m(x) = e^{-\frac{1}{2}\left(\frac{x-M}{\Sigma}\right)^2} \tag{2.83}$$

and

$$\sigma(x) = cm(x) \tag{2.84}$$

where $G(m, \sigma)$ indicates a Gaussian with mean m and standard deviation σ; and, c is a constant, which is generally in $(0, 1)$.

Let us now find an expression for $\tilde{C}_{\tilde{\tilde{A}}}$, the centroid of $\tilde{\tilde{A}}$, in terms of M, Σ and c. From Corollary 2.2, we know that $\tilde{C}_{\tilde{\tilde{A}}}$ is approximately a Gaussian type-1 set with mean $\mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}})$ and standard deviation $\Sigma(\tilde{C}_{\tilde{\tilde{A}}})$, where

$$\mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}}) = \frac{\sum_{i=1}^{N} x_i m(x_i)}{\sum_{i=1}^{N} m(x_i)} \tag{2.85}$$

and

$$\Sigma(\tilde{C}_{\tilde{\tilde{A}}}) = \frac{\sum_{i=1}^{N} \left| x_i - \mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}}) \right| \sigma(x_i)}{\sum_{i=1}^{N} m(x_i)} \tag{2.86}$$

provided that

$$k \frac{\sum_{i=1}^{N} \sigma(x_i)}{\sum_{i=1}^{N} m(x_i)} \ll 1 \tag{2.87}$$

where k is the number of standard deviations considered significant, 2 or 3, and the domain of $\tilde{\tilde{A}}$ is assumed to be discretized into N points.

Since $m(x)$ for $x \in X$ is the principal membership function of $\tilde{\tilde{A}}$, we see, from (2.85), that $\mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}})$ is the same as the centroid of the principal membership function, which is equal to M, i.e.,

$$\mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}}) = M \tag{2.88}$$

To find $\Sigma(\tilde{C}_{\tilde{\tilde{A}}})$, let us assume that X is not discretized (i.e., it is continous), so that (2.86) can be rewritten as (using the fact that $\mathcal{M}(\tilde{C}_{\tilde{\tilde{A}}}) = M$)

$$\Sigma(\tilde{C}_{\tilde{\tilde{A}}}) = \frac{\int_{x \in X} |x - M| \sigma(x) dx}{\int_{x \in X} m(x) dx} \tag{2.89}$$

The fuzzy sets we deal with are generally subsets of the real line, so that $X = \Re$. Observe, from (2.83) and (2.89), that the denominator of (2.89) is the area under a Gaussian with mean M and standard deviation Σ. Recall from probability theory that the area under a probability density function is unity; therefore,

$$\int_{x \in \Re} m(x) dx = \sqrt{2\pi}\Sigma \tag{2.90}$$

Let the numerator of (2.89) be equal to I. It can be computed as follows.

$$I = I_1 + I_2 \tag{2.91}$$

where

$$I_1 = \int_{x \le M} (M - x) \sigma(x) dx \tag{2.92}$$

and

$$I_2 = \int_{x > M} (x - M) \sigma(x) dx \tag{2.93}$$

Substituting (2.83), (2.84) and $\frac{1}{2}[(x - M)/\Sigma]^2 = t$ into (2.92) and (2.93), it is easy to see that

$$I_1 = I_2 = c\Sigma^2 \int_0^\infty e^{-t} dt = c\Sigma^2 \tag{2.94}$$

Using (2.90), (2.91) and (2.94) in (2.89), we find that

$$\Sigma(\tilde{C}_{\tilde{\tilde{A}}}) = \frac{2c\Sigma^2}{\sqrt{2\pi}\Sigma} = \sqrt{\frac{2}{\pi}}c\Sigma \tag{2.95}$$

From (2.88) and (2.95), we see that $\tilde{C}_{\tilde{\tilde{A}}}$ is approximately a Gaussian type-1 set with mean M and standard deviation $\sqrt{2/\pi}c\Sigma$. The condition (2.87) requires that $c \ll 1/k$ [since $\sigma(x) = cm(x)$].

Observe that, if we set $c = 0$, $\sigma(x) = 0$ for all $x \in X$, implying that $\tilde{\tilde{A}}$ collapses to its principal membership. Now, from (2.95) , we get $\Sigma(\tilde{C}_{\tilde{\tilde{A}}}) = 0$, which means that the centroid of $\tilde{\tilde{A}}$ collapses to a single point, equal to M. This is consistent with the fact that M is the centroid of the principal membership function of $\tilde{\tilde{A}}$. □

2.5.2.2 Centroid of an Interval Type-2 Set

Appendix D.1 describes a computational procedure to compute the exact result of a weighted average of interval type-1 sets. This procedure can be used to compute the centroid of an interval type-2 set. Appendix D.2 also gives a result similar to Theorem 2.5.

2.5.2.3 Centroid of a Triangular Type-2 Set

For a result similar to Theorem 2.5 for triangular type-2 sets, see [8].

Chapter 3

Properties of Membership Grades

Mizumoto and Tanaka discuss the properties of membership grades of type-2 fuzzy sets in great detail in [17]. They examine the following properties which are satisfied by membership grades of type-1 fuzzy sets : ($\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ are type-1 fuzzy subsets of a set X; and $\vee$ and $\wedge$ represent *max* and *min* respectively.)

$$\begin{aligned}
&\textit{Reflexive} \text{ law} &:\quad & \mu_{\tilde{A}} \le \mu_{\tilde{A}} && (3.1)\\
&\textit{Antisymmetric} \text{ law} &:\quad & \mu_{\tilde{A}} \le \mu_{\tilde{B}},\ \mu_{\tilde{B}} \le \mu_{\tilde{A}} \Rightarrow \mu_{\tilde{A}} = \mu_{\tilde{B}} && (3.2)\\
&\textit{Transitive} \text{ law} &:\quad & \mu_{\tilde{A}} \le \mu_{\tilde{B}},\ \mu_{\tilde{B}} \le \mu_{\tilde{C}} \Rightarrow \mu_{\tilde{A}} \le \mu_{\tilde{C}} && (3.3)\\
&\textit{Idempotent} \text{ laws} &:\quad & \mu_{\tilde{A}} \vee \mu_{\tilde{A}} = \mu_{\tilde{A}} &&\\
& & & \mu_{\tilde{A}} \wedge \mu_{\tilde{A}} = \mu_{\tilde{A}} && (3.4)\\
&\textit{Commutative} \text{ laws} &:\quad & \mu_{\tilde{A}} \vee \mu_{\tilde{B}} = \mu_{\tilde{B}} \vee \mu_{\tilde{A}} &&\\
& & & \mu_{\tilde{A}} \wedge \mu_{\tilde{B}} = \mu_{\tilde{B}} \wedge \mu_{\tilde{A}} && (3.5)\\
&\textit{Associative} \text{ laws} &:\quad & (\mu_{\tilde{A}} \vee \mu_{\tilde{B}}) \vee \mu_{\tilde{C}} = \mu_{\tilde{A}} \vee (\mu_{\tilde{B}} \vee \mu_{\tilde{C}}) &&\\
& & & (\mu_{\tilde{A}} \wedge \mu_{\tilde{B}}) \wedge \mu_{\tilde{C}} = \mu_{\tilde{A}} \wedge (\mu_{\tilde{B}} \wedge \mu_{\tilde{C}}) && (3.6)\\
&\textit{Absorption} \text{ laws} &:\quad & \mu_{\tilde{A}} \wedge (\mu_{\tilde{A}} \vee \mu_{\tilde{B}}) = \mu_{\tilde{A}} &&\\
& & & \mu_{\tilde{A}} \vee (\mu_{\tilde{A}} \wedge \mu_{\tilde{B}}) = \mu_{\tilde{A}} && (3.7)\\
&\textit{Distributive} \text{ laws} &:\quad & \mu_{\tilde{A}} \wedge (\mu_{\tilde{B}} \vee \mu_{\tilde{C}}) &&\\
& & & \quad = (\mu_{\tilde{A}} \wedge \mu_{\tilde{B}}) \vee (\mu_{\tilde{A}} \wedge \mu_{\tilde{C}}) &&\\
& & & \mu_{\tilde{A}} \vee (\mu_{\tilde{B}} \wedge \mu_{\tilde{C}}) &&\\
& & & \quad = (\mu_{\tilde{A}} \vee \mu_{\tilde{B}}) \wedge (\mu_{\tilde{A}} \vee \mu_{\tilde{C}}) && (3.8)\\
&\textit{Involution} \text{ law} &:\quad & \mu_{\overline{\overline{\tilde{A}}}} = \mu_{\tilde{A}} && (3.9)\\
&\textit{De Morgan's} \text{ laws} &:\quad & \overline{\mu_{\tilde{A}} \vee \mu_{\tilde{B}}} = \mu_{\overline{\tilde{A}}} \wedge \mu_{\overline{\tilde{B}}} &&\\
& & & \overline{\mu_{\tilde{A}} \wedge \mu_{\tilde{B}}} = \mu_{\overline{\tilde{A}}} \vee \mu_{\overline{\tilde{B}}} && (3.10)\\
&\textit{Identity} \text{ laws} &:\quad & \mu_{\tilde{A}} \vee 0 = \mu_{\tilde{A}}, \mu_{\tilde{A}} \wedge 1 = \mu_{\tilde{A}} &&\\
& & & \mu_{\tilde{A}} \vee 1 = 1, \mu_{\tilde{A}} \wedge 0 = 0 && (3.11)\\
&\text{Failure of } \textit{complement} \text{ laws} &:\quad & \mu_{\tilde{A}} \vee \mu_{\overline{\tilde{A}}} \ne 1 &&\\
& & & \mu_{\tilde{A}} \wedge \mu_{\overline{\tilde{A}}} \ne 0 && (3.12)
\end{aligned}$$

For type-2 fuzzy sets, membership grades are type-1 fuzzy sets supported in $[0,1]$. Mizumoto and Tanaka show that *normal convex* type-1 fuzzy grades satisfy all the aforementioned laws for *min* t-norm and *max* t-conorm. (Non-normal and non-convex grades do satisfy some laws. See [17] for details.) We focus on the *product* t-norm and examine which of the above properties are satisfied by the type-1 fuzzy grades (membership grades of type-2 sets). We consider only *convex, normal* fuzzy grades in our work. The results are summarized in Table 3.1.

Table 3.1: This table summarizes the results of Chapter 3. In the type-2 case, we assume *convex, normal* membership grades. Additionally, for *reflexive, antisymmetric* and *transitive* laws, fuzzy set inclusion is defined as in (3.26). For product t-norm, the laws that are satisfied in the type-1 case, but not in the type-2 case, are highlighted. Results for the minimum t-norm are taken from [17].

Set Theoretic Laws		**Minimum t-norm**		**Product t-norm**	
		Type-1	**Type-2**	**Type-1**	**Type-2**
Reflexive	$\mu_{\tilde{A}} \leq \mu_{\tilde{A}}$	Yes	Yes	Yes	Yes
Anti-symmetric	$\mu_{\tilde{A}} \leq \mu_{\tilde{B}},\ \mu_{\tilde{B}} \leq \mu_{\tilde{A}} \Rightarrow \mu_{\tilde{A}} = \mu_{\tilde{B}}$	Yes	Yes	Yes	Yes
Transitive	$\mu_{\tilde{A}} \leq \mu_{\tilde{B}},\ \mu_{\tilde{B}} \leq \mu_{\tilde{C}} \Rightarrow \mu_{\tilde{A}} \leq \mu_{\tilde{C}}$	Yes	Yes	Yes	Yes
Idempotent	$\mu_{\tilde{A}} \vee \mu_{\tilde{A}} = \mu_{\tilde{A}}$	Yes	Yes	Yes	**No**
	$\mu_{\tilde{A}} \star \mu_{\tilde{A}} = \mu_{\tilde{A}}$	Yes	Yes	No	No
Commutative	$\mu_{\tilde{A}} \vee \mu_{\tilde{B}} = \mu_{\tilde{B}} \vee \mu_{\tilde{A}}$	Yes	Yes	Yes	Yes
	$\mu_{\tilde{A}} \star \mu_{\tilde{B}} = \mu_{\tilde{B}} \star \mu_{\tilde{A}}$	Yes	Yes	Yes	Yes
Associative	$(\mu_{\tilde{A}} \vee \mu_{\tilde{B}}) \vee \mu_{\tilde{C}} = \mu_{\tilde{A}} \vee (\mu_{\tilde{B}} \vee \mu_{\tilde{C}})$	Yes	Yes	Yes	Yes
	$(\mu_{\tilde{A}} \star \mu_{\tilde{B}}) \star \mu_{\tilde{C}} = \mu_{\tilde{A}} \star (\mu_{\tilde{B}} \star \mu_{\tilde{C}})$	Yes	Yes	Yes	Yes
Absorption	$\mu_{\tilde{A}} \star (\mu_{\tilde{A}} \vee \mu_{\tilde{B}}) = \mu_{\tilde{A}}$	Yes	Yes	No	No
	$\mu_{\tilde{A}} \vee (\mu_{\tilde{A}} \star \mu_{\tilde{B}}) = \mu_{\tilde{A}}$	Yes	Yes	Yes	**No**
Distributive	$\mu_{\tilde{A}} \star (\mu_{\tilde{B}} \vee \mu_{\tilde{C}}) = (\mu_{\tilde{A}} \star \mu_{\tilde{B}}) \vee (\mu_{\tilde{A}} \star \mu_{\tilde{C}})$	Yes	Yes	Yes	**No**
	$\mu_{\tilde{A}} \vee (\mu_{\tilde{B}} \star \mu_{\tilde{C}}) = (\mu_{\tilde{A}} \vee \mu_{\tilde{B}}) \star (\mu_{\tilde{A}} \vee \mu_{\tilde{C}})$	Yes	Yes	No	No
Involution	$\mu_{\overline{\overline{\tilde{A}}}} = \mu_{\tilde{A}}$	Yes	Yes	Yes	Yes
De Morgan's Laws	$\overline{\mu_{\tilde{A}} \vee \mu_{\tilde{B}}} = \mu_{\overline{\tilde{A}}} \star \mu_{\overline{\tilde{B}}}$	Yes	Yes	No	No
	$\overline{\mu_{\tilde{A}} \star \mu_{\tilde{B}}} = \mu_{\overline{\tilde{A}}} \vee \mu_{\overline{\tilde{B}}}$	Yes	Yes	No	No
Identity	$\mu_{\tilde{A}} \vee 0 = \mu_{\tilde{A}}$	Yes	Yes	Yes	Yes
	$\mu_{\tilde{A}} \star 1 = \mu_{\tilde{A}}$	Yes	Yes	Yes	Yes
	$\mu_{\tilde{A}} \vee 1 = 1$	Yes	Yes	Yes	Yes
	$\mu_{\tilde{A}} \star 0 = 0$	Yes	Yes	Yes	Yes
Complement (Failure)	$\mu_{\tilde{A}} \vee \mu_{\overline{\tilde{A}}} \neq 1$	Yes	Yes	Yes	Yes
	$\mu_{\tilde{A}} \star \mu_{\overline{\tilde{A}}} \neq 0$	Yes	Yes	Yes	Yes

Before going to type-2 sets, we review which of the above properties are satisfied by membership grades of type-1 sets for *product* t-norm and *max* t-conorm. Again consider type-1 fuzzy subsets, $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, of set X, as above.

Reflexive, *antisymmetric* and *transitive* laws do not make use of the t-norm at all and hence are satisfied for product t-norm and max t-conorm. *Commutative* and *associative* laws are also satisfied, since both, max and product operations are commutative and associative, i.e.,

$$\mu_{\tilde{A}} \vee \mu_{\tilde{B}} = \mu_{\tilde{B}} \vee \mu_{\tilde{A}} \tag{3.13}$$

$$\mu_{\tilde{A}} \mu_{\tilde{B}} = \mu_{\tilde{B}} \mu_{\tilde{A}} \tag{3.14}$$

$$(\mu_{\tilde{A}} \vee \mu_{\tilde{B}}) \vee \mu_{\tilde{C}} = \mu_{\tilde{A}} \vee (\mu_{\tilde{B}} \vee \mu_{\tilde{C}}) \tag{3.15}$$
$$(\mu_{\tilde{A}}\mu_{\tilde{B}})\mu_{\tilde{C}} = \mu_{\tilde{A}}(\mu_{\tilde{B}}\mu_{\tilde{C}}) \tag{3.16}$$

The second part of *absorption* laws is satisfied, i.e., $\mu_{\tilde{A}} \vee (\mu_{\tilde{A}}\mu_{\tilde{B}}) = \mu_{\tilde{A}}$; the first part of *distributive* laws is satisfied, i.e., product is distributive over maximum; and the first part of *idempotent* laws is also satisfied, i.e., $\mu_{\tilde{A}} \vee \mu_{\tilde{A}} = \mu_{\tilde{A}}$.

The *involution* law is satisfied, since we define complement in the same way as before, $\mu_{\overline{\tilde{A}}} = 1 - \mu_{\tilde{A}}$. Additionally, *identity* laws are satisfied, i.e.,

$$\begin{aligned} \mu_{\tilde{A}} \vee 0 = \mu_{\tilde{A}} \quad &, \quad \mu_{\tilde{A}} \times 1 = \mu_{\tilde{A}} \\ \mu_{\tilde{A}} \vee 1 = 1 \quad &, \quad \mu_{\tilde{A}} \times 0 = 0 \end{aligned} \tag{3.17}$$

None of the other laws are satisfied under product t-norm. We list the reasons for failure of these laws as follows :

$$\textit{Idempotent}\text{ laws (second part)} : \mu_{\tilde{A}}\mu_{\tilde{A}} \neq \mu_{\tilde{A}} \tag{3.18}$$
$$\begin{aligned} \textit{Absorption}\text{ laws (first part)} : \; & \mu_{\tilde{A}} > \mu_{\tilde{B}} \Rightarrow \\ & \mu_{\tilde{A}}(\mu_{\tilde{A}} \vee \mu_{\tilde{B}}) = \mu_{\tilde{A}}\mu_{\tilde{A}} \neq \mu_{\tilde{A}} \end{aligned} \tag{3.19}$$
$$\begin{aligned} \textit{Distributive}\text{ laws (second part)} : \; & \mu_{\tilde{A}} > \mu_{\tilde{B}}, \mu_{\tilde{A}} > \mu_{\tilde{C}} \Rightarrow \\ & \mu_{\tilde{A}} \vee (\mu_{\tilde{B}}\mu_{\tilde{C}}) = \mu_{\tilde{A}} \\ & \neq (\mu_{\tilde{A}} \vee \mu_{\tilde{B}})(\mu_{\tilde{A}} \vee \mu_{\tilde{C}}) = \mu_{\tilde{A}}^2 \end{aligned} \tag{3.20}$$
$$\begin{aligned} \textit{De Morgan's}\text{ laws} : \; & \overline{\mu_{\tilde{A}} \vee \mu_{\tilde{B}}} = 1 - (\mu_{\tilde{A}} \vee \mu_{\tilde{B}}) \\ & \neq \mu_{\overline{\tilde{A}}}\mu_{\overline{\tilde{B}}} = (1 - \mu_{\tilde{A}})(1 - \mu_{\tilde{B}}) \end{aligned} \tag{3.21}$$
$$\begin{aligned} & \overline{\mu_{\tilde{A}}\mu_{\tilde{B}}} = 1 - \mu_{\tilde{A}}\mu_{\tilde{B}} \\ & \neq \mu_{\overline{\tilde{A}}} \vee \mu_{\overline{\tilde{B}}} \\ & = \max\{(1 - \mu_{\tilde{A}}), (1 - \mu_{\tilde{B}})\} \end{aligned} \tag{3.22}$$

Complement laws fail in both the cases, for minimum [17] as well as product t-norms.

From our discussion so far, we see that the {max, product} t-conorm / t-norm pair, for type-1 sets, does not satisfy *complement* laws, *De Morgan's* laws and parts of *idempotent*, *absorption* and *distributive* laws (our discussion is summarized in Table 3.1). Note that this has in no way deterred users from developing and applying type-1 fuzzy logic systems using the {max, product} t-conorm / t-norm pair, because, for the most part, we don't make use of these laws in our fuzzy logic systems. If, however, the designer of a type-1 FLS did use these laws, then the design would be in error, since the laws do not apply.

Now, let's turn to type-2 sets. We use *join* and *meet* operations in place of the t-conorm and t-norms respectively, in the above definitions. The underlying t-conorm and t-norm are maximum and product, respectively.

Recall that all our operations on type-2 sets collapse to their type-1 counterparts, i.e., if we replace all the type-1 fuzzy grades by the primary memberships at which the secondary memberships are equal to 1 the results remain valid. We assume, for simplicity, that the secondary membership functions reach the value 1 at only one point. This means that we can think of type-1 fuzzy sets as special cases of type-2 fuzzy sets, where only one value of the primary memberships has a non-zero secondary membership (and this secondary membership is equal to 1); therefore, *if there are any set theoretic laws that are not satisfied by type-1 fuzzy sets, we can safely say that type-2 sets will not satisfy those laws either*; however, the converse of this statement may not be true. *If any condition is satisfied by type-1 sets, it may or may not be satisfied by type-2 sets.* So, next we examine only those aforementioned properties that are satisfied in the type-1 case.

Consider three type-2 fuzzy subsets, $\tilde{\tilde{\mathrm{P}}}$, $\tilde{\tilde{\mathrm{Q}}}$, $\tilde{\tilde{\mathrm{R}}}$, of a set X, with membership grades as follows :

$$\tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} = \int_{u\in[0,1]} f(u)/u \tag{3.23}$$

$$\tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} = \int_{v\in[0,1]} g(v)/v \tag{3.24}$$

$$\tilde{\mu}_{\tilde{\tilde{\mathrm{R}}}} = \int_{w\in[0,1]} h(w)/w \tag{3.25}$$

We define fuzzy set inclusion as follows [12] :

$$\tilde{\tilde{\mathrm{P}}} \subseteq \tilde{\tilde{\mathrm{Q}}} \Leftrightarrow \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}}(x) \leq \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}}(x) \;\; \forall x \in X \tag{3.26}$$

The following generalized versions of *reflexive*, *antisymmetric* and *transitive* laws are satisfied. Observe that these laws do not make use of the t-norm or t-conorm.

$$\textit{Reflexive}\text{ law} : \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \subseteq \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \tag{3.27}$$

$$\textit{Antisymmetric}\text{ law} : \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \subseteq \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}}\,,\, \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} \subseteq \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \Rightarrow \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} = \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} \tag{3.28}$$

$$\textit{Transitive}\text{ law} : \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \subseteq \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}}\,,\, \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} \subseteq \tilde{\mu}_{\tilde{\tilde{\mathrm{R}}}} \Rightarrow \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \subseteq \tilde{\mu}_{\tilde{\tilde{\mathrm{R}}}} \tag{3.29}$$

where $\tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} = \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} \Leftrightarrow f(u) = g(u)\, \forall u \in [0,1]$ [see Eqs. (3.28), (3.23) and (3.24)].

Commutative laws are satisfied for max t-conorm and product t-norm, because :

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \sqcup \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} &= [\int_u f(u)/u] \sqcup [\int_v g(v)/v] \\
&= \int_u \int_v f(u)g(v)/(u \vee v) \\
&= \int_v \int_u g(v)f(u)/(v \vee u) \\
&= [\int_v g(v)/v] \sqcup [\int_u f(u)/u] \\
&= \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} \sqcup \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}}
\end{aligned} \tag{3.30}$$

A very similar proof can be presented for $\tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \sqcap \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} = \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} \sqcap \tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}}$.

Associative laws are satisfied for max t-conorm and product t-norm, because :

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \sqcap (\tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} \sqcap \tilde{\mu}_{\tilde{\tilde{\mathrm{R}}}}) &= [\int_u f(u)/u] \sqcap [\int_v \int_w g(v)h(w)/vw] \\
&= \int_u \int_v \int_w f(u)[g(v)h(w)]/u(vw) \\
&= \int_u \int_v \int_w [f(u)g(v)]h(w)/(uv)w \\
&= [\int_u \int_v f(u)g(v)/uv] \sqcap [\int_w h(w)/w] \\
&= (\tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \sqcap \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}}) \sqcap \tilde{\mu}_{\tilde{\tilde{\mathrm{R}}}}
\end{aligned} \tag{3.31}$$

A similar proof can be presented for $\tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \sqcup (\tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}} \sqcup \tilde{\mu}_{\tilde{\tilde{\mathrm{R}}}}) = (\tilde{\mu}_{\tilde{\tilde{\mathrm{P}}}} \sqcup \tilde{\mu}_{\tilde{\tilde{\mathrm{Q}}}}) \sqcup \tilde{\mu}_{\tilde{\tilde{\mathrm{R}}}}$. The proofs of (3.30) and (3.31) are very similar to those in [17].

The proof for the *involution* law, given in [17], remains unchanged because we do not change the definition of *negation* when we change from min t-norm to product t-norm, i.e., $\tilde{\mu}_{\bar{\bar{\mathrm{P}}}} = \neg\tilde{\mu}_{\bar{\mathrm{P}}}$.

Before proving the *identity* laws we explain the concepts of 0 and 1 membership grades in case of type-2 sets [17]. In type-2 sets, 0 and 1 memberships are represented as 1/0 and 1/1 respectively. This is in accordance with our earlier discussion about type-1 sets being a special case of type-2 sets. Zero membership in type-2 sets is equivalent to having a "zero" fuzzy set of type-1 (since the membership grade is a fuzzy set of type-1), i.e., an element is said to have a zero membership in a type-2 set if it has a secondary membership equal to 1 corresponding to the primary membership of 0, and if it has all other secondary memberships equal to 0. Similarly, an element is said to have a membership grade equal to 1 in a type-2 set, if it has a secondary membership equal to 1 corresponding to the primary membership of 1, and if all other secondary memberships are zero. Now, we proceed to prove the *identity* laws.

Identity laws are satisfied for product t-norm and max t-conorm, if we use only normalized membership grades, because :

$$\begin{aligned}
\tilde{\mu}_{\bar{\mathrm{P}}} \sqcup 0 &= [\int_u f(u)/u] \sqcup 1/0 \\
&= \int_u f(u) \times 1/(u \vee 0) \\
&= \int_u f(u)/u \\
&= \tilde{\mu}_{\bar{\mathrm{P}}} \qquad (3.32)
\end{aligned}$$

$$\begin{aligned}
\tilde{\mu}_{\bar{\mathrm{P}}} \sqcap 1 &= [\int_u f(u)/u] \sqcap 1/1 \\
&= \int_u f(u) \times 1/(u \times 1) \\
&= \int_u f(u)/u \\
&= \tilde{\mu}_{\bar{\mathrm{P}}} \qquad (3.33)
\end{aligned}$$

$$\begin{aligned}
\tilde{\mu}_{\bar{\mathrm{P}}} \sqcup 1 &= [\int_u f(u)/u] \sqcup 1/1 \\
&= \int_u f(u) \times 1/(u \vee 1) \\
&= \int_u f(u)/1 \\
&= [\sup_u f(u)]/1 \\
&= 1/1 \\
&= 1 \qquad (3.34)
\end{aligned}$$

$$\begin{aligned}
\tilde{\mu}_{\bar{\mathrm{P}}} \sqcap 0 &= [\int_u f(u)/u] \sqcap 1/0 \\
&= \int_u f(u) \times 1/(u \times 0) \\
&= \int_u f(u)/0 \\
&= [\sup_u f(u)]/0 \\
&= 1/0
\end{aligned}$$

$$= \quad 0 \tag{3.35}$$

We have made use of the fact that the secondary membership functions are normalized only in (3.34) and (3.35). We do not need to use this fact in the proofs of (3.32) and (3.33), which means that (3.32) and (3.33) hold even in the case of non-normalized secondary membership functions.

Next we show, by means of an example, that the first parts of the *idempotent* and *distributive* laws, and the second part of the *absorption* law, which are satisfied in the type-1 case, may not be satisfied for product t-norm and max t-conorm in the type-2 case.

Example 3.1 Consider the following three normalized, convex, type-1 fuzzy grades :

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{P}}} &= 0.5/0.1 + 1/0.7 \\
\tilde{\mu}_{\tilde{\tilde{Q}}} &= 0.6/0.3 + 1/0.7 \\
\tilde{\mu}_{\tilde{\tilde{R}}} &= 0.4/0.2 + 1/0.8
\end{aligned}$$

Failure of the first part of the *idempotent* law :

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{P}}} \sqcup \tilde{\mu}_{\tilde{\tilde{P}}} &= (0.5/0.1 + 1/0.7) \sqcup (0.5/0.1 + 1/0.7) \\
&= 0.25/0.1 + 0.5/0.7 + 0.5/0.7 + 1/0.7 \\
&= 0.25/0.1 + 1/0.7 \\
&\neq \tilde{\mu}_{\tilde{\tilde{P}}} \qquad (3.36)
\end{aligned}$$

Failure of the first part of the *distributive* law :

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{Q}}} \sqcup \tilde{\mu}_{\tilde{\tilde{R}}} &= (0.6/0.3 + 1/0.7) \sqcup (0.4/0.2 + 1/0.8) \\
&= 0.24/0.3 + 0.6/0.8 + 0.4/0.7 + 1/0.8 \\
&= 0.24/0.3 + 0.4/0.7 + 1/0.8 \qquad (3.37) \\
\tilde{\mu}_{\tilde{\tilde{P}}} \sqcap (\tilde{\mu}_{\tilde{\tilde{Q}}} \sqcup \tilde{\mu}_{\tilde{\tilde{R}}}) &= (0.5/0.1 + 1/0.7) \sqcap (0.24/0.3 + 0.4/0.7 + 1/0.8) \\
&= 0.12/0.03 + 0.2/0.07 + 0.5/0.08 \\
&\quad +0.24/0.21 + 0.4/0.49 + 1/0.56 \qquad (3.38) \\
\tilde{\mu}_{\tilde{\tilde{P}}} \sqcap \tilde{\mu}_{\tilde{\tilde{Q}}} &= (0.5/0.1 + 1/0.7) \sqcap (0.6/0.3 + 1/0.7) \\
&= 0.3/0.03 + 0.5/0.07 + 0.6/0.21 + 1/0.49 \qquad (3.39) \\
\tilde{\mu}_{\tilde{\tilde{P}}} \sqcap \tilde{\mu}_{\tilde{\tilde{R}}} &= (0.5/0.1 + 1/0.7) \sqcap (0.4/0.2 + 1/0.8) \\
&= 0.2/0.02 + 0.5/0.08 + 0.4/0.14 + 1/0.56 \qquad (3.40) \\
(\tilde{\mu}_{\tilde{\tilde{P}}} \sqcap \tilde{\mu}_{\tilde{\tilde{Q}}}) \sqcup (\tilde{\mu}_{\tilde{\tilde{P}}} \sqcap \tilde{\mu}_{\tilde{\tilde{R}}}) &= (0.3/0.03 + 0.5/0.07 + 0.6/0.21 + 1/0.49) \\
&\quad \sqcup(0.2/0.02 + 0.5/0.08 + 0.4/0.14 + 1/0.56) \\
&= (0.3/0.03) \sqcup (0.2/0.02 + 0.5/0.08 + 0.4/0.14 + 1/0.56) \\
&\quad +(0.5/0.07) \sqcup (0.2/0.02 + 0.5/0.08 + 0.4/0.14 + 1/0.56) \\
&\quad +(0.6/0.21) \sqcup (0.2/0.02 + 0.5/0.08 + 0.4/0.14 + 1/0.56) \\
&\quad +(1/0.49) \sqcup (0.2/0.02 + 0.5/0.08 + 0.4/0.14 + 1/0.56) \\
&= 0.06/0.03 + 0.15/0.08 + 0.12/0.14 + 0.3/0.56 \\
&\quad +0.1/0.07 + 0.25/0.08 + 0.2/0.14 + 0.5/0.56 \\
&\quad +0.12/0.21 + 0.3/0.21 + 0.24/0.21 + 0.6/0.56 \\
&\quad +0.2/0.49 + 0.5/0.49 + 0.4/0.49 + 1/0.56 \\
&= 0.06/0.03 + 0.1/0.07 + 0.25/0.08 + 0.2/0.14
\end{aligned}$$

$$+0.3/0.21 + 0.5/0.49 + 1/0.56$$
$$\neq \tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcap (\tilde{\mu}_{\tilde{\bar{\text{Q}}}} \sqcup \tilde{\mu}_{\tilde{\bar{\text{R}}}}) \tag{3.41}$$

This shows that under the product *t*-norm, even with convex, normal sets, *meet* is not distributive over *join*, i.e.,

$$\tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcap (\tilde{\mu}_{\tilde{\bar{\text{Q}}}} \sqcup \tilde{\mu}_{\tilde{\bar{\text{R}}}}) \neq (\tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcap \tilde{\mu}_{\tilde{\bar{\text{Q}}}}) \sqcup (\tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcap \tilde{\mu}_{\tilde{\bar{\text{R}}}}) \tag{3.42}$$

Failure of the second part of the *absorption* law : Using (3.39),

$$\begin{aligned}
\tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcup (\tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcap \tilde{\mu}_{\tilde{\bar{\text{Q}}}}) &= (0.5/0.1 + 1/0.7) \sqcup (0.3/0.03 + 0.5/0.07 + 0.6/0.21 + 1/0.49) \\
&= 0.15/0.1 + 0.25/0.1 + 0.3/0.21 + 0.5/0.49 \\
&\quad +0.3/0.7 + 0.5/0.7 + 0.6/0.7 + 1/0.7 \\
&= 0.25/0.1 + 0.3/0.21 + 0.5/0.49 + 1/0.7 \\
&\neq \tilde{\mu}_{\tilde{\bar{\text{P}}}}
\end{aligned} \tag{3.43}$$

□

Observe that both of the *distributive* and *absorption* laws do not hold. So, under product *t*-norm, *meet* for type-2 case is totally non-distributive and non-absorptive, and is only partially distributive and absorptive for the type-1 case.

If the membership grades are crisp sets, i.e., if the secondary memberships can be only ones or zeros, then the first *idempotent* law, which fails in the general type-2 case, is satisfied. To see this, consider $\tilde{\mu}_{\tilde{\bar{\text{P}}}} = \sum_{i=1}^{n} 1/u_i$, which is an interval set; then

$$\begin{aligned}
\tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcup \tilde{\mu}_{\tilde{\bar{\text{P}}}} &= (\sum_{i=1}^{n} 1/u_i) \sqcup (\sum_{j=1}^{n} 1/u_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} 1/(u_i \vee u_j) \\
&= \sum_{i=1}^{n} 1/(u_i \vee u_i) + \sum_{\substack{i=1 \\ i \neq j}}^{n}\sum_{j=1}^{n} 1/(u_i \vee u_j) \\
&= \sum_{i=1}^{n} 1/u_i + \sum_{\substack{i=1 \\ i \neq j}}^{n}\sum_{j=1}^{n} 1/(u_i \vee u_j)
\end{aligned} \tag{3.44}$$

Observe that in the second term on the RHS of line 4 above, for each pair $\{i, j\}$ with $i, j = 1, 2, \cdots, n$, the result of $u_i \vee u_j$ is equal to either u_i or u_j, where $i, j = 1, 2, \ldots, n$ and $i \neq j$. Let $u' = u_i \vee u_j$. Now, since the first term contains all the u_i's, with $i = 1, 2, \cdots, n$, u' is contained in it too. Also, since both the secondary memberships of u', in the first term as well as in the second term, are 1, the effective secondary membership of u' in $\tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcup \tilde{\mu}_{\tilde{\bar{\text{P}}}}$ is again 1. So, essentially, $\tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcup \tilde{\mu}_{\tilde{\bar{\text{P}}}}$ contains all the u_i's ($i = 1, 2, \cdots, n$), each with a secondary membership equal to 1, i.e.,

$$\tilde{\mu}_{\tilde{\bar{\text{P}}}} \sqcup \tilde{\mu}_{\tilde{\bar{\text{P}}}} = \sum_{i=1}^{n} 1/u_i = \tilde{\mu}_{\tilde{\bar{\text{P}}}} \tag{3.45}$$

Observe that the set considered above, where membership grades are crisp sets in $[0, 1]$, is still a type-2 set (i.e., it is *not* a type-1 set); however since the first *idempotent* law is not satisfied for the general type-2 case, we say that it is not satisfied in the type-2 case.

As we have already seen, the second part of the *absorption* law and the first part of the *distributive* law, which fail in the general type-2 case are satisfied in case of type-1 sets, i.e., to satisfy these laws, we need the membership grades of type-2 sets to be type-1 singletons, i.e., they must consist of only one primary membership having secondary membership equal to 1 and all other secondary memberships must equal 0.

We have seen that the {max, product} t-conorm / t-norm pair, for type-2 sets, does not satisfy *idempotent, absorption, distributive, De Morgan's* and *complement* laws; hence, if the design of a type-2 fuzzy logic system involves the use of these laws, it will be in error. Observe also that some laws that are satisfied in the type-1 case under {max, product} t-conorm / t-norm pair are not satisfied in the type-2 case (see Table 3.1).

Chapter 4

Relations and Compositions

In this chapter, we examine type-2 fuzzy relations and their compositions. They play an important role in a type-2 fuzzy logic system, which is the subject of Chapter 5.

4.1 Introduction

A relation among crisp sets X_1, X_2, ..., X_n is a subset of the Cartesian product $X_1 \times X_2 \times ... \times X_n$. Just as a type-1 fuzzy relation is a type-1 fuzzy subset of the product space, a type-2 fuzzy relation is a type-2 fuzzy set [12], i.e., every element of the product space has a membership grade which is a fuzzy set of type-1 supported in $[0, 1]$. All the previously discussed type-2 operations like *join*, *meet* and *negation* can be used with type-2 relations.

Example 4.1 Consider the example in [15] of a type-1 fuzzy relation between real numbers, "u is close to v". In [15], the membership function chosen for this example was $\mu_{\widetilde{\text{close}}}(\mid u-v \mid) = \max\{(5- \mid u-v \mid)/5, 0\}$. This membership function is shown in Fig. 4.1 (a).

If one is not sure of the exact nature of the membership function, one could, for example, define the following type-2 relation : for $0 \leq \mid u-v \mid \leq 2.5$, the membership grade $\tilde{\mu}_{\approx\atop\text{close}}(\mid u-v \mid)$ is a type-1 fuzzy set with support $[1-0.3 \mid u-v \mid, 1-0.1 \mid u-v \mid]$; and, for $2.5 \leq \mid u-v \mid \leq 5$, the membership grade $\tilde{\mu}_{\approx\atop\text{close}}(\mid u-v \mid)$ is a type-1 fuzzy set with support $[0.5-0.1 \mid u-v \mid, 1.5-0.3 \mid u-v \mid]$. This membership function is depicted in Fig. 4.1 (b). The secondary membership functions have been chosen so that when $\mid u-v \mid= 0$, there is no additional fuzziness about "close", and similarly, when $\mid u-v \mid= 5$, this is so far from $\mid u-v \mid= 0$ that, again, there is no additional fuzziness about it. When $\mid u-v \mid= 2.5$, the secondary membership reaches unity when the primary membership value equals 0.5 and the support of the secondary membership function is $[0.25, 0.75]$. Figure 4.1 (c) shows an example of a triangular secondary membership function corresponding to $\mid u-v \mid= 2.5$. Other choices are possible for both the support and the shape of the secondary membership function. □

4.2 Relations on the Same Product Space

Consider two universes of discourse, U and V. If $\tilde{\tilde{\text{R}}}(u, v)$ and $\tilde{\tilde{\text{S}}}(u, v)$ are two type-2 fuzzy relations on the same product space, the membership grades of their union and intersection can be represented as

$$\tilde{\mu}_{\tilde{\tilde{\text{R}}}\cup\tilde{\tilde{\text{S}}}}(u, v) = \tilde{\mu}_{\tilde{\tilde{\text{R}}}}(u, v) \sqcup \tilde{\mu}_{\tilde{\tilde{\text{S}}}}(u, v) \quad (4.1)$$

$$\tilde{\mu}_{\tilde{\tilde{\text{R}}}\cap\tilde{\tilde{\text{S}}}}(u, v) = \tilde{\mu}_{\tilde{\tilde{\text{R}}}}(u, v) \sqcap \tilde{\mu}_{\tilde{\tilde{\text{S}}}}(u, v) \quad (4.2)$$

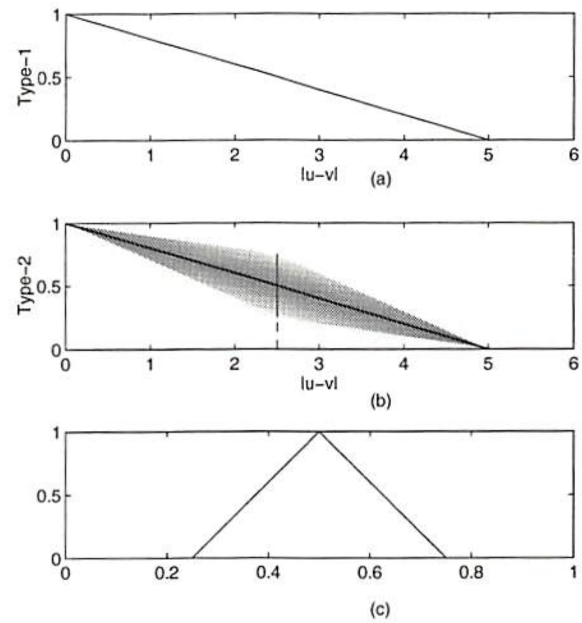


Figure 4.1: Examples of membership functions for (a) type-1, and (b) type-2 fuzzy relations. In (b), the thick dark line shows the primary memberships which have secondary membership equal to 1. The intensity of color in the grey area is approximately proportional to the value of the secondary membership grades. Darker color represents higher secondary membership. The domain of the type-1 fuzzy set corresponding to $|u - v| = 2.5$ is also indicated in the figure. Figure (c) shows the secondary membership function corresponding to $|u - v| = 2.5$.

Example 4.2 Consider the two somewhat contradictory fuzzy relations "u is close to v" and "u is smaller than v". Both are on the same product space $U \times V$. For simplicity, let us assume that $U = \{u_1, u_2\} = \{2, 12\}$ and $V = \{v_1, v_2, v_3\} = \{1, 7, 13\}$. To begin with, we assume that both the relations are type-1 and calculate the membership grades for their union and intersection. Then, we will "extend" this example to the type-2 case. Let the membership grades for the type-1 case be

$$\mu_{\widetilde{\text{close}}}(u,v) = \begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{pmatrix} 0.9 & 0.4 & 0.1 \\ 0.1 & 0.4 & 0.9 \end{pmatrix} \end{array} \tag{4.3}$$

$$\mu_{\widetilde{\text{smaller}}}(u,v) = \begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{pmatrix} 0 & 0.6 & 1 \\ 0 & 0 & 0.3 \end{pmatrix} \end{array} \tag{4.4}$$

The membership grades for the union and intersection of these relations (assuming min t-norm and max t-conorm) can be found as

$$\mu_{\widetilde{\text{close}} \cup \widetilde{\text{smaller}}}(u_i, v_j) = \mu_{\widetilde{\text{close}}}(u_i, v_j) \vee \mu_{\widetilde{\text{smaller}}}(u_i, v_j) \tag{4.5}$$

$$\mu_{\widetilde{\text{close}} \cap \widetilde{\text{smaller}}}(u_i, v_j) = \mu_{\widetilde{\text{close}}}(u_i, v_j) \wedge \mu_{\widetilde{\text{smaller}}}(u_i, v_j) \tag{4.6}$$

where $i = 1, 2$ and $j = 1, 2, 3$. Using (4.5) and (4.6), we have

$$\mu_{\widetilde{\text{close}} \cup \widetilde{\text{smaller}}}(u,v) = \begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{pmatrix} 0.9 & 0.6 & 1 \\ 0.1 & 0.4 & 0.9 \end{pmatrix} \end{array} \tag{4.7}$$

$$\mu_{\widetilde{\text{close}} \cap \widetilde{\text{smaller}}}(u,v) = \begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{pmatrix} 0 & 0.4 & 0.1 \\ 0 & 0 & 0.3 \end{pmatrix} \end{array} \tag{4.8}$$

From (4.7) and (4.8), we see that "u is close to v" *or* """u is smaller than v" is much more sensible than "u is close to v" *and* "u is smaller than v", because membership values in $\mu_{\widetilde{\text{close}} \cup \widetilde{\text{smaller}}}(u,v)$ are fairly large, whereas those in $\mu_{\widetilde{\text{close}} \cap \widetilde{\text{smaller}}}(u,v)$ are mostly very small.

Now, let's consider a type-2 version of the above relations, i.e., let us assume that there is some uncertainty in the membership grades. Let the membership grades for the type-2 case be as follows :

$$\tilde{\mu}_{\underset{\approx}{\text{close}}}(u,v) = \begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{pmatrix} 0.3/0.8 + 1/0.9 + 0.7/1 & 0.7/0.3 + 1/0.4 + 0.1/0.5 & 0.5/0 + 1/0.1 \\ 0.5/0 + 1/0.1 & 0.7/0.3 + 1/0.4 + 0.1/0.5 & 0.3/0.8 + 1/0.9 + 0.7/1 \end{pmatrix} \end{array} \tag{4.9}$$

$$\tilde{\mu}_{\underset{\text{smaller}}{\approx}}(u,v) = \begin{array}{c} \\ u_1 \\ \\ u_2 \end{array} \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{pmatrix} 1/0+0.9/0.1+0.4/0.5 & 0.8/0.3+0.8/0.4+0.9/0.5+1/0.6 & 0.9/0.9+1/1 \\ 1/0+0.1/0.1+0.1/0.2 & 1/0+0.3/0.1 & 1/0.3+0.9/0.4+0.4/0.5 \end{pmatrix} \end{array} \tag{4.10}$$

Observe that we have collected the type-2 membership grades also in a matrix, just like the type-1 case. The only difference in the type-2 case is that each element of the membership matrix is a type-1 fuzzy set rather than a crisp number. Observe, also, that the numbers to the right of the slash represent primary memberships and the ones to the left represent corresponding secondary memberships. We show only those primary memberships which have non-zero secondary memberships. We have purposely chosen the grades so that the primary memberships which correspond to the membership grades in the type-1 case have unity secondary memberships. This can be interpreted as "perturbing" the membership matrices in (4.7) and (4.8) a little bit. We have chosen normal membership grades in this example, just to stress the fact these type-2 grades are related to the type-1 grades in (4.7) and (4.8). Observe, also, that pairs having the same membership in the type-1 case need not necessarily have the same membership in the type-2 case.

The membership grades for the union and intersection of these relations can be found as follows :

$$\tilde{\mu}_{\underset{\text{close}}{\approx} \cup \underset{\text{smaller}}{\approx}}(u_i,v_j) = \tilde{\mu}_{\underset{\text{close}}{\approx}}(u_i,v_j) \sqcup \tilde{\mu}_{\underset{\text{smaller}}{\approx}}(u_i,v_j) \tag{4.11}$$

$$\tilde{\mu}_{\underset{\text{close}}{\approx} \cap \underset{\text{smaller}}{\approx}}(u_i,v_j) = \tilde{\mu}_{\underset{\text{close}}{\approx}}(u_i,v_j) \sqcap \tilde{\mu}_{\underset{\text{smaller}}{\approx}}(u_i,v_j) \tag{4.12}$$

where $i = 1, 2$ and $j = 1, 2, 3$. Using (4.11) and (4.12), we get (using *min* t-norm and *max* t-conorm)

$$\begin{aligned}
\tilde{\mu}_{\underset{\text{close}}{\approx} \cup \underset{\text{smaller}}{\approx}}(u_1,v_1) &= (0.3/0.8+1/0.9+0.7/1) \sqcup (1/0+0.9/0.1+0.4/0.5) \\
&= 0.3/0.8+0.3/0.8+0.3/0.8+1/0.9+0.9/0.9+0.4/0.9 \\
&\quad +0.7/1+0.7/1+0.4/1 \\
&= 0.3/0.8+1/0.9+0.7/1 \\
\tilde{\mu}_{\underset{\text{close}}{\approx} \cup \underset{\text{smaller}}{\approx}}(u_1,v_2) &= (0.7/0.3+1/0.4+0.1/0.5) \\
&\quad \sqcup(0.8/0.3+0.8/0.4+0.9/0.5+1/0.6) \\
&= 0.7/0.3+0.7/0.4+0.7/0.5+0.7/0.6+0.8/0.4+0.8/0.4 \\
&\quad +0.9/0.5+1/0.6+0.1/0.5+0.1/0.5+0.1/0.5+0.1/0.6 \\
&= 0.7/0.3+0.8/0.4+0.9/0.5+1/0.6 \\
\tilde{\mu}_{\underset{\text{close}}{\approx} \cup \underset{\text{smaller}}{\approx}}(u_1,v_3) &= (0.5/0+1/0.1) \sqcup (0.9/0.9+1/1) \\
&= 0.5/0.9+0.5/1+0.9/0.9+1/1 \\
&= 0.9/0.9+1/1 \\
\tilde{\mu}_{\underset{\text{close}}{\approx} \cup \underset{\text{smaller}}{\approx}}(u_2,v_1) &= (0.5/0+1/0.1) \sqcup (1/0+0.1/0.1+0.1/0.2) \\
&= 0.5/0+0.1/0.1+0.1/0.2+1/0.1+0.1/0.1+0.1/0.2 \\
&= 0.5/0+1/0.1+0.1/0.2 \\
\tilde{\mu}_{\underset{\text{close}}{\approx} \cup \underset{\text{smaller}}{\approx}}(u_2,v_2) &= (0.7/0.3+1/0.4+0.1/0.5) \sqcup (1/0+0.3/0.1) \\
&= 0.7/0.3+0.3/0.3+1/0.4+0.3/0.4+0.1/0.5+0.1/0.5 \\
&= 0.7/0.3+1/0.4+0.1/0.5 \\
\tilde{\mu}_{\underset{\text{close}}{\approx} \cup \underset{\text{smaller}}{\approx}}(u_2,v_3) &= (0.3/0.8+1/0.9+0.7/1) \sqcup (1/0.3+0.9/0.4+0.4/0.5)
\end{aligned}$$

$$
\begin{aligned}
&= 0.3/0.8 + 0.3/0.8 + 0.3/0.8 + 1/0.9 + 0.9/0.9 + 0.4/0.9 \\
&\quad +0.7/1 + 0.7/1 + 0.4/1 \\
&= 0.3/0.8 + 1/0.9 + 0.7/1
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{\mu}_{\overset{\approx}{\text{close}} \cap \overset{\approx}{\text{smaller}}}(u_1, v_1) &= (0.3/0.8 + 1/0.9 + 0.7/1) \sqcap (1/0 + 0.9/0.1 + 0.4/0.5) \\
&= 0.3/0 + 0.3/0.1 + 0.3/0.5 + 1/0 + 0.9/0.1 + 0.4/0.5 \\
&\quad +0.7/0 + 0.7/0.1 + 0.4/0.5 \\
&= 1/0 + 0.9/0.1 + 0.4/0.5 \\
\tilde{\mu}_{\overset{\approx}{\text{close}} \cap \overset{\approx}{\text{smaller}}}(u_1, v_2) &= (0.7/0.3 + 1/0.4 + 0.1/0.5) \\
&\quad \sqcap(0.8/0.3 + 0.8/0.4 + 0.9/0.5 + 1/0.6) \\
&= 0.7/0.3 + 0.7/0.3 + 0.7/0.3 + 0.7/0.3 + 0.8/0.3 + 0.8/0.4 \\
&\quad 0.9/0.4 + 1/0.4 + 0.1/0.3 + 0.1/0.4 + 0.1/0.5 + 0.1/0.5 \\
&= 0.8/0.3 + 1/0.4 + 0.1/0.5 \\
\tilde{\mu}_{\overset{\approx}{\text{close}} \cap \overset{\approx}{\text{smaller}}}(u_1, v_3) &= (0.5/0 + 1/0.1) \sqcap (0.9/0.9 + 1/1) \\
&= 0.5/0 + 0.5/0 + 0.9/0.1 + 1/0.1 \\
&= 0.5/0 + 1/0.1 \\
\tilde{\mu}_{\overset{\approx}{\text{close}} \cap \overset{\approx}{\text{smaller}}}(u_2, v_1) &= (0.5/0 + 1/0.1) \sqcap (1/0 + 0.1/0.1 + 0.1/0.2) \\
&= 0.5/0 + 0.1/0 + 0.1/0 + 1/0 + 0.1/0.1 + 0.1/0.1 \\
&= 1/0 + 0.1/0.1 \\
\tilde{\mu}_{\overset{\approx}{\text{close}} \cap \overset{\approx}{\text{smaller}}}(u_2, v_2) &= (0.7/0.3 + 1/0.4 + 0.1/0.5) \sqcap (1/0 + 0.3/0.1) \\
&= 0.7/0 + 0.3/0.1 + 1/0 + 0.3/0.1 + 0.1/0 + 0.1/0.1 \\
&= 1/0 + 0.3/0.1 \\
\tilde{\mu}_{close \cap smaller}(u_2, v_3) &= (0.3/0.8 + 1/0.9 + 0.7/1) \sqcap (1/0.3 + 0.9/0.4 + 0.4/0.5) \\
&= 0.3/0.3 + 0.3/0.4 + 0.3/0.5 + 1/0.3 + 0.9/0.4 + 0.4/0.5 \\
&\quad +0.7/0.3 + 0.7/0.4 + 0.4/0.5 \\
&= 1/0.3 + 0.9/0.4 + 0.4/0.5
\end{aligned}
$$

Collecting the above results in matrices, we have

$$
\tilde{\mu}_{\overset{\approx}{\text{close}} \cup \overset{\approx}{\text{smaller}}}(u, v) =
$$

$$
\begin{array}{c} \\ u_1 \\ \\ u_2 \\ \\ \end{array}
\begin{array}{c}
\begin{array}{ccc} \qquad v_1 \qquad & \qquad v_2 \qquad & \qquad v_3 \qquad \end{array} \\
\begin{pmatrix}
\begin{array}{c} 0.3/0.8 + 1/0.9 \\ +0.7/1 \end{array} & \begin{array}{c} 0.7/0.3 + 0.8/0.4 \\ +0.9/0.5 + 1/0.6 \end{array} & 0.9/0.9 + 1/1 \\
\begin{array}{c} 0.5/0 + 1/0.1 \\ +0.1/0.2 \end{array} & \begin{array}{c} 0.7/0.3 + 1/0.4 \\ +0.1/0.5 \end{array} & \begin{array}{c} 0.3/0.8 + 1/0.9 \\ +0.7/1 \end{array}
\end{pmatrix}
\end{array}
\qquad (4.13)
$$

$$
\tilde{\mu}_{\overset{\approx}{\text{close}} \cap \overset{\approx}{\text{smaller}}}(u, v) =
$$

$$\begin{array}{c} \\ u_1 \\ \\ u_2 \end{array} \begin{array}{c} v_1 \qquad\qquad\qquad v_2 \qquad\qquad\qquad v_3 \\ \begin{pmatrix} \begin{array}{c} 1/0+0.9/0.1 \\ +0.4/0.5 \end{array} & \begin{array}{c} 0.8/0.3+1/0.4 \\ +0.1/0.5 \end{array} & 0.5/0+1/0.1 \\ 1/0+0.1/0.1 & 1/0+0.3/0.1 & \begin{array}{c} 1/0.3+0.9/0.4 \\ +0.4/0.5 \end{array} \end{pmatrix} \end{array} \tag{4.14}$$

Comparing the above results with those in the type-1 case, we observe the following. Since we chose type-2 membership grades in such a way that the primary memberships corresponding to the membership grades in the type-1 case have unity secondary memberships, the memberships of union and intersection also exhibit a similar structure, i.e., primary membership grades for the union (intersection) which correspond to the membership grades for the union (intersection) in the type-1 case, have unity secondary memberships. For example, $\mu_{\widetilde{\text{close}} \cap \widetilde{\text{smaller}}}(u_1, v_2) = 0.4$ and it's type-2 counterpart is $\tilde{\mu}_{\overset{\approx}{\text{close}} \cap \overset{\approx}{\text{smaller}}}(u_1, v_2) = 0.8/0.3 + 1/0.4 + 0.1/0.5$.

Observe that the membership grades for the union are, in general, higher than those for the intersection, i.e., the values of primary memberships of the union that have non-zero secondary memberships are, in general, higher than those of the intersection, again indicating that the intersection ("close" and "smaller") of the above two relations is treated with a higher degree of disbelief than their union ("close" or "smaller"). □

4.3 Relations on Different Product Spaces

Consider two different product spaces, $U \times V$ and $V \times W$, that share a common set and let $R(U, V)$ and $S(V, W)$ be two *crisp* relations on these spaces. The composition of these relations is defined [15] as "a subset $T(U, W)$ of $U \times W$ such that $(u, w) \in T$ if and only if $(u, v) \in R$ and $(v, w) \in S$". This can be expressed as a max-min, max-product or in general, a *sup-star* composition as follows (where $\star$ indicates any suitable t-norm operation) :

$$\mu_{R \circ S}(u, w) = \sup_{v \in V} [\mu_R(u, v) \star \mu_S(v, w)] \tag{4.15}$$

The validity of the sup-star composition for crisp sets is shown in [25]. If R and S are two crisp relations on $U \times V$ and $V \times W$ respectively, then the membership for any pair (u, w), $u \in U$ and $w \in W$, is 1 if and only if there exists at least one $v \in V$ such that $\mu_R(u, v) = 1$ and $\mu_S(v, w) = 1$. In [25], it is shown that this condition is equivalent to having the sup-star composition equal to 1.

When we enter the fuzzy domain, set memberships belong to the interval $[0, 1]$ rather than just being 0 or 1. So, now we can think of an element as belonging to a set if it has a non-zero membership in that set. In this respect, the aforementioned condition on the composition of relations can be rephrased as follows :

> If $\tilde{\text{R}}$ and $\tilde{\text{S}}$ are two type-1 fuzzy relations on $U \times V$ and $V \times W$ respectively, then the membership for any pair (u, w), $u \in U$ and $w \in W$, is non-zero if and only if there exists at least one $v \in V$ such that $\mu_{\tilde{\text{R}}}(u, v) \neq 0$ and $\mu_{\tilde{\text{S}}}(v, w) \neq 0$.

It can be easily shown that this condition is equivalent to the sup-star composition,

$$\mu_{\tilde{\text{R}} \circ \tilde{\text{S}}}(u, w) = \sup_{v \in V} [\mu_{\tilde{\text{R}}}(u, v) \star \mu_{\tilde{\text{S}}}(v, w)] \tag{4.16}$$

In the proof of this, given next, we use the following method. Let **A** be the statement "$\mu_{\tilde{\text{R}} \circ \tilde{\text{S}}}(u, w) \neq 0$", and **B** be the statement "there exists at least one $v \in V$ such that $\mu_{\tilde{\text{R}}}(u, v) \neq 0$ and $\mu_{\tilde{\text{S}}}(v, w) \neq 0$". We prove that "**A** iff **B**" by first proving that $\overline{\mathbf{B}} \Rightarrow \overline{\mathbf{A}}$ (which is equivalent to proving $\mathbf{A} \Rightarrow \mathbf{B}$, i.e., necessity of **B**) and then proving that $\overline{\mathbf{A}} \Rightarrow \overline{\mathbf{B}}$ (which is equivalent to proving $\mathbf{B} \Rightarrow \mathbf{A}$, i.e., sufficiency of **B**).

Proof: Necessity If there exists no $v \in V$ such that $\mu_{\tilde{R}}(u,v) \neq 0$ and $\mu_{\tilde{S}}(v,w) \neq 0$, then this means that for every $v \in V$, either $\mu_{\tilde{R}}(u,v)$ or $\mu_{\tilde{S}}(v,w)$ is equal to zero (or both are zero), which in turn implies that $\mu_{\tilde{R}}(u,v) \star \mu_{\tilde{S}}(v,w) = 0$ for every $v \in V$, i.e., the supremum of $\mu_{\tilde{R}}(u,v) \star \mu_{\tilde{S}}(v,w)$ over $v \in V$ is zero.

Sufficiency If the sup-star composition is zero, then it must be true that $\mu_{\tilde{R}}(u,v) \star \mu_{\tilde{S}}(v,w) = 0$ for every $v \in V$, which means that for every $v \in V$, either $\mu_{\tilde{R}}(u,v)$ or $\mu_{\tilde{S}}(v,w)$ (or both) is zero. This means that there is no $v \in V$ such that $\mu_{\tilde{R}}(u,v) \neq 0$ and $\mu_{\tilde{S}}(v,w) \neq 0$. □

This shows that the sup-star composition is valid for type-1 sets. Of course, this argument does not let us decide whether the actual numerical value arrived at by the sup-star composition is correct in some sense. Also, there might be other expressions involving $\mu_{\tilde{R}}(u,v)$ and $\mu_{\tilde{S}}(v,w)$ that might prove to be equivalent to the condition on the composition of relations.

Now, we show that an equation similar to (4.16) holds for the composition of type-2 fuzzy relations (provided all the secondary membership functions are normalized). The same condition on the composition of relations that we used for the type-1 case can be used for the type-2 case, i.e.,

> If $\tilde{\tilde{R}}$ and $\tilde{\tilde{S}}$ are two type-2 fuzzy relations on $U \times V$ and $V \times W$ respectively, then the membership for any pair (u,w), $u \in U$ and $w \in W$, is non-zero if and only if there exists at least one $v \in V$ such that $\tilde{\mu}_{\tilde{\tilde{R}}}(u,v) \neq 0$ and $\tilde{\mu}_{\tilde{\tilde{S}}}(v,w) \neq 0$.

Recall the concept of 0 and 1 memberships in type-2 sets, discussed earlier. We will also make use of *identity* laws (which hold for minimum and product t-norm as shown in Chapter 3) in the proof of the sup-star composition.

Consider two type-2 fuzzy relations, $\tilde{\tilde{R}}$ and $\tilde{\tilde{S}}$, on $U \times V$ and $V \times W$, respectively. We next show that their composition is equivalent to the following "extended" version of the sup-star composition.

$$\tilde{\mu}_{\tilde{\tilde{R}} \circ \tilde{\tilde{S}}}(u,w) = \sqcup_{v \in V}[\tilde{\mu}_{\tilde{\tilde{R}}}(u,v) \sqcap \tilde{\mu}_{\tilde{\tilde{S}}}(v,w)] \tag{4.17}$$

The method of proof is very similar to that in the type-1 case.

Proof: Necessity If there exists no $v \in V$ such that $\tilde{\mu}_{\tilde{\tilde{R}}}(u,v) \neq 0$ and $\tilde{\mu}_{\tilde{\tilde{S}}}(v,w) \neq 0$, then this means that either $\tilde{\mu}_{\tilde{\tilde{R}}}(u,v) = 0$ or $\tilde{\mu}_{\tilde{\tilde{S}}}(v,w) = 0$ (or both are zero) for every $v \in V$. So, from (3.35), we have that $\tilde{\mu}_{\tilde{\tilde{R}}}(u,v) \sqcap \tilde{\mu}_{\tilde{\tilde{S}}}(v,w) = 0$ for every $v \in V$. From (3.32), we see that $0 \sqcup 0 = 0$, which implies that the RHS of (4.17) is equal to 0.

Sufficiency If the extended *sup-star* composition is zero, then, from (4.17),

$$\sqcup_{v \in V}\ [\tilde{\mu}_{\tilde{\tilde{R}}}(u,v) \sqcap \tilde{\mu}_{\tilde{\tilde{S}}}(v,w)] = 0 \tag{4.18}$$

Observe, from (2.14), that for two normal membership grades, $\mu_{\tilde{P}} = \int_u f(u)/u$ and $\mu_{\tilde{Q}} = \int_v g(v)/v$,

$$\mu_{\tilde{P}} \sqcup \mu_{\tilde{Q}} = 0 \quad \Leftrightarrow \quad \int_u \int_v [f(u) \star g(v)]/(u \vee v) = 1/0 \tag{4.19}$$

which implies that $u = v = 0$, i.e., $\mu_{\tilde{P}} = \mu_{\tilde{Q}} = 1/0$; therefore, from (4.18), it must be true that $\tilde{\mu}_{\tilde{\tilde{R}}}(u,v) \sqcap \tilde{\mu}_{\tilde{\tilde{S}}}(v,w) = 0$ for every $v \in V$, which means that for every $v \in V$, either $\tilde{\mu}_{\tilde{\tilde{R}}}(u,v)$ or $\tilde{\mu}_{\tilde{\tilde{S}}}(v,w)$ (or both) is zero. This means that there is no $v \in V$ such that $\tilde{\mu}_{\tilde{\tilde{R}}}(u,v) \neq 0$ and $\tilde{\mu}_{\tilde{\tilde{S}}}(v,w) \neq 0$. □

Since we generally use normalized membership functions, we can use (4.17) for the composition of type-2 relations. Also, since (3.32) and (3.35) hold for minimum as well as product t-norm, (4.17) will hold for both the choices.

Observe that the proofs of the sup-star composition for type-1 and type-2 cases are very similar. In fact, the only difference in the two is that in the type-2 case we use the more general *meet* operation in place of the $\star$ operation that is used in the type-1 case, and, the concept of 0 and 1 in case of type-2 sets is more general than that in case of type-1 sets.

Dubois and Prade [4, 5] give a formula for the composition of type-2 relations under minimum t-norm as an extension of the type-1 sup-min composition (without proof). This formula is the same as (4.17).

Example 4.3 Consider the type-1 relation "u is close to v" on $U \times V$, where $U = \{u_1, u_2\}$ and $V = \{v_1, v_2, v_3\}$ are as given in Example 4.2. We restate them here for convenience : $U = \{2, 12\}$, $V = \{1, 7, 13\}$ and

$$\mu_{\widetilde{\text{close}}}(u,v) = \begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{pmatrix} 0.9 & 0.4 & 0.1 \\ 0.1 & 0.4 & 0.9 \end{pmatrix} \end{array}$$

Now consider another type-1 fuzzy relation "v is bigger than w" on $V \times W$ where $W = \{w_1, w_2\} = \{4, 8\}$, with the following membership function :

$$\mu_{\widetilde{\text{bigger}}}(v,w) = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array} \begin{array}{c} \begin{array}{cc} w_1 & w_2 \end{array} \\ \begin{pmatrix} 0 & 0 \\ 0.6 & 0 \\ 1 & 0.7 \end{pmatrix} \end{array} \tag{4.20}$$

The statement "u is close to v and v is bigger than w" indicates the composition of these two relations, which can be found by using (4.16) and the minimum t-norm, as follows :

$$\begin{aligned} \mu_{\widetilde{\text{close}} \circ \widetilde{\text{bigger}}}(u_i, w_j) &= [\mu_{\widetilde{\text{close}}}(u_i, v_1) \wedge \mu_{\widetilde{\text{bigger}}}(v_1, w_j)] \\ &\quad \vee [\mu_{\widetilde{\text{close}}}(u_i, v_2) \wedge \mu_{\widetilde{\text{bigger}}}(v_2, w_j)] \\ &\quad \vee [\mu_{\widetilde{\text{close}}}(u_i, v_3) \wedge \mu_{\widetilde{\text{bigger}}}(v_3, w_j)] \end{aligned} \tag{4.21}$$

where $i = 1, 2$ and $j = 1, 2, 3$. Using (4.21), we have

$$\begin{aligned} \mu_{\widetilde{\text{close}} \circ \widetilde{\text{bigger}}}(u_1, w_1) &= [\mu_{\widetilde{\text{close}}}(u_1, v_1) \wedge \mu_{\widetilde{\text{bigger}}}(v_1, w_1)] \\ &\quad \vee [\mu_{\widetilde{\text{close}}}(u_1, v_2) \wedge \mu_{\widetilde{\text{bigger}}}(v_2, w_1)] \\ &\quad \vee [\mu_{\widetilde{\text{close}}}(u_1, v_3) \wedge \mu_{\widetilde{\text{bigger}}}(v_3, w_1)] \\ &= [0.9 \wedge 0] \vee [0.4 \wedge 0.6] \vee [0.1 \wedge 1] \\ &= 0 \vee 0.4 \vee 0.1 \\ &= 0.4 \end{aligned}$$

Doing all the calculations in a similar manner, we get

$$\mu_{\widetilde{\text{close}} \circ \widetilde{\text{bigger}}}(u,w) = \begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{array}{cc} w_1 & w_2 \end{array} \\ \begin{pmatrix} 0.4 & 0.1 \\ 0.9 & 0.7 \end{pmatrix} \end{array} \tag{4.22}$$

Now, just as in Example 4.2, let us consider the type-2 relation "v is bigger than w" which is obtained by adding some uncertainty to the type-1 relation in (4.20). Let the membership grades be as follows :

$$\tilde{\mu}_{\underset{\text{bigger}}{\approx}}(v,w) = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array} \begin{array}{c} w_1 \qquad\qquad\qquad w_2 \\ \left(\begin{array}{cc} 1/0+0.6/0.1 & 1/0+0.1/0.1 \\ 0.4/0.5+1/0.6 & 1/0+0.8/0.1 \\ 0.9/0.7 & 0.2/0.2 \\ 0.7/0.9+1/1 & 0.5/0.6+1/0.7 \\ & 0.7/0.8 \end{array} \right) \end{array} \tag{4.23}$$

The composition of the type-2 relations "u is close to v and v is bigger than w" can be found by using (4.17) as follows :

$$\begin{aligned} \tilde{\mu}_{\underset{\text{close}\circ\text{bigger}}{\approx\ \approx}}(u_i,w_j) &= [\tilde{\mu}_{\underset{\text{close}}{\approx}}(u_i,v_1) \sqcap \tilde{\mu}_{\underset{\text{bigger}}{\approx}}(v_1,w_j)] \\ &\quad \sqcup[\tilde{\mu}_{\underset{\text{close}}{\approx}}(u_i,v_2) \sqcap \tilde{\mu}_{\underset{\text{bigger}}{\approx}}(v_2,w_j)] \\ &\quad \sqcup[\tilde{\mu}_{\underset{\text{close}}{\approx}}(u_i,v_3) \sqcap \tilde{\mu}_{\underset{\text{bigger}}{\approx}}(v_3,w_j)] \end{aligned} \tag{4.24}$$

where $i = 1,2$ and $j = 1,2$. Using (4.24), (4.23) and (4.9), we have (using min t-norm and max t-conorm)

$$\begin{aligned} \tilde{\mu}_{\underset{\text{close}\circ\text{bigger}}{\approx\ \approx}}(u_1,w_1) &= [(0.3/0.8+1/0.9+0.7/1)\sqcap(1/0+0.6/0.1)] \\ &\quad \sqcup[(0.7/0.3+1/0.4+0.1/0.5)\sqcap(0.4/0.5+1/0.6+0.9/0.7)] \\ &\quad \sqcup[(0.5/0+1/0.1)\sqcap(0.7/0.9+1/1)] \\ &= [0.3/0+0.3/0.1+1/0+0.6/0.1+0.7/0+0.6/0.1] \\ &\quad \sqcup[0.4/0.3+0.7/0.3+0.7/0.3+0.4/0.4+1/0.4+0.9/0.4] \\ &\quad \sqcup[0.5/0+0.5/0+0.7/0.1+1/0.1] \\ &= (1/0+0.6/0.1)\sqcup(0.7/0.3+1/0.4)\sqcup(0.5/0+1/0.1) \\ &= [0.7/0.3+1/0.4+0.6/0.3+0.6/0.4]\sqcup(0.5/0+1/0.1) \\ &= (0.7/0.3+1/0.4)\sqcup(0.5/0+1/0.1) \\ &= 0.5/0.3+0.7/0.3+0.5/0.4+1/0.4 \\ &= 0.7/0.3+1/0.4 \\ \tilde{\mu}_{\underset{\text{close}\circ\text{bigger}}{\approx\ \approx}}(u_1,w_2) &= [(0.3/0.8+1/0.9+0.7/1)\sqcap(1/0+0.1/0.1)] \\ &\quad \sqcup[(0.7/0.3+1/0.4+0.1/0.5) \\ &\quad \sqcap(1/0+0.8/0.1+0.2/0.2)] \\ &\quad \sqcup[(0.5/0+1/0.1)\sqcap(0.5/0.6+1/0.7+0.7/0.8)] \\ &= [0.3/0+0.1/0.1+1/0+0.1/0.1+0.7/0+0.1/0.1] \\ &\quad \sqcup[0.7/0+0.7/0.1+0.2/0.2+1/0+0.8/0.1+0.2/0.2 \\ &\quad +0.1/0+0.1/0.1+0.1/0.2] \\ &\quad \sqcup[0.5/0+0.5/0+0.5/0+0.5/0.1+1/0.1+0.7/0.1] \\ &= (1/0+0.1/0.1)\sqcup(1/0+0.8/0.1+0.2/0.2) \\ &\quad \sqcup(0.5/0+1/0.1) \end{aligned}$$

$$
\begin{aligned}
&= [1/0 + 0.8/0.1 + 0.2/0.2 + 0.1/0.1 + 0.1/0.1 + 0.1/0.2] \\
&\quad \sqcup(0.5/0 + 1/0.1) \\
&= (1/0 + 0.8/0.1 + 0.2/0.2) \sqcup (0.5/0 + 1/0.1) \\
&= 0.5/0 + 1/0.1 + 0.5/0.1 + 0.8/0.1 + 0.2/0.2 + 0.2/0.2 \\
&= 0.5/0 + 1/0.1 + 0.2/0.2 \\
\tilde{\mu}_{\underset{\approx}{\text{close}} \circ \underset{\approx}{\text{bigger}}}(u_2, w_1) &= [(0.5/0 + 1/0.1) \sqcap (1/0 + 0.6/0.1)] \\
&\quad \sqcup[(0.7/0.3 + 1/0.4 + 0.1/0.5) \\
&\quad\quad \sqcap(0.4/0.5 + 1/0.6 + 0.9/0.7)] \\
&\quad \sqcup[(0.3/0.8 + 1/0.9 + 0.7/1) \sqcap (0.7/0.9 + 1/1)] \\
&= [0.5/0 + 0.5/0 + 1/0 + 0.6/0.1] \\
&\quad \sqcup[0.4/0.3 + 0.7/0.3 + 0.7/0.3 + 0.4/0.4 + 1/0.4 + 0.9/0.4 \\
&\quad +0.1/0.5 + 0.1/0.5 + 0.1/0.5] \\
&\quad \sqcup[0.3/0.8 + 0.3/0.8 + 0.7/0.9 + 1/0.9 + 0.7/0.9 + 0.7/1] \\
&= (1/0 + 0.6/0.1) \sqcup (0.7/0.3 + 1/0.4 + 0.1/0.5) \\
&\quad \sqcap(0.3/0.8 + 1/0.9 + 0.7/1) \\
&= [0.7/0.3 + 1/0.4 + 0.1/0.5 + 0.6/0.3 + 0.6/0.4 + 0.1/0.5] \\
&\quad \sqcap(0.3/0.8 + 1/0.9 + 0.7/1) \\
&= (0.7/0.3 + 1/0.4 + 0.1/0.5) \sqcap (0.3/0.8 + 1/0.9 + 0.7/1) \\
&= 0.3/0.8 + 0.7/0.9 + 0.7/1 + 0.3/0.8 + 1/0.9 + 0.7/1 \\
&\quad +0.1/0.8 + 0.1/0.9 + 0.1/1 \\
&= 0.3/0.8 + 1/0.9 + 0.7/1 \\
\tilde{\mu}_{\underset{\approx}{\text{close}} \circ \underset{\approx}{\text{bigger}}}(u_2, w_2) &= [(0.5/0 + 1/0.1) \sqcap (1/0 + 0.1/0.1)] \\
&\quad \sqcup[(0.7/0.3 + 1/0.4 + 0.1/0.5) \sqcap (1/0 + 0.8/0.1 + 0.2/0.2)] \\
&\quad \sqcup[(0.3/0.8 + 1/0.9 + 0.7/1) \sqcap (0.5/0.6 + 1/0.7 + 0.7/0.8)] \\
&= [0.5/0 + 0.1/0 + 1/0 + 0.1/0.1] \\
&\quad \sqcup[0.7/0 + 0.7/0.1 + 0.2/0.2 + 1/0 + 0.8/0.1 + 0.2/0.2 \\
&\quad +0.1/0 + 0.1/0.1 + 0.1/0.2] \\
&\quad \sqcup[0.3/0.6 + 0.3/0.7 + 0.3/0.8 + 0.5/0.6 + 1/0.7 + 0.7/0.8 \\
&\quad +0.5/0.6 + 1/0.7 + 0.7/0.8] \\
&= (1/0 + 0.1/0.1) \sqcup (1/0 + 0.8/0.1 + 0.2/0.2) \\
&\quad \sqcup(0.5/0.6 + 1/0.7 + 0.7/0.8) \\
&= [1/0 + 0.8/0.1 + 0.2/0.2 + 0.1/0.1 + 0.1/0.1 + 0.1/0.2] \\
&\quad \sqcup(0.5/0.6 + 1/0.7 + 0.7/0.8) \\
&= (1/0 + 0.8/0.1 + 0.2/0.2) \sqcup (0.5/0.6 + 1/0.7 + 0.7/0.8) \\
&= 0.5/0.6 + 1/0.7 + 0.7/0.8 + 0.5/0.6 + 0.8/0.7 + 0.7/0.8 \\
&\quad +0.2/0.6 + 0.2/0.7 + 0.2/0.8 \\
&= 0.5/0.6 + 1/0.7 + 0.7/0.8
\end{aligned}
$$

Collecting the results in a matrix, we have

$$\tilde{\mu}_{\underset{\approx}{\text{close}} \circ \underset{\approx}{\text{bigger}}}(u,w) = \begin{matrix} \\ u_1 \\ u_2 \end{matrix} \begin{matrix} w_1 & w_2 \\ \left(\begin{matrix} 0.7/0.3 + 1/0.4 \\ 0.3/0.8 + 1/0.9 \\ +0.7/1 \end{matrix} \right. & \left. \begin{matrix} 0.5/0 + 1/0.1 \\ +0.2/0.2 \\ 0.5/0.6 + 1/0.7 \\ +0.7/0.8 \end{matrix} \right) \end{matrix} \tag{4.25}$$

Comparing (4.25) and (4.22), we observe that the results are indeed quite similar to the results of the type-1 sup-star composition, i.e., in the type-2 results, primary memberships corresponding to the memberships in the type-1 results, have unity secondary memberships. □

4.3.1 Composition of a Set with a Relation

Consider the case where one of the relations involved in the composition is just a fuzzy set. The composition of the type-1 set $\tilde{R} \in U$ and type-1 fuzzy relation $\tilde{S}(U,V)$ is given as [15]

$$\mu_{\tilde{R} \circ \tilde{S}}(v) = \sup_{u \in U} [\mu_{\tilde{R}}(u) \star \mu_{\tilde{S}}(u,v)], \tag{4.26}$$

which is a type-1 fuzzy set on V. Similarly, in the type-2 case, the composition of a type-2 fuzzy set in $\tilde{\tilde{R}} \in U$ and a type-2 relation $\tilde{\tilde{S}}(U,V)$ is given by

$$\tilde{\mu}_{\tilde{\tilde{R}} \circ \tilde{\tilde{S}}}(v) = \sqcup_{u \in U} [\tilde{\mu}_{\tilde{\tilde{R}}}(u) \sqcap \tilde{\mu}_{\tilde{\tilde{S}}}(u,v)] \tag{4.27}$$

Equation (4.27) is used in Chapter 5 for the inference mechanism of a rule of a type-2 fuzzy logic system.

Example 4.4 Consider again the type-1 relation "u is close to v" on $U \times V$ in Example 4.2, where $U = \{2, 12\}$ and $V = \{1, 7, 13\}$. We restate it here

$$\mu_{\widetilde{\text{close}}}(u,v) = \begin{matrix} \\ u_1 \\ u_2 \end{matrix} \begin{matrix} \ v_1 \quad v_2 \quad v_3 \\ \begin{pmatrix} 0.9 & 0.4 & 0.1 \\ 0.1 & 0.4 & 0.9 \end{pmatrix} \end{matrix}$$

Also consider the type-1 fuzzy set $\widetilde{\text{small}}$ on U, defined as $\widetilde{\text{small}} = 0.9/2 + 0.1/12$. $\widetilde{\text{small}}$ can be written as a vector

$$\mu_{\widetilde{\text{small}}}(u) = \begin{matrix} \ u_1 \quad u_2 \\ \begin{pmatrix} 0.9 & 0.1 \end{pmatrix} \end{matrix} \tag{4.28}$$

The composition of the two statements "u is small and u is close to v" can, now, be obtained by using (4.26) as follows :

$$\mu_{\widetilde{\text{small}} \circ \widetilde{\text{close}}}(v_j) = [\mu_{\widetilde{\text{small}}}(u_1) \wedge \mu_{\widetilde{\text{close}}}(u_1, v_j)] \vee [\mu_{\widetilde{\text{small}}}(u_2) \wedge \mu_{\widetilde{\text{close}}}(u_2, v_j)] \tag{4.29}$$

where $j = 1, 2, 3$. Using (4.29), we have

$$\begin{aligned} \mu_{\widetilde{\text{small}} \circ \widetilde{\text{close}}}(v_1) &= [0.9 \wedge 0.9] \vee [0.1 \wedge 0.1] \\ &= 0.9 \vee 0.1 \\ &= 0.9 \end{aligned}$$

Doing similar calculations for v_2 and v_3, we get

$$\mu_{\widetilde{\text{small}} \circ \widetilde{\text{close}}}(v) = \begin{array}{c} \begin{matrix} v_1 & v_2 & v_3 \end{matrix} \\ \begin{pmatrix} 0.9 & 0.4 & 0.1 \end{pmatrix} \end{array} \tag{4.30}$$

Now, consider the type-2 fuzzy relation "u is close to v" on $U \times V$, given in Example 4.2. We restate its membership function here for convenience

$$\tilde{\mu}_{\underset{\approx}{\text{close}}}(u,v) =$$

$$\begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{matrix} v_1 & v_2 & v_3 \end{matrix} \\ \begin{pmatrix} 0.3/0.8 + 1/0.9 + 0.7/1 & 0.7/0.3 + 1/0.4 + 0.1/0.5 & 0.5/0 + 1/0.1 \\ 0.5/0 + 1/0.1 & 0.7/0.3 + 1/0.4 + 0.1/0.5 & 0.3/0.8 + 1/0.9 + 0.7/1 \end{pmatrix} \end{array} \tag{4.31}$$

Also consider type-2 fuzzy set $\overset{\approx}{\text{small}}$ on U, whose membership function is obtained by adding some uncertainty to the membership function of $\widetilde{\text{small}}$ as follows :

$$\tilde{\mu}_{\underset{\approx}{\text{small}}}(u_1) = 0.5/0.7 + 1/0.9 \tag{4.32}$$

$$\tilde{\mu}_{\underset{\approx}{\text{small}}}(u_2) = 1/0.1 + 0.3/0.4 \tag{4.33}$$

The composition of the type-2 set $\overset{\approx}{\text{small}}$ and type-2 relation $\overset{\approx}{\text{close}}$, which stands for the statement "u is small and u is close to v" can now be obtained using (4.27) as follows :

$$\tilde{\mu}_{\underset{\approx}{\text{small}} \circ \underset{\approx}{\text{close}}}(v_j) = [\tilde{\mu}_{\underset{\approx}{\text{small}}}(u_1) \sqcap \tilde{\mu}_{\underset{\approx}{\text{close}}}(u_1, v_j)] \sqcup [\tilde{\mu}_{\underset{\approx}{\text{small}}}(u_2) \sqcap \tilde{\mu}_{\underset{\approx}{\text{close}}}(u_2, v_j)] \tag{4.34}$$

where $j = 1, 2, 3$. Using (4.34), we obtain

$$\begin{aligned}
\tilde{\mu}_{\underset{\approx}{\text{small}} \circ \underset{\approx}{\text{close}}}(v_1) &= [(0.5/0.7 + 1/0.9) \sqcap (0.3/0.8 + 1/0.9 + 0.7/1)] \\
&\quad \sqcup [(1/0.1 + 0.3/0.4) \sqcap (0.5/0 + 1/0.1)] \\
&= [0.3/0.7 + 0.5/0.7 + 0.5/0.7 + 0.3/0.8 + 1/0.9 + 0.7/0.9] \\
&\quad \sqcup [0.5/0 + 1/0.1 + 0.3/0 + 0.3/0.1] \\
&= [0.5/0.7 + 0.3/0.8 + 1/0.9] \sqcup [0.5/0 + 1/0.1] \\
&= 0.5/0.7 + 0.5/0.7 + 0.3/0.8 + 0.3/0.8 + 0.5/0.9 + 1/0.9 \\
&= 0.5/0.7 + 0.3/0.8 + 1/0.9 && (4.35) \\
\tilde{\mu}_{\underset{\approx}{\text{small}} \circ \underset{\approx}{\text{close}}}(v_2) &= [(0.5/0.7 + 1/0.9) \sqcap (0.7/0.3 + 1/0.4 + 0.1/0.5)] \\
&\quad \sqcup [(1/0.1 + 0.3/0.4) \sqcap (0.7/0.3 + 1/0.4 + 0.1/0.5)] \\
&= [0.5/0.3 + 0.5/0.4 + 0.1/0.5 + 0.7/0.3 + 1/0.4 + 0.1/0.5] \\
&\quad \sqcup [0.7/0.1 + 1/0.1 + 0.1/0.1 + 0.3/0.3 + 0.3/0.4 + 0.1/0.4] \\
&= [0.7/0.3 + 1/0.4 + 0.1/0.5] \sqcup [1/0.1 + 0.3/0.3 + 0.3/0.4] \\
&= 0.7/0.3 + 0.3/0.3 + 0.3/0.4 + 1/0.4 + 0.3/0.4 + 0.3/0.4 \\
&\quad + 0.1/0.5 + 0.1/0.5 + 0.1/0.5 \\
&= 0.7/0.3 + 1/0.4 + 0.1/0.5 && (4.36) \\
\tilde{\mu}_{\underset{\approx}{\text{small}} \circ \underset{\approx}{\text{close}}}(v_3) &= [(0.5/0.7 + 1/0.9) \sqcap (0.5/0 + 1/0.1)]
\end{aligned}$$

$$
\begin{aligned}
& \sqcup[(1/0.1 + 0.3/0.4) \sqcap (0.3/0.8 + 1/0.9 + 0.7/1)] \\
= \; & [0.5/0 + 0.5/0.1 + 0.5/0 + 1/0.1] \\
& \sqcup[0.3/0.1 + 1/0.1 + 0.7/0.1 + 0.3/0.4 + 0.3/0.4 + 0.3/0.4] \\
= \; & [0.5/0 + 1/0.1] \sqcup [1/0.1 + 0.3/0.4] \\
= \; & 0.5/0.1 + 0.3/0.4 + 1/0.1 + 0.3/0.4 \\
= \; & 1/0.1 + 0.3/0.4
\end{aligned} \tag{4.37}
$$

Collecting the results in a vector, we have

$$
\tilde{\mu}_{\underset{\approx}{\text{small}} \circ \underset{\approx}{\text{close}}}(v) = \begin{pmatrix} \overset{v_1}{0.5/0.7 + 0.3/0.8} & \overset{v_2}{0.7/0.3 + 1/0.4} & \overset{v_3}{1/0.1 + 0.3/0.4} \\ +1/0.9 & +0.1/0.5 & \end{pmatrix} \tag{4.38}
$$

Comparing (4.30) and (4.38), we observe that the type-1 and type-2 results are quite similar. In the type-2 results, primary memberships corresponding to the memberships in the type-1 results, have unity secondary memberships. □

4.4 Cartesian Product of Membership Functions

If U and V are domains of type-1 fuzzy sets $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$, respectively, characterized by certain membership functions, then the Cartesian product of these membership functions will be supported over $U \times V$, and each point $\{(u, v); u \in U, v \in V\}$ in this plane will have a membership grade (a crisp number in $[0, 1]$) which could be the result of a t-norm operation between the membership grade of u in $\tilde{\mathrm{F}}$ and the membership grade of v in $\tilde{\mathrm{G}}$.

Now consider the case of type-2 sets $\tilde{\tilde{\mathrm{F}}}$ and $\tilde{\tilde{\mathrm{G}}}$. The membership grades of every element in $\tilde{\tilde{\mathrm{F}}}$ and $\tilde{\tilde{\mathrm{G}}}$ are now type-1 fuzzy sets. So, now the t-norm operation in the type-1 case has to be replaced by the *meet* operation; therefore, when we find the Cartesian product, it is still supported over $U \times V$ as in the type-1 case; but now each point $(u, v) \in U \times V$ will have a membership grade which is again a type-1 fuzzy set (result of a *meet* operation between membership grades of u and v).

Example 4.5 Consider two universes of discourse, U and V, where $U = \{u_1, u_2, u_3\}$ and $V = \{v_1, v_2\}$. Let $\tilde{\tilde{\mathrm{F}}}$ be a type-2 fuzzy set on U with membership function

$$
\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{\mathrm{F}}}}(u_1) &= 0.9/0.2 + 0.9/0.8 + 0.4/1 \\
\tilde{\mu}_{\tilde{\tilde{\mathrm{F}}}}(u_2) &= 0.1/0.4 + 1/0.7 + 1/1 \\
\tilde{\mu}_{\tilde{\tilde{\mathrm{F}}}}(u_3) &= 0.6/0 + 0.8/0.2
\end{aligned} \tag{4.39}
$$

and let $\tilde{\tilde{\mathrm{G}}}$ be a type-2 fuzzy set on V with membership function

$$
\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{\mathrm{G}}}}(v_1) &= 0.4/0.5 + 0.3/0.6 \\
\tilde{\mu}_{\tilde{\tilde{\mathrm{G}}}}(v_2) &= 0.7/0.6 + 0.6/0.8 + 0.1/0.9
\end{aligned} \tag{4.40}
$$

The membership function for the Cartesian product of these two membership functions can be obtained as

$$
\tilde{\mu}_{\tilde{\tilde{\mathrm{F}}} \times \tilde{\tilde{\mathrm{G}}}}(u_i, v_j) = \tilde{\mu}_{\tilde{\tilde{\mathrm{F}}}}(u_i) \sqcap \tilde{\mu}_{\tilde{\tilde{\mathrm{G}}}}(v_j) \tag{4.41}
$$

where $i = 1, 2, 3$ and $j = 1, 2$. Using (4.41), we get (using min t-norm and max t-conorm)

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{F}} \times \tilde{\tilde{G}}}(u_1, v_1) &= (0.9/0.2 + 0.9/0.8 + 0.4/1) \sqcap (0.4/0.5 + 0.3/0.6) \\
&= 0.4/0.2 + 0.3/0.2 + 0.4/0.5 + 0.3/0.6 + 0.4/0.5 + 0.3/0.6 \\
&= 0.4/0.2 + 0.4/0.5 + 0.3/0.6 \\
\tilde{\mu}_{\tilde{\tilde{F}} \times \tilde{\tilde{G}}}(u_1, v_2) &= (0.9/0.2 + 0.9/0.8 + 0.4/1) \sqcap (0.7/0.6 + 0.6/0.8 + 0.1/0.9) \\
&= 0.7/0.2 + 0.6/0.2 + 0.1/0.2 + 0.7/0.6 + 0.6/0.8 \\
&\quad + 0.1/0.8 + 0.4/0.6 + 0.4/0.8 + 0.1/0.9 \\
&= 0.7/0.2 + 0.7/0.6 + 0.6/0.8 + 0.1/0.9 \\
\tilde{\mu}_{\tilde{\tilde{F}} \times \tilde{\tilde{G}}}(u_2, v_1) &= (0.1/0.4 + 1/0.7 + 1/1) \sqcap (0.4/0.5 + 0.3/0.6) \\
&= 0.1/0.4 + 0.1/0.4 + 0.4/0.5 + 0.3/0.6 + 0.4/0.5 + 0.3/0.6 \\
&= 0.1/0.4 + 0.4/0.5 + 0.3/0.6 \\
\tilde{\mu}_{\tilde{\tilde{F}} \times \tilde{\tilde{G}}}(u_2, v_2) &= (0.1/0.4 + 1/0.7 + 1/1) \sqcap (0.7/0.6 + 0.6/0.8 + 0.1/0.9) \\
&= 0.1/0.4 + 0.1/0.4 + 0.1/0.4 + 0.7/0.6 + 0.6/0.7 \\
&\quad + 0.1/0.7 + 0.7/0.6 + 0.6/0.8 + 0.1/0.9 \\
&= 0.1/0.4 + 0.7/0.6 + 0.6/0.7 + 0.6/0.8 + 0.1/0.9 \\
\tilde{\mu}_{\tilde{\tilde{F}} \times \tilde{\tilde{G}}}(u_3, v_1) &= (0.6/0 + 0.8/0.2) \sqcap (0.4/0.5 + 0.3/0.6) \\
&= 0.4/0 + 0.3/0 + 0.4/0.2 + 0.3/0.2 \\
&= 0.4/0 + 0.4/0.2 \\
\tilde{\mu}_{\tilde{\tilde{F}} \times \tilde{\tilde{G}}}(u_3, v_2) &= (0.6/0 + 0.8/0.2) \sqcap (0.7/0.6 + 0.6/0.8 + 0.1/0.9) \\
&= 0.6/0 + 0.6/0 + 0.1/0 + 0.7/0.2 + 0.6/0.2 + 0.1/0.2 \\
&= 0.6/0 + 0.7/0.2
\end{aligned}$$

Summarizing the calculations in a matrix form, we have

$$\tilde{\mu}_{\tilde{\tilde{F}} \times \tilde{\tilde{G}}}(u, v) = \begin{array}{c} \\ u_1 \\ \\ u_2 \\ \\ u_3 \end{array} \begin{array}{c} v_1 \qquad\qquad\qquad\qquad v_2 \\ \begin{pmatrix} \begin{array}{c} 0.4/0.2 + 0.4/0.5 \\ +0.3/0.6 \end{array} & \begin{array}{c} 0.7/0.2 + 0.7/0.6 + 0.6/0.8 \\ +0.1/0.9 \end{array} \\ \begin{array}{c} 0.1/0.4 + 0.4/0.5 \\ +0.3/0.6 \end{array} & \begin{array}{c} 0.1/0.4 + 0.7/0.6 + 0.6/0.7 \\ +0.6/0.8 + 0.1/0.9 \end{array} \\ 0.4/0 + 0.4/0.2 & 0.6/0 + 0.7/0.2 \end{pmatrix} \end{array} \qquad (4.42)$$

□

Chapter 5

Fuzzy Logic Systems

5.1 Fuzzy Logic

The tenets of fuzzy logic do not change from type-1 to type-2 fuzzy sets, and in general, will not change for any type-n. A higher-type number just indicates a higher *degree of fuzziness.* Since a higher type changes the nature of the membership functions, the operations that depend on the membership functions change; however, the basic principles of fuzzy logic are independent of the nature of membership functions and hence, do not change. Rules of inference like Generalized Modus Ponens or Generalized Modus Tollens continue to apply.

5.1.1 Engineering Membership Functions for Implication

Traditional membership functions for implication, when used for engineering problems, may give counter-intuitive results. For example [15], consider the IF-THEN statement "IF x is $\tilde{\text{A}}$, THEN y is $\tilde{\text{B}}$" having a membership function $\mu_{\tilde{\text{A}}\to\tilde{\text{B}}}(x,y)$. If we use any of the traditional membership functions for $\mu_{\tilde{\text{A}}\to\tilde{\text{B}}}(x,y)$ [e.g., Lukasiewicz implication : $\mu_{\tilde{\text{A}}\to\tilde{\text{B}}}(x,y) = \min\{1, 1-\mu_{\tilde{\text{A}}}(x)+\mu_{\tilde{\text{B}}}(y)\}$], then for any (x, y) pair such that $x \notin \tilde{\text{A}}$ and $y \in \tilde{\text{B}}$, $\mu_{\tilde{\text{A}}\to\tilde{\text{B}}}(x,y) = 1$, which does not make much sense from an engineering perspective, where cause should lead to effect and noncause should not lead to anything. Also, even if the consequent $\tilde{\text{B}}$ in the above rule is associated with a fuzzy set of finite support, as a result of firing the rule, we may get a fuzzy set with infinite support, which again is not an intuitive result.

As we have already seen in Section 2.1, a type-2 membership function can be visualized as a collection of many different type-1 membership functions. Hence, we can see that the difficulties caused by traditional membership functions for implication in the type-1 case, are caused in the type-2 case too. To avoid these difficulties, we may again use *minimum* or *product* inference as in the type-1 case, i.e., we may use the *meet* operation for implication with *minimum* (i.e., Mamdani) or *product* (i.e., Larsen) t-norm. This is the approach that we take.

5.2 Type-2 Fuzzy Logic Systems

This section discusses the structure of a type-2 fuzzy logic system (FLS) in detail. Knowledge about the structure of a type-1 FLS is assumed. Figure 5.1 shows the structure of a type-1 FLS (see [15] for more details) and Figure 5.2 shows that of a type-2 FLS. We assume that both antecedent and consequent sets are type-2; however, this need not necessarily be the case in practice. All the results remain valid even if any or all of the fuzzy sets are type-1.

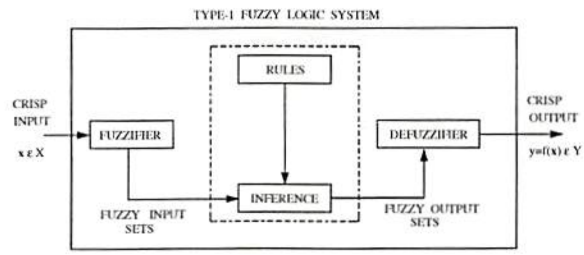


Figure 5.1: The structure of a type-1 FLS [15].

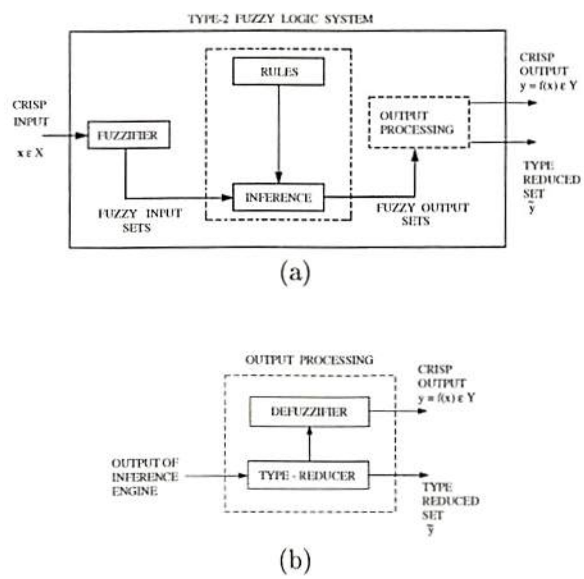


Figure 5.2: The structure of a type-2 FLS. The structure of the *output processing* block is shown in Fig. (b). In order to emphasize the importance of the type-reduced set, we have shown two outputs for the FLS, the type-reduced set and the crisp defuzzified value.

5.2.1 Fuzzification

In this report, we will consider only *singleton* fuzzification, i.e., in the input fuzzy set, a single point will have a non-zero membership grade and this membership grade will be equal to 1. The extension of our work to non-singleton type-2 FLSs is left as future work.

5.2.2 Rules

In the type-1 case, we generally have "IF-THEN" rules, where the lth rule has the form "R^l : IF x_1 is $\tilde{F}_1^l$ and x_2 is $\tilde{F}_2^l$ and $\cdots$ and x_p is $\tilde{F}_p^l$, THEN y is $\tilde{G}^l$", where x_is are inputs; $\tilde{F}_i^l$s are antecedent sets ($i = 1, \dots, p$); y is the output; and $\tilde{G}^l$s are consequent sets.

The distinction between type-1 and type-2 is associated with the nature of the membership function, which is not important while forming rules; hence, the structure of the rules remains exactly the same in the type-2 case, the only difference being that now some or all of the sets involved are of type-2; so, the lth rule in a type-2 FLS has the form "R^l : IF x_1 is $\tilde{\tilde{F}}_1^l$ and x_2 is $\tilde{\tilde{F}}_2^l$ and $\cdots$ and x_p is $\tilde{\tilde{F}}_p^l$, THEN y is $\tilde{\tilde{G}}^l$". It is not necessary that all the antecedents and the consequent be type-2 fuzzy sets. As long as one antecedent or the consequent set is type-2, we will have a type-2 FLS.

5.2.3 Inference Engine

In general, the rules we use will have multiple antecedents connected by *and*'s. Just as in the type-1 case, we can connect these multiple antecedents by the *meet* operation (corresponding to t-norm in the type-1 case). The input membership grade can be combined with the antecedent membership grades by the type-2 version of the *sup-star* composition, described in Chapter 4. Different rules can be combined using the *join* operation (corresponding to t-conorm in the type-1 case), or during defuzzification.

The output of the inference engine consists of the fired consequent fuzzy sets, each one of which is modified from a consequent fuzzy set by a degree of firing. This degree of firing is obtained, in general, as a result of t-norm (*meet*) and t-conorm (*join*) operations on membership grades of the inputs.

Consider a type-2 FLS having p inputs, $x_1 \in X_1, x_2 \in X_2, \dots, x_p \in X_p$ and one output $y \in Y$. Let us suppose that it has M rules, where the lth rule has the form

$$R^l : \text{IF } x_1 \text{ is } \tilde{\tilde{F}}_1^l \text{ and } x_2 \text{ is } \tilde{\tilde{F}}_2^l \text{ and} \dots \text{and } x_p \text{ is } \tilde{\tilde{F}}_p^l, \text{ THEN } y \text{ is } \tilde{\tilde{G}}^l.$$

This rule represents a type-2 fuzzy relation between the input space, $X_1 \times X_2 \times \cdots \times X_p$, and the output space, Y, of the FLS. Let's denote the membership function of this type-2 relation as $\tilde{\mu}_{\tilde{\tilde{F}}_1^l \times \cdots \times \tilde{\tilde{F}}_p^l \to \tilde{\tilde{G}}^l}(\mathbf{x}, y)$, where $\tilde{\tilde{F}}_1^l \times \cdots \times \tilde{\tilde{F}}_p^l$ denotes the Cartesian product of $\tilde{\tilde{F}}_1^l, \tilde{\tilde{F}}_2^l, \dots, \tilde{\tilde{F}}_p^l$, and $\mathbf{x} = \{x_1, x_2, \dots, x_p\}$.

When an input $\mathbf{x}'$ is applied, the composition of the fuzzy set $\tilde{\tilde{\mathbf{X}}}'$, to which $\mathbf{x}'$ belongs and the rule R^l is found by using the extended sup-star composition [see (4.27)]

$$\tilde{\mu}_{\tilde{\tilde{\mathbf{X}}}' \circ \tilde{\tilde{F}}_1^l \times \cdots \times \tilde{\tilde{F}}_p^l \to \tilde{\tilde{G}}^l}(y) = \sqcup_{\mathbf{x} \in \tilde{\tilde{\mathbf{X}}}'} \left[\tilde{\mu}_{\tilde{\tilde{\mathbf{X}}}'}(\mathbf{x}) \sqcap \tilde{\mu}_{\tilde{\tilde{F}}_1^l \times \cdots \times \tilde{\tilde{F}}_p^l \to \tilde{\tilde{G}}^l}(\mathbf{x}, y) \right] \tag{5.1}$$

Since we use singleton fuzzification, the fuzzy set $\tilde{\tilde{\mathbf{X}}}'$ is such that it has a membership grade 1 corresponding to $\mathbf{x} = \mathbf{x}'$ and has zero membership grades for all other inputs; therefore, (5.1) reduces to

$$\tilde{\mu}_{\tilde{\tilde{\mathbf{X}}}' \circ \tilde{\tilde{F}}_1^l \times \cdots \times \tilde{\tilde{F}}_p^l \to \tilde{\tilde{G}}^l}(y) = \tilde{\mu}_{\tilde{\tilde{F}}_1^l \times \cdots \times \tilde{\tilde{F}}_p^l \to \tilde{\tilde{G}}^l}(\mathbf{x}', y) \tag{5.2}$$

We denote $\tilde{\mathbf{X}}' \circ \tilde{\mathrm{F}}_1^l \times \cdots \times \tilde{\mathrm{F}}_p^l \to \tilde{\mathrm{G}}^l$ as $\tilde{\mathrm{B}}^l$, the output set corresponding to the lth rule. The RHS of (5.2) is computed using the implication membership functions (see Section 5.1.1). Since we generally use product or minimum implication (corresponding to the *meet* operation with product or minimum t-norm in the type-2 case), (5.2) can be rewritten as

$$\tilde{\mu}_{\tilde{\mathrm{B}}^l}(y) = \tilde{\mu}_{\tilde{\mathrm{F}}_1^l \times \cdots \times \tilde{\mathrm{F}}_p^l}(\mathbf{x}') \sqcap \tilde{\mu}_{\tilde{\mathrm{G}}^l}(y) \tag{5.3}$$

where $\tilde{\mathrm{F}}_1^l \times \cdots \times \tilde{\mathrm{F}}_p^l$ denotes the Cartesian product of $\tilde{\mathrm{F}}_1^l, \tilde{\mathrm{F}}_2^l, \ldots, \tilde{\mathrm{F}}_p^l$. Using (4.41), (5.3) can be rewritten as

$$\begin{aligned} \tilde{\mu}_{\tilde{\mathrm{B}}^l}(y) &= \tilde{\mu}_{\tilde{\mathrm{F}}_1^l}(x_1) \sqcap \tilde{\mu}_{\tilde{\mathrm{F}}_2^l}(x_2) \sqcap \cdots \sqcap \tilde{\mu}_{\tilde{\mathrm{F}}_p^l}(x_p) \sqcap \tilde{\mu}_{\tilde{\mathrm{G}}^l}(y) \\ &= \tilde{\mu}_{\tilde{\mathrm{G}}^l}(y) \sqcap \left[\sqcap_{i=1}^p \tilde{\mu}_{\tilde{\mathrm{F}}_i^l}(x_i) \right] \end{aligned} \tag{5.4}$$

where $\sqcap$ denotes the *meet* operation based on the t-norm used (assuming that we are using the same operation for the t-norm and the inference ... product or minimum) and we have used the fact that type-2 membership functions commute for minimum or product t-norms (see Table 3.1).

Example 5.1 Figure 5.3 shows an example of product and minimum inference for an arbitrary single-input single-output type-2 FLS, using Gaussian type-2 sets. The product inference function in Fig. 5.3 (c) was obtained by finding the *meet* under product t-norm of the membership grade of $x = 4$ with the membership grade of every point of the consequent function in Fig. 5.3 (b). Let the membership grade at $x = 4$ in the antecedent function be a Gaussian with mean μ and standard deviation Δ, and let the membership grade of every domain point y of the consequent curve be a Gaussian with mean $m(y)$ and standard deviation $\sigma(y)$. The Gaussians are contained in $[0, 1]$ and may, therefore, be clipped. We ignore the effect of this clipping for simplicity. Using the Gaussian approximation in (2.48), the inferred output set is Gaussian type-2, in which the membership grade of every domain point y is a Gaussian with mean $\mu m(y)$ and standard deviation $\sqrt{\Delta^2 m(y)^2 + \mu^2 \sigma(y)^2}$. Similarly, the minimum inference function in Fig. 5.3 (d) was obtained by finding the *meet* under minimum t-norm of the membership grade of $x = 4$ with the membership grade of every point of the consequent function. We used Theorem 2.1 for this. Observe that, in both cases, the result of the inference is a type-2 set [Figs. 5.3 (c) and (d)]. We can interpret the banded behavior about the output's principal membership function as an indication of combined antecedent and consequent uncertainties. □

5.2.4 Type-Reduction

Observe, from Figs. 5.1 and 5.2, that the defuzzifier block in the type-1 FLS is replaced by two blocks : type-reducer and defuzzifier. We consider type-reduction in this subsection.

In a type-1 FLS, where the output sets are type-1 fuzzy sets, we perform defuzzification in order to get a number which is in some sense a crisp (type-0) representative of the combined output sets. In the type-2 case, the output sets are type-2; so we have to use "extended versions" [using the Extension Principle (see Chapter 2)] of type-1 defuzzification methods. Since type-1 defuzzification gives a crisp number at the output of the FLS, the extended defuzzification operation in the type-2 case gives a type-1 fuzzy set at the output. Since this operation takes us from the type-2 output sets of the FLS to a type-1 set, we call this operation "type-reduction" and call the type-1 set so obtained a "type-reduced set". The *type-reduced* fuzzy set may then be defuzzified to obtain a single crisp number; however, in many applications, the type reduced set may be more important than a single crisp number. Output processing is depicted pictorially in Fig. 5.4. Type-reduction is treated, quite extensively, in Sections 5.3 and 5.4.

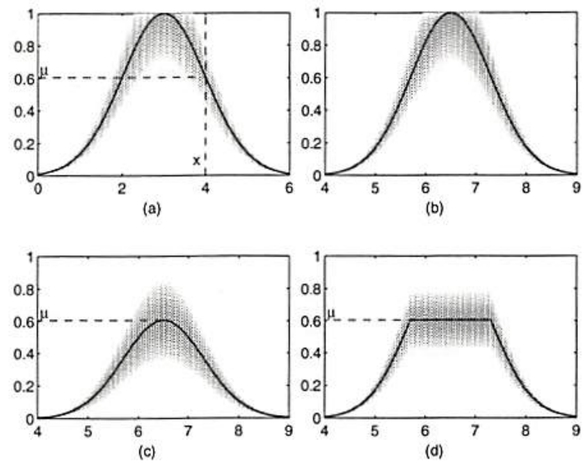


Figure 5.3: Illustrations of product and minimum inference in the type-2 case. (a) A Gaussian type-2 antecedent set for a single input system. The membership of a certain input $x = 4$ in the principal membership function is also shown, equal to μ. (b) The consequent set corresponding to the antecedent set shown in (a). (c) The scaled consequent set, for $x = 4$, using product inference. Observe that the secondary membership functions of the consequent set also change depending upon the standard deviation of the membership grade of x. (d) The clipped consequent set, for $x = 4$, using minimum inference.

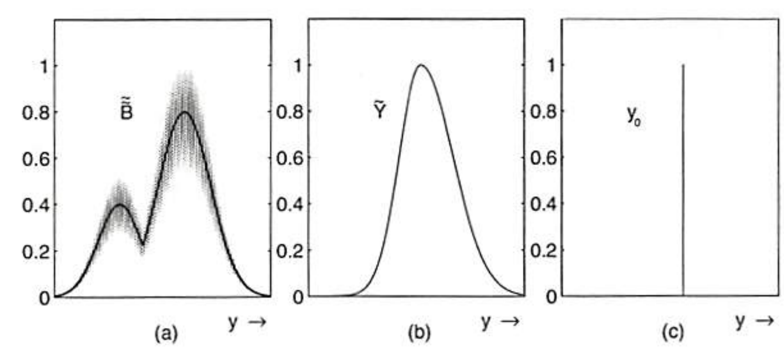


Figure 5.4: A pictorial representation of the *output processing* in a type-2 FLS. For an applied input $\mathbf{x}$, the type-reducer first combines the individual rule output sets in some manner to obtain a combined output set $\tilde{\tilde{B}}$ [Fig. (a)], and then, from $\tilde{\tilde{B}}$, creates a type-1 set, $\tilde{Y}$ [Fig. (b)], which we call the type-reduced set. Finally, the defuzzifier produces a crisp output, y_0 [Fig. (c)], from the type-reduced set.

5.2.5 Defuzzification

To obtain a crisp output from the type-2 FLS, we can defuzzifiy the type-reduced set. The most natural way of doing this seems to be by finding the centroid of the type-reduced set; however, there exist other possibilities, like choosing the unity membership point in the type-reduced set. Defuzzification is treated in Section 5.5.

5.3 Type-Reduction

In a type-1 FLS, the output corresponding to each fired rule is a type-1 set in the output space. The defuzzifier combines the output sets corresponding to all the fired rules in some way to obtain a single output set and then finds a crisp number that is representative of this combined output set, e.g., the centroid defuzzifier finds the union of all the output sets and uses the centroid of the union as the crisp output. *In all the defuzzifiers of interest to us, the crisp number is obtained as the centroid of some combined output set.*

The output set corresponding to each rule of the type-2 FLS is a type-2 set. The type-reducer combines all these output sets in some way (just like a type-1 defuzzifier combines the type-1 rule output sets) and then performs a centroid calculation on this type-2 set, which leads to a type-1 set that we call the "type-reduced" set. See Section 2.5 for discussion about the concept of the centroid of a type-2 set. As explained in Section 2.5.1, we will always use minimum t-norm for calculating the centroid of a type-2 set with a continuous domain, even though we use product t-norm everywhere else.

We now discuss some commonly used type-1 defuzzification methods and their associated type-reduction methods. In each case, we consider a p-input single-output FLS that uses singleton fuzzification and product or minimum inference and has M rules of the form

$$R^l : \text{IF } x_1 \text{ is } \tilde{\tilde{\mathrm{F}}}_1^l \text{ and } x_2 \text{ is } \tilde{\tilde{\mathrm{F}}}_2^l \text{ and } \cdots \text{ and } x_p \text{ is } \tilde{\tilde{\mathrm{F}}}_p^l, \text{ THEN } y \text{ is } \tilde{\tilde{\mathrm{G}}}^l.$$

where $x_i \in X_i$ and $y \in Y$.

Though we have shown all the antecedent and consequent sets to be type-2, all our discussions are valid even when some or all of the sets are type-1. In the latter case, the type-2 system reduces to a type-1 system.

In this section, the FLS is assumed to be type-1 (i.e., all the antecedent and consequent sets are assumed type-1) when we discuss type-1 defuzzification methods; and it is assumed to be type-2 (i.e., some or all of the antecedent and consequent sets are assumed type-2) when we discuss type-reduction methods.

5.3.1 Centroid Type-Reduction

The *centroid* defuzzifier [12] combines the output type-1 sets using a t-conorm (e.g., the maximum t-conorm) and then finds the centroid of this set. If we denote the composite output fuzzy set as $\tilde{\mathrm{B}}$, the centroid defuzzifier is given as

$$y_c(\mathbf{x}) = \frac{\sum_{i=1}^{N} y_i \mu_{\tilde{\mathrm{B}}}(y_i)}{\sum_{i=1}^{N} \mu_{\tilde{\mathrm{B}}}(y_i)} \tag{5.5}$$

where the output set $\tilde{\mathrm{B}}$ is discretized into N points.

The centroid type-reducer combines all the output type-2 sets by finding their union. The membership grade of $y \in Y$ is given as

$$\tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}}(y) = \sqcup_{l=1}^{M} \tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}^l}(y) \tag{5.6}$$

where $\tilde{\mu}_{\tilde{\tilde{B}}^l}(y)$ is as defined in (5.4). The centroid type-reducer then calculates the centroid of $\tilde{\tilde{B}}$. The expression for the centroid type-reduced set is an extended version of (5.5), i.e.,

$$\tilde{Y}_c(\mathbf{x}) = \int_{\theta_1} \cdots \int_{\theta_N} [\mu_{\tilde{D}_1}(\theta_1) \star \cdots \star \mu_{\tilde{D}_N}(\theta_N)] \Big/ \frac{\sum_{i=1}^N y_i\theta_i}{\sum_{i=1}^N \theta_i} \tag{5.7}$$

where $\tilde{D}_i = \tilde{\mu}_{\tilde{\tilde{B}}}(y_i)$ and $\theta_i \in \tilde{\mu}_{\tilde{\tilde{B}}}(y_i)$ $(i = 1, \ldots, N)$. Equation (5.7) can be rewritten as

$$\tilde{Y}_c(\mathbf{x}) = \int_y \sup_{\{\theta_1,\ldots,\theta_N\}} [\mu_{\tilde{D}_1}(\theta_1) \star \cdots \star \mu_{\tilde{D}_N}(\theta_N)] \Big/ y \tag{5.8}$$

where $\{\theta_1, \ldots, \theta_N\}$ are such that $\sum_{i=1}^N y_i\theta_i \Big/ \sum_{j=1}^N \theta_j = y$.

The sequence of computations needed to obtain $\tilde{Y}_c(\mathbf{x})$ is as follows :

1. Compute the combined output set using (5.6). This is possible, because we know $\tilde{\mu}_{\tilde{\tilde{B}}^l}(y)$ $(l = 1, \ldots, M)$ for all $y \in Y$. Theorem 2.3 is used to do this step.
2. Discretize the output space Y into N points, $y_1, \ldots, y_N$.
3. Discretize the domain of each $\tilde{\mu}_{\tilde{\tilde{B}}}(y_i)$ $(i = 1, \ldots, n)$ into a suitable number of points. Figure 5.5 shows a type-2 output set of an arbitrary FLS discretized for type-reduction purposes.
4. Enumerate all the embedded sets (see Section 2.5 for the definition of an embedded type-1 set). For example, if each $\tilde{\mu}_{\tilde{\tilde{B}}}(y)$ is discretized into M points, there will be M^N embedded type-1 sets or, if $\tilde{\mu}_{\tilde{\tilde{B}}}(y_i)$ is discretized into M_i points, there will be $\prod_{j=1}^N M_j$ embedded sets.
5. Compute the centroid type-reduced set using (5.7), i.e., compute the centroid of each enumerated set and assign it a membership grade equal to the t-norm of the secondary memberships corresponding to the enumerated set. We must use the minimum t-norm here as explained in Section 2.5.1.

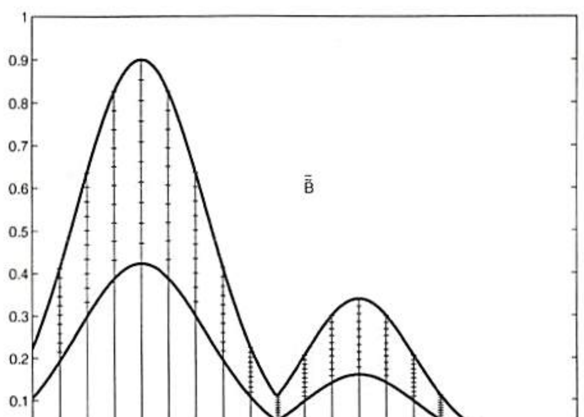


Figure 5.5: An arbitrary type-2 output set $\tilde{\tilde{B}}$ discretized for type-reduction purposes. The independent axis (y) is discretized into N points and the domains of the membership grades corresponding to each y are also discretized. The vertical axis shows primary memberships of y in $\tilde{\tilde{B}}$.

In step 5, the centroid and membership computation has to be repeated $\prod_{j=1}^{N} M_j$ times and so, in general, will involve an enormous amount of computation. These computations lend themselves to parallel processing. Some approximations that reduce the amount of computation tremendously are described in Sections 5.7, 5.9, and in [8]. If the output set is an interval type-2 set, the centroid can also be computed exactly using the computational procedure described in Appendix D. This procedure requires much less computation than the one described above.

5.3.2 Center-of-Sums Type-Reduction

The *center-of-sums* defuzzifier [2] combines the output type-1 sets by adding them and then finds the centroid of this set. The center-of-sums defuzzifier can be expressed as

$$y_a(\mathbf{x}) = \frac{\sum_{l=1}^{M} C_{\tilde{\mathrm{B}}^l} A_{\tilde{\mathrm{B}}^l}}{\sum_{l=1}^{M} A_{\tilde{\mathrm{B}}^l}} \tag{5.9}$$

where, $C_{\tilde{\mathrm{B}}_l}$ denotes the centroid of the lth output set and $A_{\tilde{\mathrm{B}}^l}$ denotes the area of the lth output set. The subscript a on y indicates "additive" combining. If we use product inference, (since the output sets are scaled versions of the consequent sets) $C_{\tilde{\mathrm{B}}^l}$ equals the centroid of the lth consequent set and $A_{\tilde{\mathrm{B}}^l}$ equals the area of the lth consequent set multiplied by the degree of firing of the lth consequent set, $\mathcal{T}_{i=1}^{p} \mu_{\tilde{\mathrm{F}}_i^l}(x_i)$; therefore, (5.9) can be expressed as

$$y_a(\mathbf{x}) = \frac{\sum_{l=1}^{M} c^l a^l \mathcal{T}_{i=1}^{p} \mu_{\tilde{\mathrm{F}}_i^l}(x_i)}{\sum_{l=1}^{M} a^l \mathcal{T}_{i=1}^{p} \mu_{\tilde{\mathrm{F}}_i^l}(x_i)} \tag{5.10}$$

where $\mathcal{T}$ indicates the chosen t-norm, and, c^l is the centroid and a^l is the area of the lth consequent set.

Closely related to the center-of-sums defuzzifier is Kosko's *Standard Additive Model* (SAM) [13], which uses a weighted additive combining of the output sets and product inference. The output of this system for an input $\mathbf{x}$ is

$$y_{SAM}(\mathbf{x}) = \frac{\sum_{l=1}^{M} w_l c^l a^l \mathcal{T}_{i=1}^{p} \mu_{\tilde{\mathrm{F}}_i^l}(x_i)}{\sum_{l=1}^{M} w_l a^l \mathcal{T}_{i=1}^{p} \mu_{\tilde{\mathrm{F}}_i^l}(x_i)} \tag{5.11}$$

Observe that (5.11) is the same as (5.10) except for weights w_l. Observe, also, that since $w_l a^l$ always appears as a product, one really does *not* have 2 degrees of freedom. The weights can be absorbed into the areas or vice-versa. So, without loss of generality, we can set $w_l = 1$ for all l, in which case $y_{SAM}(\mathbf{x}) = y_a(\mathbf{x})$.

The center-of-sums type-reducer combines the type-2 output sets by adding them and then finds the centroid of the resulting type-2 set, which corresponds to the center-of-sums type-reduced set $\tilde{\mathrm{Y}}_a(\mathbf{x})$. In the type-1 case, since the center-of-sums defuzzifier can be expressed in terms of the centroids and areas of each individual output set [see (5.9)], calculations are simplified (as compared to the centroid defuzzifier). In the type-2 case, however, no such simplification occurs. We demonstrate this by means of an example.

Example 5.2 Consider a type-2 set $\tilde{\tilde{\mathrm{A}}} = \tilde{\tilde{\mathrm{A}}}_1 + \tilde{\tilde{\mathrm{A}}}_2$. Let us suppose that the domain of $\tilde{\tilde{\mathrm{A}}}$ is discretized into N points, $x_1, \ldots, x_N$, and that $x_1, \ldots, x_{N_1} \in \tilde{\tilde{\mathrm{A}}}_1$ and $x_{N_1+1} \ldots, x_N \in \tilde{\tilde{\mathrm{A}}}_2$. If we let $\tilde{\mathrm{D}}^i = \tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}}(x_i)$, $\tilde{\mathrm{D}}_1^i = \tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}_1}(x_i)$ and $\tilde{\mathrm{D}}_2^i = \tilde{\mu}_{\tilde{\tilde{\mathrm{A}}}_2}(x_i)$, the centroid of $\tilde{\tilde{\mathrm{A}}}$ is [(2.69)]

$$
\begin{aligned}
\tilde{C}_{\tilde{\tilde{\mathrm{A}}}} &= \int_{\theta_1} \cdots \int_{\theta_N} \mathcal{T}_{i=1}^{N} \mu_{\tilde{\mathrm{D}}^i}(\theta_i) \Big/ \frac{\sum_{i=1}^{N} x_i \theta_i}{\sum_{i=1}^{N} \theta_i} \\
&= \int_{\theta_1} \cdots \int_{\theta_{N_1}} \int_{\theta_{N_1+1}} \cdots \int_{\theta_N} \left[\mathcal{T}_{i=1}^{N_1} \mu_{\tilde{\mathrm{D}}_1^i}(\theta_i)\right] \star \left[\mathcal{T}_{i=N_1}^{N} \mu_{\tilde{\mathrm{D}}_2^i}(\theta_i)\right] \Big/ \\
&\qquad \frac{\left(\frac{\sum_{i=1}^{N_1} x_i\theta_i}{\sum_{i=1}^{N_1} \theta_i}\right)(\sum_{i=1}^{N_1} \theta_i) + \left(\frac{\sum_{i=N_1+1}^{N} x_i\theta_i}{\sum_{i=N_1+1}^{N} \theta_i}\right)(\sum_{i=N_1+1}^{N} \theta_i)}{(\sum_{i=1}^{N_1} \theta_i) + (\sum_{i=N_1+1}^{N} \theta_i)} \\
&= \int_{\theta_1} \cdots \int_{\theta_{N_1}} \int_{\theta_{N_1+1}} \cdots \int_{\theta_N} \left[\mathcal{T}_{i=1}^{N_1} \mu_{\tilde{\mathrm{D}}_1^i}(\theta_i)\right] \star \left[\mathcal{T}_{i=N_1}^{N} \mu_{\tilde{\mathrm{D}}_2^i}(\theta_i)\right] \Big/ \\
&\qquad \frac{c_1 a_1 + c_2 a_2}{a_1 + a_2}
\end{aligned}
\tag{5.12}
$$

where $\{c_1, a_1\}$ and $\{c_2, a_2\}$ are the centroids and areas of those embedded type-1 sets in $\tilde{\tilde{\mathrm{A}}}^1$ and $\tilde{\tilde{\mathrm{A}}}^2$, which are determined by the primary memberships $\{\theta_1, \ldots, \theta_{N_1}\}$ and $\{\theta_{N_1+1}, \ldots, \theta_N\}$, respectively.

Equation (5.12) can not be simplified any further. Even though, (5.12) gives the centroid of $\tilde{\tilde{\mathrm{A}}}$ in terms of centroids and areas of embedded type-1 sets in $\tilde{\tilde{\mathrm{A}}}_1$ and $\tilde{\tilde{\mathrm{A}}}_2$, this does not serve any great purpose from a computational point of view. If we want to use (5.12), we still have to consider all the possible combinations of type-1 embedded sets in $\tilde{\tilde{\mathrm{A}}}_1$ and $\tilde{\tilde{\mathrm{A}}}_2$. For each such combination, we have to calculate the centroid and area for each type-1 set to get one point ($[c_1a_1 + c_2a_2]/[a_1 + a_2]$) in $\tilde{C}_{\tilde{\tilde{\mathrm{A}}}}$; and, then calculate the membership grade of this point by finding the t-norm of all the secondary memberships of $\{\theta_1, \ldots, \theta_N\}$ that make up the embedded curves. This procedure is no simpler than a straightforward centroid calculation as in the centroid type-reducer. □

The most straightforward way of computing the center-of-sums type-reduced set seems to be by finding the centroid of the sum of the output sets, just as the centroid type-reducer finds the centroid of the union of output sets. Equations (5.7) or (5.8) can be used for this purpose with $\tilde{\tilde{\mathrm{B}}}$ indicating the sum of output sets, i.e.,

$$
\tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}}(y) = \sum_{l=1}^{M} \tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}^l}(y); \quad y \in Y \tag{5.13}
$$

See Section 2.4 for more discussion about addition of fuzzy numbers. The SAM type-reducer would also use (5.7) or (5.8) with $\tilde{\tilde{\mathrm{B}}}$ denoting the weighted addition of output sets instead of their union.

The sequence of computations (as well as the intense computational load) needed to obtain $\tilde{\mathrm{Y}}_a(\mathbf{x})$ is exactly the same as we just described for $\tilde{\mathrm{Y}}_c(\mathbf{x})$, except that in step 1, we compute the combined output set using (5.13).

5.3.3 Height Type-Reduction

The height defuzzifier [2] replaces each rule output set by a singleton at the point having maximum membership in that output set, and then calculates the centroid of the type-1 set comprised of these singletons. The output of a height defuzzifier is given as

$$y_h(\mathbf{x}) = \frac{\sum_{l=1}^{M} \bar{y}^l \mu_{\mathrm{B}^l}(\bar{y}^l)}{\sum_{l=1}^{M} \mu_{\mathrm{B}^l}(\bar{y}^l)} \tag{5.14}$$

where : $\bar{y}^l$ is the point having maximum membership in the lth output set (if there is more than one such point, their average can be taken as $\bar{y}^l$); and, its membership grade in the lth output set, $\mu_{\mathrm{B}^l}(\bar{y}^l)$, is given as [15]

$$\mu_{\mathrm{B}^l}(\bar{y}^l) = \mu_{\mathrm{G}^l}(\bar{y}^l) \star \mathcal{T}_{i=1}^{p} \mu_{\mathrm{F}_i^l}(x_i)\,; \tag{5.15}$$

where $\star$ and $\mathcal{T}$ indicate the chosen t-norm (assuming the inference uses the same operation as the t-norm ... product or minimum).

The height type-reducer replaces each output set by a type-2 singleton, i.e., by a fuzzy set whose domain consists of a single point, the membership of which is a type-1 set in $[0, 1]$. The lth output set is replaced by a singleton situated at $\bar{y}^l$, where $\bar{y}^l$ can be chosen to be the point having the highest membership in the principal membership function of output set $\tilde{\mathrm{B}}^l$. If $\tilde{\mathrm{B}}^l$ is such that a principal membership function cannot be defined (for an example of such a type-2 set, see Fig. 1.2), one may choose $\bar{y}^l$ as the point having the highest primary membership with a secondary membership equal to 1, or as a point satisfying some similar criterion (in Fig. 1.2, we can choose $\bar{y}^l = 0.5$, the mid-point of the interval of uncertainty of the mean). The membership grade of $\bar{y}^l$ in the type-2 case is obtained from (5.4) as follows :

$$\tilde{\mu}_{\tilde{\mathrm{B}}^l}(\bar{y}^l) = \tilde{\mu}_{\tilde{\mathrm{G}}^l}(\bar{y}^l) \sqcap \left[\sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{\mathrm{F}}_i^l}(x_i)\right] \tag{5.16}$$

If we let $\tilde{\mathrm{D}}^l = \tilde{\mu}_{\tilde{\mathrm{B}}^l}(\bar{y}^l)$, the expression for the type-reduced set is obtained as an extension of (5.14), as

$$\tilde{\mathrm{Y}}_h(\mathbf{x}) = \int_{\theta_1}\int_{\theta_2}\cdots\int_{\theta_M} [\mu_{\tilde{\mathrm{D}}^1}(\theta_1) \star \cdots \star \mu_{\tilde{\mathrm{D}}^M}(\theta_M)] \Big/ \frac{\sum_{l=1}^{M} \bar{y}^l \theta_l}{\sum_{l=1}^{M} \theta_l} \tag{5.17}$$

where $\theta_l \in \tilde{\mathrm{D}}^l$ for $l = 1, \ldots, M$. If we let $\sum_{l=1}^{M} \bar{y}^l\theta_l \big/ \sum_{l=1}^{M} \theta_l = y$, then (5.17) can also be written as

$$\tilde{\mathrm{Y}}_h(\mathbf{x}) = \int_y \sup_{\{\theta_1,\ldots,\theta_M\}} [\mu_{\tilde{\mathrm{D}}^1}(\theta_1) \star \cdots \star \mu_{\tilde{\mathrm{D}}^M}(\theta_M)] \big/ y \tag{5.18}$$

where $\{\theta_1, \ldots, \theta_M\}$ are such that $\sum_{l=1}^{M} \bar{y}^l\theta_l \big/ \sum_{l=1}^{M} \theta_l = y$.

The sequence of computations needed to obtain $\tilde{\mathrm{Y}}_h(\mathbf{x})$ is as follows :

1. Choose $\bar{y}^l$ for each output set, $l = 1, 2, \ldots, M$.

2. Discretize the domain of each $\tilde{\mu}_{\tilde{\mathrm{B}}^l}(\bar{y}^l)$ into a suitable number of points. The discretization is carried out in a manner similar to that for centroid or center-of-sums type-reduction (see Fig. 5.5), the only difference is that the number of points on the horizontal axis is now M instead of N.

3. Enumerate all the possible combinations $\{\theta_1, \ldots, \theta_M\}$, such that $\theta_l \in \tilde{\mu}_{\tilde{\mathrm{B}}^l}(\bar{y}^l)$. For example, if $\tilde{\mu}_{\tilde{\mathrm{B}}^l}(\bar{y}^l)$ is discretized into N_l points, there will be $\prod_{j=1}^{M} N_j$ combinations.

4. Compute the height type-reduced set using (5.17). Since the domain of the combined output set is discrete, we can now use product or minimum t-norm in (5.17), as explained in Section 2.5.

In step 4, the weighted sum and membership computations in (5.17) have to be repeated $\prod_{j=1}^{M} N_j$ times; but, generally, $\prod_{j=1}^{M} N_j \ll \prod_{j=1}^{N} N_j$ (where N is the number of discrete y-points in case of centroid or center-of-sums type-reduction), so computing the height type-reduced set involves much fewer computations than computing the centroid or center-of-sums type-reduced set. Again, parallel processing is possible, and, some approximations that reduce the amount of computation tremendously are described in Sections 5.7, 5.9, and in [8]. If the FLS uses interval type-2 sets, the type-reduced can also be computed exactly using the computational procedure described in Appendix D. This procedure requires much less computation than the one described above.

5.3.4 Modified Height Type-Reduction

The modified height defuzzifier [15] is very similar to the height defuzzifier, the only difference being that the modified height defuzzifier scales each $\mu_{\tilde{\mathrm{B}}^l}(\bar{y}^l)$ by the inverse of the spread (or some measure of the spread) of the lth consequent set. Its output can be expressed as

$$y_{mh}(\mathbf{x}) = \frac{\sum_{l=1}^{M} \bar{y}^l \mu_{\tilde{\mathrm{B}}^l}(\bar{y}^l)/\delta^{l^2}}{\sum_{l=1}^{M} \mu_{\tilde{\mathrm{B}}^l}(\bar{y}^l)/\delta^{l^2}} \tag{5.19}$$

where δ^l is some measure of the spread of the lth consequent set, and $\bar{y}^l$ and $\mu_{\tilde{\mathrm{B}}^l}(\bar{y}^l)$ have the same meanings as in (5.14).

When all the consequent sets are normal, convex and symmetric, and we use product inference, then $\mu_{\tilde{\mathrm{B}}^l}(\bar{y}^l) = \mathcal{T}_{i=1}^{p} \mu_{\tilde{\mathrm{F}}_i^l}(x_i)$ and $\bar{y}^l = c^l$, the centroid of the lth consequent set. Now, observe by comparing (5.19) with (5.10) and (5.11), that : (1) if $1/\delta^{l^2} = a^l$, then $y_{mh}(\mathbf{x}) = y_a(\mathbf{x})$; and (2) if $1/\delta^{l^2} = w_l a^l$, then $y_{mh}(\mathbf{x}) = y_{SAM}(\mathbf{x})$.

The only difference between the modified height type-reducer and the height type-reducer is that each output set membership, $\tilde{\mu}_{\tilde{\bar{\mathrm{B}}}^l}(\bar{y}^l)$, in the modified height type-reducer is scaled by $1/\delta^{l^2}$, where δ^l is some measure of the spread of the lth consequent set (δ^l can, for example, be taken as the spread of the principal membership function of the lth consequent set). The expression for the type-reduced set is given as

$$\tilde{\mathrm{Y}}_{mh}(\mathbf{x}) = \int_{\theta_1} \int_{\theta_2} \cdots \int_{\theta_M} [\mu_{\tilde{\mathrm{D}}^1}(\theta_1) \star \cdots \star \mu_{\tilde{\mathrm{D}}^M}(\theta_M)] \Big/ \frac{\sum_{l=1}^{M} \bar{y}^l \theta_l/\delta^{l^2}}{\sum_{l=1}^{M} \theta_l/\delta^{l^2}} \tag{5.20}$$

where all symbols have the same meaning as in (5.17).

The sequence of computations needed to obtain $\tilde{\mathrm{Y}}_{mh}(\mathbf{x})$ is as follows :

1. Choose $\bar{y}^l$ and δ^l for each output set, $l = 1, 2, \ldots, M$.

2. Discretize the domain of each $\tilde{\mu}_{\tilde{\bar{\mathrm{B}}}^l}(\bar{y}^l)$ into a suitable number of points.

3. Enumerate all the possible combinations $\{\theta_1, \ldots, \theta_M\}$, such that $\theta_l \in \tilde{\mu}_{\tilde{\bar{\mathrm{B}}}^l}(\bar{y}^l)$. For example, if $\tilde{\mu}_{\tilde{\bar{\mathrm{B}}}^l}(\bar{y}^l)$ is discretized into N_l points, there will be $\prod_{j=1}^{M} N_j$ combinations.

4. Compute the modified height type-reduced set using (5.20). Since the domain of the combined output set is discrete, we can now use product or minimum t-norm in (5.20), as explained in Section 2.5.

Our discussions at the end of Section 5.3.3, about computational complexity, apply here as well.

5.3.5 Center-of-Sets Type-Reduction

For reasons that will be explained in Section 5.3.6, we introduce a new center-of-sets defuzzification method, in which we replace each rule consequent set by a singleton situated at its *centroid* and then find the centroid of the type-1 set comprised of these singletons. The expression for the output is given as

$$y_{cos}(\mathbf{x}) = \frac{\sum_{l=1}^{M} c^l \mathcal{T}_{i=1}^{p} \mu_{\tilde{\mathrm{F}}_i^l}(x_i)}{\sum_{l=1}^{M} \mathcal{T}_{i=1}^{p} \mu_{\tilde{\mathrm{F}}_i^l}(x_i)} \tag{5.21}$$

where $\mathcal{T}$ indicates the chosen t-norm, and, c^l is the centroid of the lth consequent set. Observe that if each consequent set is symmetric, normal and convex, $c^l = \bar{y}^l$ and $\mu_{\tilde{\mathrm{G}}^l}(\bar{y}^l) = 1$ for $l = 1, \ldots, M$; consequently, $y_{cos}(\mathbf{x}) = y_h(\mathbf{x})$. Observe, also, that (5.21) is similar to the expression for the center-of-sums defuzzifier using product inference [Eq. (5.10)], except that in (5.21) we do not consider areas of the individual output sets.

The center-of-sets type-reducer replaces each consequent set by its centroid (which itself is a type-1 set, if the consequent set is type-2), and finds a weighted average of these centroids, where the weight associated with the lth centroid is the degree of firing corresponding to the lth rule, namely $\sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{\mathrm{F}}_i^l}(x_i)$. The expression for the type-reduced set is the following extension of (5.21),

$$\tilde{\mathrm{Y}}_{cos}(\mathbf{x}) = \int_{d_1} \cdots \int_{d_M} \int_{e_1} \cdots \int_{e_M} \mathcal{T}_{l=1}^{M} \mu_{\tilde{\mathrm{C}}_l}(d_l) \star \mathcal{T}_{l=1}^{M} \mu_{\tilde{\mathrm{E}}_l}(e_l) \Bigg/ \frac{\sum_{l=1}^{M} d_l e_l}{\sum_{l=1}^{M} e_l} \tag{5.22}$$

where : $\mathcal{T}$ and $\star$ indicate the chosen t-norm; $d_l \in \tilde{\mathrm{C}}_l = \tilde{C}_{\tilde{\mathrm{G}}^l}$, the centroid of the lth consequent set (we have used d_l as the variable name instead of c_l to avoid confusion later on in Appendix D); and, $e_l \in \tilde{\mathrm{E}}_l = \sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{\mathrm{F}}_i^l}(x_i)$, the degree of firing associated with the lth consequent set, for $l = 1, \ldots, M$.

The sequence of computations needed to obtain $\tilde{\mathrm{Y}}_{cos}(\mathbf{x})$ is as follows :

1. Discretize the output space Y into a suitable number of points, and compute the centroid $\tilde{\mathrm{C}}_l$ of each consequent set on the discretized output space using (2.69). This is possible, because we know $\tilde{\mu}_{\tilde{\mathrm{G}}^l}(y)$ $(l = 1, 2, \ldots, M)$ for all $y \in Y$. *These consequent centroid sets can be computed ahead of time and stored for future use.*

2. Compute the degree of firing, $\tilde{\mathrm{E}}_l = \sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{\mathrm{F}}_i^l}(x_i)$, associated with the lth consequent set, using the results in Chapter 2 (Theorem 2.1 for minimum t-norm, the approximation in Section 2.3 for product t-norm, or results in Appendix D for interval sets).

3. Discretize the domain of each $\tilde{\mathrm{C}}_l$ ($\tilde{\mathrm{C}}_l$ is the centroid of the lth consequent set) into a suitable number of points, say M_l $(l = 1, 2, \ldots, M)$.

4. Discretize the domain of each $\tilde{\mathrm{E}}_l$ into a suitable number of points, say N_l $(l = 1, 2, \ldots, M)$.

5. Enumerate all the possible combinations $\{d_1, \ldots, d_M, e_1, \ldots, e_M\}$ such that $d_l \in \tilde{\mathrm{C}}_l$ and $e_l \in \tilde{\mathrm{E}}_l$. The total number of combinations will be $\prod_{j=1}^{M} M_j N_j$.

6. Compute the center-of-sums type-reduced set using (5.22).

In step 6, the weighted sum and t-norm operations in (5.22) have to be repeated $\prod_{j=1}^{M} M_j N_j$ times. Observe, from Sections 5.3.1 and 5.3.3, that this number is, in general, larger than that required for a height (or modified height) type-reducer, but is less than that required for a centroid (or center-of-sums) type-reducer. Parallel processing is possible, and, some approximations that reduce the amount of computation tremendously are described in Sections 5.7, 5.9, and in [8]. If the FLS uses interval type-2 sets, the center-of-sets type-reduced set can also be computed exactly using the computational procedure described in Appendix D. This procedure requires much less computation than the one described above.

Figure 5.6 (a) shows the centroids of consequent sets, $\tilde{C}_1$, $\tilde{C}_2$ and $\tilde{C}_3$, for an arbitrary type-2 FLS and Fig. 5.6 (b), (c) and (d) show their corresponding degrees of firing, $\tilde{E}_1$, $\tilde{E}_2$ and $\tilde{E}_3$, for some applied input. Both, the centroids and their degrees of firing are discretized for type-reduction purposes. In this example, $M_1 = 15$, $M_2 = 16$, $M_3 = 14$, $N_1 = 8$, $N_2 = 7$ and $N_3 = 11$, so that step 5 will involve $2,069,760$ combinations.

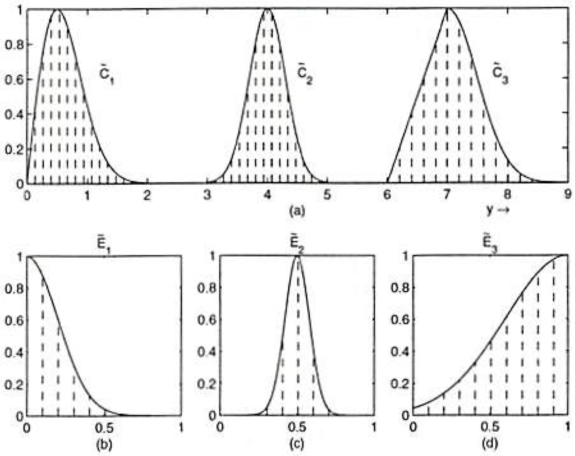


Figure 5.6: (a) The centroids of consequent sets for an arbitrary type-2 FLS. The corresponding degrees of firing, for some applied input, are shown in (b), (c) and (d), respectively [each of (b), (c) and (d) shows the plot of secondary versus primary memberships]. All the centroids and the degrees of firing are discretized for center-of-sets type-reduction.

Next, we describe a situation when computation of (5.22) is simplified considerably. Observe that (5.22) can also be written as

$$\begin{aligned}\tilde{Y}_{cos}(\mathbf{x}) &= \int_{e_1} \cdots \int_{e_M} \mathcal{T}_{l=1}^{M} \mu_{\tilde{E}_l}(e_l) \Big/ \frac{\sum_{l=1}^{M} \tilde{C}_l e_l}{\sum_{l=1}^{M} e_l} \\ &= \int_{e_1} \cdots \int_{e_M} \mathcal{T}_{l=1}^{M} \mu_{\tilde{E}_l}(e_l) \Big/ \tilde{Y}'(e_1, \ldots, e_M) \end{aligned} \quad (5.23)$$

where, for every combination of $\{e_1, \ldots, e_M\}$, the quantity to the left of the slash $[\mathcal{T}_{l=1}^{M} \mu_{\tilde{E}_l}(e_l)]$, is t-normed with the membership grade of every point in $\tilde{Y}'(e_1, \ldots, e_M)$; and, $\tilde{Y}_{cos}(\mathbf{x})$ is given by the union of all these scaled (or clipped if the t-norm used is minimum) $\tilde{Y}'(e_1, \ldots, e_M)$'s. When the consequent centroids ($\tilde{C}_l$'s) are Gaussian or interval sets, Theorems 2.4 or D.1 can be used to compute $\tilde{Y}'(e_1, \ldots, e_M)$. This reduces the number of computations required to compute $\tilde{Y}_{cos}(\mathbf{x})$ considerably. We demonstrate the use of Eq. (5.23) with an example.

Example 5.3 Consider a hypothetical situation, where for a certain input, [see (5.22) and (5.23)] $M = 2$, $\tilde{C}_1 = 1/[1, 2]$, $\tilde{C}_2 = 1/[4, 5]$, $\tilde{E}_1 = 0.5/0.6 + 1/0.8$, and $\tilde{E}_2 = 1/0.4$. Observe that $\tilde{C}_1$ and

$\tilde{C}_2$ are interval type-1 sets, and $\tilde{E}_2$ is crisp. Using (5.23) and product t-norm, the type-reduced set is computed as follows :

$$\begin{aligned}\tilde{Y}_{cos}(\mathbf{x}) &= \sum_{e_1}\sum_{e_2} \mu_{\tilde{E}_1}(e_1) \star \mu_{\tilde{E}_2}(e_2) \Big/ \frac{e_1\tilde{C}_1 + e_2\tilde{C}_2}{e_1 + e_2} \\ &= \left[0.5 \Big/ \frac{0.6\tilde{C}_1 + 0.4\tilde{C}_2}{0.6 + 0.4}\right] \cup \left[1 \Big/ \frac{0.8\tilde{C}_1 + 0.4\tilde{C}_2}{0.8 + 0.4}\right] \end{aligned} \quad (5.24)$$

where the "+" signs indicate algebraic sum, and "∪" indicates union. Using Theorem D.1, we get

$$\frac{0.6\tilde{C}_1 + 0.4\tilde{C}_2}{0.6 + 0.4} = [2.2, 3.2] \quad (5.25)$$

and

$$\frac{0.8\tilde{C}_1 + 0.4\tilde{C}_2}{0.8 + 0.4} = [2, 3] \quad (5.26)$$

Using (5.25) and (5.26) in (5.24), we get

$$\tilde{Y}_{cos}(\mathbf{x}) = \tilde{Y}_1 \cup \tilde{Y}_2 \quad (5.27)$$

where

$$\tilde{Y}_1 = 0.5/[2.2, 3.2] \quad (5.28)$$

and

$$\tilde{Y}_2 = 1/[2, 3] \quad (5.29)$$

$\tilde{Y}_1$, $\tilde{Y}_2$ and $\tilde{Y}_{cos}(\mathbf{x})$ are depicted in Fig. 5.7 (a).

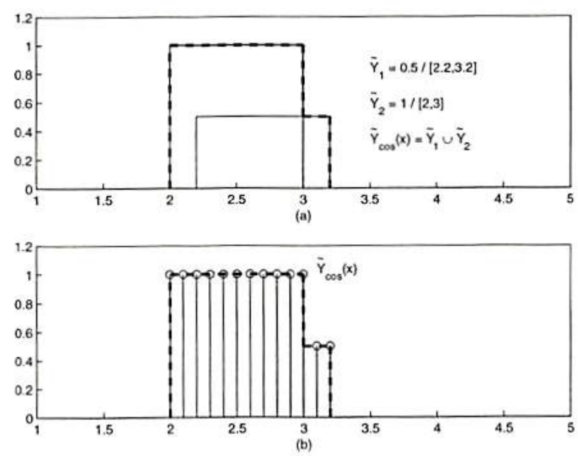


Figure 5.7: The type-reduced set, $\tilde{Y}_{cos}(\mathbf{x})$, for Example 5.3 computed using Eq. (5.23) [Fig. (a)] and Eq. (5.22) [Fig. (b)]. In Fig. (a), $\tilde{Y}_{cos}(\mathbf{x})$ (thick dashed line) is obtained as the union of two type-1 sets (thin solid lines); and, in Fig. (b), it is obtained as the union of many fuzzy singletons.

If we use (5.22) to compute $\tilde{Y}_{cos}(\mathbf{x})$, it discretizes $\tilde{C}_1$ and $\tilde{C}_2$, and follows the center-of-sets computational procedure described earlier in this section. This procedure computes $\tilde{Y}_{cos}(\mathbf{x})$ point by point, and the same type-reduced set, given by (5.27), is obtained as a union of many fuzzy singletons [see Fig. 5.7 (b)], where by a "fuzzy singleton", we mean a fuzzy set consisting of a single point. The exact number of fuzzy singletons depends upon the number of points in the discretized centroids, $\tilde{C}_1$ and $\tilde{C}_2$. Obviously, in this case, (5.23) computes $\tilde{Y}_{cos}(\mathbf{x})$ much more efficiently than (5.22). $\square$

If only the consequents are type-2, i.e., if all the degrees of firing are crisp [$\tilde{E}_l = \mathcal{T}_{i=1}^{p} \mu_{\tilde{F}_i^l}(x_i)$], (5.23) reduces to

$$\tilde{Y}_{cos}(\mathbf{x}) = \frac{\sum_{l=1}^{M} \tilde{C}_l \mathcal{T}_{i=1}^{p} \mu_{\tilde{F}_i^l}(x_i)}{\sum_{l=1}^{M} \mathcal{T}_{i=1}^{p} \mu_{\tilde{F}_i^l}(x_i)} \tag{5.30}$$

where $\mathcal{T}$ indicates the chosen t-norm and both the summations indicate algebraic sum. If the $\tilde{C}_l$'s are interval sets or Gaussians, $\tilde{Y}_{cos}(\mathbf{x})$, in this case, can be computed by one application of Theorem D.1 or 2.4. See Example 5.7 for an example of this case.

As mentioned earlier, if, in (5.22), the $\tilde{E}_l$'s and $\tilde{C}_l$'s both are interval type-1 sets, the computational procedure described in Appendix D can be used to compute the type-reduced set. In Example 5.3, $\tilde{E}_1$ is not an interval type-1 set, which is why we used (5.23) to compute the type-reduced set.

5.3.6 A Comparison of Height and Center-of-Sets Type-Reducers

In a type-1 FLS, height defuzzification is computationally inexpensive and gives satisfactory results (see, for example, [2, 24]). In a type-2 FLS, however, height type-reduction may not perform so well. The center-of-sets type-reduction does a better job. Here we explain why this is so.

For expressions for the height and center-of-sets type-reduced sets see (5.17) and (5.22), respectively.

When only one rule is fired, corresponding to $l = l'$, the height type-reduced set is

$$\begin{aligned}
\tilde{Y}_h(\mathbf{x}) &= \int_{\theta_{l'}} \mu_{\tilde{D}^{l'}}(\theta_{l'}) \Bigg/ \frac{\bar{y}^{l'}\theta_{l'}}{\theta_{l'}} \\
&= \int_{\theta_{l'}} \mu_{\tilde{D}^{l'}}(\theta_{l'}) \Bigg/ \bar{y}^{l'} \\
&= \left[\sup_{l'} \mu_{\tilde{D}^{l'}}(\theta_{l'})\right] / \bar{y}^{l'} \\
&= 1/\bar{y}^{l'} \\
&= \bar{y}^{l'}
\end{aligned} \tag{5.31}$$

The next to the last step assumes that all the involved membership grades are normal, which is usually the case. Equation (5.31) shows that when a single rule is fired, the height type-reduced set collapses to a single point ! This is certainly undesirable, because it means that when the input is such that only one rule is fired, no uncertainty is associated with the output, which is generally not true.

We invented the center-of-sets type-reducer to avoid this problem. It is easy to see that, even when only a single rule is fired, $\tilde{Y}_{cos}(\mathbf{x})$ is a type-1 fuzzy set, which equals the centroid of the type-2 output set corresponding to the fired rule, i.e.,

$$\begin{aligned}
\tilde{Y}_{cos}(\mathbf{x}) &= \int_{d_{l'}} \int_{e_{l'}} \mu_{\tilde{C}_{l'}}(d_{l'}) \mu_{\tilde{E}_{l'}}(e_{l'}) \Bigg/ \frac{d_{l'} e_{l'}}{e_{l'}} \\
&= \int_{d_{l'}} \int_{e_{l'}} \mu_{\tilde{C}_{l'}}(d_{l'}) \mu_{\tilde{E}_{l'}}(e_{l'}) \Bigg/ d_{l'}
\end{aligned}$$

$$
\begin{aligned}
&= \int_{d_{l'}} \mu_{\tilde{C}_{l'}}(d_{l'}) \int_{e_{l'}} \mu_{\tilde{E}_{l'}}(e_{l'}) \Big/ d_{l'} \\
&= \int_{d_{l'}} \mu_{\tilde{C}_{l'}}(d_{l'}) \big[\sup_{e_{l'}} \mu_{\tilde{E}_{l'}}(e_{l'})\big] \Big/ d_{l'} \\
&= \int_{d_{l'}} \mu_{\tilde{C}_{l'}}(d_{l'}) \Big/ d_{l'} \\
&= \tilde{C}_{l'}
\end{aligned}
\tag{5.32}
$$

where we have again assumed that the degree of firing associated with the l'th consequent is normal. (If each $\tilde{\mu}_{\tilde{F}_i^l}(x_i)$ is normal, $\tilde{E}_l = \sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{F}_i^l}(x_i)$ will also be normal. See Chapter 2 for a detailed discussion of the *meet* operation.)

5.3.7 Center-of-Sets Method and TSK Systems

The most popular form of a TSK fuzzy logic system [21, 22] uses rules of the form :

$$R^l : \text{IF } x_1 \text{ is } \tilde{F}_1^l \text{ and } x_2 \text{ is } \tilde{F}_2^l \text{ and } \cdots \text{ and } x_p \text{ is } \tilde{F}_p^l, \text{ THEN } y^l = c_0^l + c_1^l x_1 + \cdots + c_p^l x_p.$$

Given an input $\mathbf{x}$, the output is inferred as

$$y = \frac{\sum_{l=1}^{M} y^l \mathcal{T}_{l=1}^{M} \mu_{\tilde{F}_i^l}(x_i)}{\sum_{l=1}^{M} \mathcal{T}_{l=1}^{M} \mu_{\tilde{F}_i^l}(x_i)} \tag{5.33}$$

where M is the total number of rules. Observe that, if we set $c_i^l = 0$ for $i = 1, \ldots, p$, then (5.33) reduces to

$$y = \frac{\sum_{l=1}^{M} c_0^l \mathcal{T}_{l=1}^{M} \mu_{\tilde{F}_i^l}(x_i)}{\sum_{l=1}^{M} \mathcal{T}_{l=1}^{M} \mu_{\tilde{F}_i^l}(x_i)}, \tag{5.34}$$

which is the same as the output of a type-1 Mamdani FLS using center-of-sets defuzzification [see (5.21)], with c_0^l's replacing the centroids of the consequent sets. We can, therefore, think of a type-1 Mamdani FLS using center-of-sets defuzzification as a TSK system in its simplest form. Note, from (5.14), (5.15) and (5.34), that the TSK defuzzifier is structurally different from a height defuzzifier, because of the factor $\mu_{\tilde{G}^l}(\bar{y}^l)$, which appears in the height defuzzifier.

Now imagine a type-2 version of the TSK FLS in (5.34), which uses type-1 fuzzy sets $\tilde{C}_0^l$'s in place of the crisp numbers c_0^l's, everything else being the same as in (5.34). We show, next, that this type-2 TSK FLS, again, has the same structure as a Mamdani type-2 FLS using center-of-sets type-reduction, when only the consequent sets are type-2.

The lth rule of the type-2 TSK FLS is of the form :

$$R_{TSK}^l : \text{IF } x_1 \text{ is } \tilde{F}_1^l \text{ and } x_2 \text{ is } \tilde{F}_2^l \text{ and } \cdots \text{ and } x_p \text{ is } \tilde{F}_p^l, \text{ THEN } y^l = \tilde{C}_0^l;$$

and, given an input $\mathbf{x}$, the output is inferred as

$$\tilde{Y}_{TSK} = \frac{\sum_{l=1}^{M} \tilde{C}_0^l \mathcal{T}_{l=1}^{M} \mu_{\tilde{F}_i^l}(x_i)}{\sum_{l=1}^{M} \mathcal{T}_{l=1}^{M} \mu_{\tilde{F}_i^l}(x_i)}, \tag{5.35}$$

The expression for $\tilde{Y}_{cos}(\mathbf{x})$, when only the consequents are type-2 is given by (5.30). We reproduce it here for convenience.

$$\tilde{Y}_{cos}(\mathbf{x}) = \frac{\sum_{l=1}^{M} \tilde{C}_l \mathcal{T}_{i=1}^{p} \mu_{\tilde{F}_i^l}(x_i)}{\sum_{l=1}^{M} \mathcal{T}_{i=1}^{p} \mu_{\tilde{F}_i^l}(x_i)} \tag{5.36}$$

where $\mathcal{T}$ indicates the chosen t-norm and both the summations indicate algebraic sum.

From (5.35) and (5.36), it is clear that the type-2 TSK FLS which has type-1 fuzzy sets as consequents (see the rule R_{TSK}^l above) has the same structure as a Mamdani type-2 FLS using center-of-sets type-reduction, when only the consequent sets are type-2.

Since the contribution from the antecedents (i.e., the degree of firing of each consequent) is also computed in exactly the same manner in the type-2 TSK and Mamdani FLS's, both the systems will still have the same form when the antecedent sets are also type-2. *In short, a type-2 Mamdani FLS using center-of-sets type-reduction, having rules of the form "IF $\cdots$, THEN y is $\tilde{\tilde{G}}^l$", is the same as the type-2 TSK FLS, whose rules have the same "IF" part as the Mamdani FLS and whose rule consequents are of the form "THEN $y = \tilde{C}^l$", provided that $\tilde{C}^l$'s are centroids of the sets $\tilde{\tilde{G}}^l$'s.*

The TSK method is much more direct than the center-of-sets method. Rather than going through the exercise of creating a consequent fuzzy set, finding its centroid, and then using the centroid to represent it, we can just start with the centroid (i.e., a crisp number in the type-1 case and a type-1 set in the type-2 case) in the TSK method. This also gives us more flexibility in the sense that it may be difficult to create a fuzzy set having a particular centroid for the center-of-sets method; but in the TSK method, we don't have to worry about creating the complete consequent set; all we need is just the centroid.

5.3.8 Some Examples that Illustrate Type-Reduction Methods

Example 5.4 [9] Consider a type-1 FLS having consequent fuzzy sets as shown in Fig. 5.8 (a). Suppose that for some particular input $\mathbf{x}$ the fired outputs are as shown in Fig. 5.8 (b) (assuming product inference). The numbers 0.9, 0.8 and 0.2 indicate the degree of firing of each of the consequent sets. The outputs of the aforementioned defuzzifiers for this example are listed on this figure. For the modified height defuzzifier, δ^l was set equal to the standard deviation of the lth consequent set. The standard deviations for the three fired consequent sets are equal to 0.4, 0.2 and 0.2, respectively. Observe that for the consequent set centered at 5, $\bar{y}^l = 5$, but $c^l = 4.8436$; therefore, the outputs of the height and center-of-sets defuzzifiers for this example are slightly different. □

Example 5.5 [9] Now, suppose that we have a type-2 version of the problem considered in Example 5.4, where the antecedent sets are type-2 and consequent sets are type-1 and are the same as those shown in Fig. 5.8 (a) [shown again in Fig. 5.9 (a)]. The fired output sets for some input $\mathbf{x}$ are shown in Fig. 5.9 (b). The degrees of firing $[\sqcap_{i=1}^{p} \mu_{\tilde{F}_i^l}(x_i)]$ of the three fired consequents are $1/0.9 + 0.5/0.8$, $1/0.8 + 0.8/0.6$ and $1/0.2 + 0.4/0.1$, respectively. We assume product inference and product t-norm. The output sets shown in Fig. 5.9 (b) are computed using (5.4). We show the calculation for $\tilde{\tilde{B}}^1$ here. Let $G(y; m, \sigma) = \exp\left\{ -\frac{1}{2}\left(\frac{y-m}{\sigma}\right)^2 \right\}$; then,

$$\begin{aligned}
\tilde{\mu}_{\tilde{\tilde{B}}^1}(y) &= \mu_{\tilde{G}^1}(y) \sqcap (1/0.9 + 0.5/0.8) \\
&= 1\big/G(y; 2, 0.4) \sqcap (1/0.9 + 0.5/0.8) \\
&= 1\big/[0.9G(y; 2, 0.4)] + 0.5\big/[0.8G(y; 2, 0.4)]
\end{aligned} \tag{5.37}$$

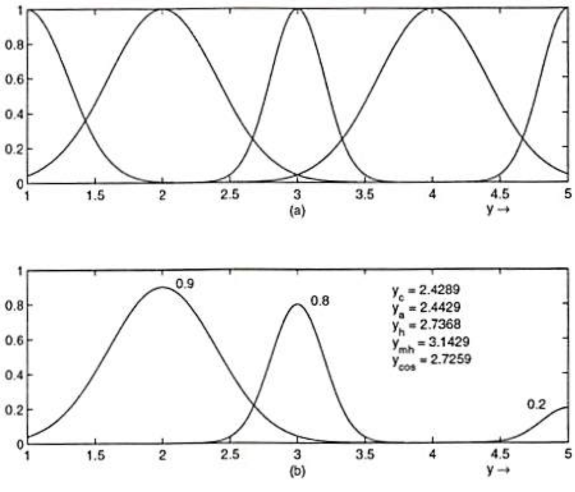


Figure 5.8: (a) Consequent sets for Example 5.4. (b) Fired consequent sets for some input **x**. Outputs of different defuzzifiers are listed : centroid (y_c), center-of-sums (y_a), height (y_h), modified height (y_{mh}) and center-of-sets (y_{cos}).

Each point y in the domain of $\tilde{\tilde{B}}^1$ has two primary memberships, one equal to $0.9G(y;2,0.4)$ and the other equal to $0.8G(y;2,0.4)$. The corresponding secondary memberships are 1 and 0.5, respectively. Output sets $\tilde{\tilde{B}}^2$ and $\tilde{\tilde{B}}^3$ are computed in a similar manner. The primary and secondary memberships in each case are shown in Fig. 5.9 (b). Observe the difference between Fig. 5.9 (b) and Fig. 5.8 (b), where, in the latter figure, each output set has a fixed height.

We now discuss in detail how each of the aforementioned type-reducers computes the type-reduced set.

1. **Centroid type-reducer** : To use a centroid type-reducer, we begin by finding a composite output fuzzy set which equals the union of the individual output sets. At every point y in the range of the type-2 FLS, we find the *join* between the membership grades of y in all three fired rule sets. See Fig. 5.10 (a).

   While finding the union of all these sets, since $\tilde{\tilde{B}}^3$ does not overlap with any of the other two sets, it will remain exactly as it is. Let's see how to compute the union of $\tilde{\tilde{B}}^1$ and $\tilde{\tilde{B}}^2$. For some point y (along the horizontal axis), let $\tilde{\mu}_{\tilde{\tilde{B}}^1}(y) = 0.5/a_1 + 1/a_2$ and $\tilde{\mu}_{\tilde{\tilde{B}}^2}(y) = 0.8/b_1 + 1/b_2$, where the sum indicates logical union. Then, from (2.17),

$$\begin{aligned}\tilde{\mu}_{\tilde{\tilde{B}}^1}(y) \sqcup \tilde{\mu}_{\tilde{\tilde{B}}^2}(y) &= (0.5/a_1 + 1/a_2) \sqcup (0.8/b_1 + 1/b_2) \\ &= (0.5 \times 0.8)/(a_1 \vee b_1) + (0.5 \times 1)/(a_1 \vee b_2) \\ &\quad +(1 \times 0.8)/(a_2 \vee b_1) + (1 \times 1)/(a_2 \vee b_2) \end{aligned} \tag{5.38}$$

   Observe that, to the left of the two dashed lines in Fig. 5.10 (a), each of a_1 and a_2 is greater than either b_1 or b_2, respectively; consequently, in this region, (5.38) gives

$$\begin{aligned}\tilde{\mu}_{\tilde{\tilde{B}}^1}(y) \sqcup \tilde{\mu}_{\tilde{\tilde{B}}^2}(y) &= 0.4/a_1 + 0.5/a_1 + 0.8/a_2 + 1/a_2 \\ &= 0.5/a_1 + 1/a_2 \\ &= \tilde{\mu}_{\tilde{\tilde{B}}^1}(y) \end{aligned} \tag{5.39}$$

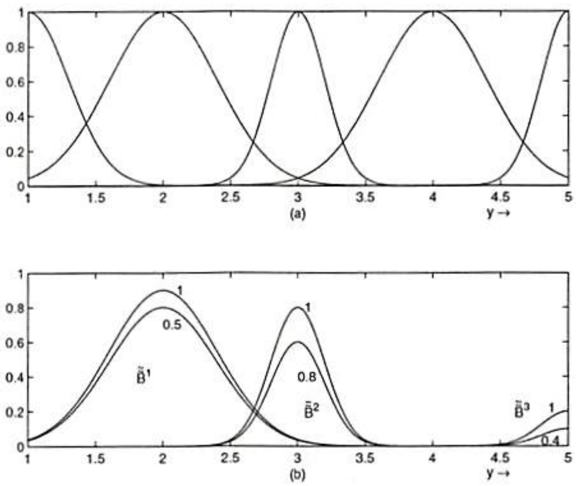


Figure 5.9: (a) Consequent sets for Example 5.5 (which are the same as the consequent sets for Example 5.4). (b) Fired consequent sets for some input **x**. The fired consequent sets are type-2, since the degrees of firing associated with each of the consequent sets are assumed to be type-1 sets. The vertical axis shows primary memberships of y in the output sets. Secondary memberships are indicated on the figure.

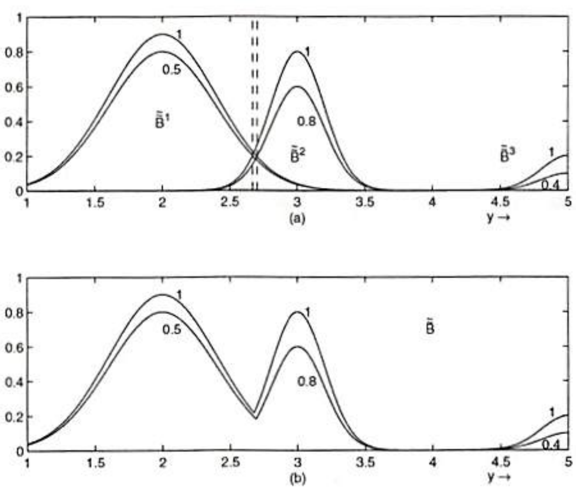


Figure 5.10: (a) The fired consequent sets for Example 5.5. The dashed lines show the region within which a point may have more than two primary memberships. (b) Approximate output set for the centroid type-reducer, found by taking the union of all the fired consequent sets. In both (a) and (b), the vertical axis shows primary memberships of y in the output sets. Secondary memberships are indicated on the figure.

Similarly, to the right of the two dashed lines, each of b_1 and b_2 is greater than either a_1 or a_2, respectively; consequently,

$$\tilde{\mu}_{\tilde{\tilde{B}}^1}(y) \sqcup \tilde{\mu}_{\tilde{\tilde{B}}^2}(y) = \tilde{\mu}_{\tilde{\tilde{B}}^2}(y) \tag{5.40}$$

In between the two dashed lines, the union will, in general, be different from either $\tilde{\mu}_{\tilde{\tilde{B}}^1}(y)$ or $\tilde{\mu}_{\tilde{\tilde{B}}^2}(y)$. In this region, a point may have more than two primary memberships [as given by (5.38)]; however, since this region is very small compared to the range of the FLS, we assume, for simplicity, that even in between the two dashed lines, every point has only two primary memberships [shown in Fig. 5.10 (b)] and that the composite output set $\tilde{\tilde{B}}$ is as depicted in Fig. 5.10 (b).

Now, for the purpose of centroid calculation, we discretize $\tilde{\tilde{B}}$ into 10 points [see Fig. 5.11 (a)]. Each of these points has 2 primary memberships; consequently, the total number of combinations that will need to be calculated (i.e., number of points in the type-reduced set) will be equal to 2^{10}, which is more than one thousand ! This shows how computationally intensive the centroid type-reducer can become. Figure 5.11 (b) shows the centroid type-reduced set for this example. Each point in this set is calculated by choosing one value of primary membership for each of the 10 points on the horizontal axis and finding the centroid of the resulting curve. The membership of this point is set equal to the minimum of all the secondary memberships corresponding to the chosen primary memberships (recall that we use the minimum t-norm for calculating the centroid of a type-2 set having a continuous domain - see Section 2.5.1). For a sample calculation, see the height type-reducer (item 3) below. [The differences in the calculations for centroid and height type-reducers are : the height type-reducer considers only three points on the horizontal axis (analogous to discretizing the entire set into only three points), whereas the centroid type-reducer considers 10 points on the horizontal axis; and, the t-norm operation between the secondary memberships in case of the centroid type-reducer is minimum instead of product.] The point having unity membership in the type-reduced set, 2.4729, corresponds to the centroid of the principal membership function of the combined output set (which corresponds to the upper curve in this case), which equals the centroid defuzzified value in Example 5.4. The slight difference in these two values is due to a slight difference in discretization of the domain.

2. **Center-of-sums type-reducer** : The center-of-sums type-reducer combines the output sets by summing them and then finds the centroid of this combined output set. The type-reduced set is, then, computed using (5.7). Figure 5.12 (a) shows the combined output set found by summing the individual sets (see Section 2.4 for more discussion on addition of fuzzy quantities). Its centroid is found by discretizing it in the same manner as in the case of the centroid type-reducer. The type-reduced set is depicted in Fig. 5.12 (b). The point having unity membership in the type-reduced set, 2.4172, corresponds to the centroid of the principal membership function of the combined output set, which equals the center-of-sums defuzzified value in Example 5.4. The slight difference in these two values is due to a slight difference in discretization of the domain.

3. **Height type-reducer** : The height type-reducer replaces each fired rule output set by a singleton at the point having maximum membership in that output set. In this example, there are 3 $\bar{y}^l$'s corresponding to the three fired output sets, each $\bar{y}^l$ having two possible primary memberships. The height type-reducer considers each of the eight possible combinations and performs height defuzzification on them using (5.17) to get points in the type-reduced set. The membership of each point in the type-reduced set is calculated by performing the t-norm (product, in this example) between the corresponding secondary membership values.

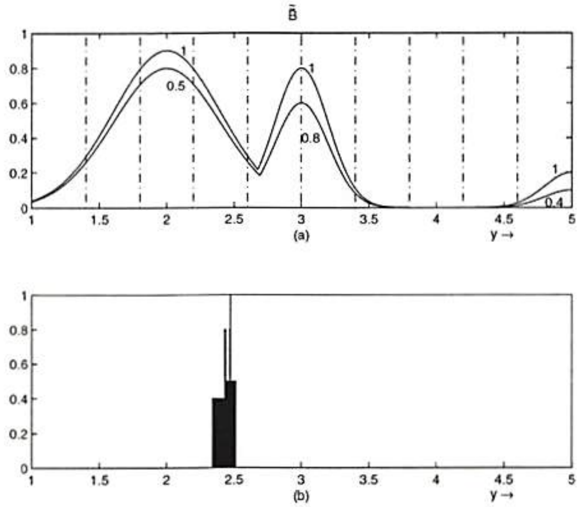


Figure 5.11: (a) The discretized combined output set, and (b) corresponding centroid type-reduced set for Example 5.5. The point having unity membership in the type-reduced set is equal to 2.4729 and the centroid of the type-reduced set is equal to 2.4321. In (a), the vertical axis shows primary memberships of y in the combined output set $\tilde{\tilde{B}}$; and in (b), the vertical axis shows memberships of y in the centroid type-reduced set.

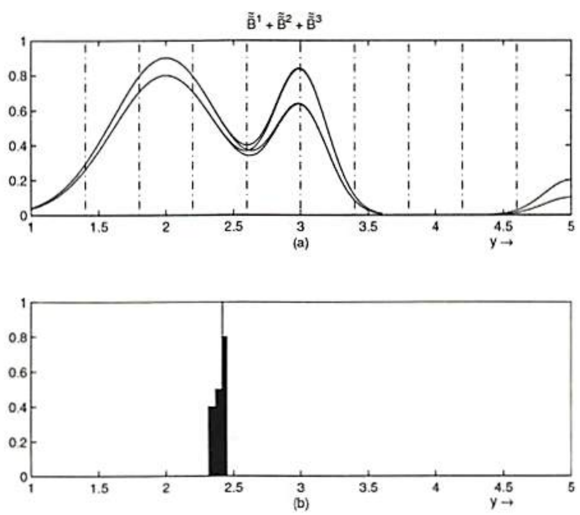


Figure 5.12: (a) The discretized combined output set, and (b) corresponding center-of-sums type-reduced set for Example 5.5. The point having unity membership in the type-reduced set is equal to 2.4172 and the centroid of the type-reduced set is equal to 2.3856. In (a), the vertical axis shows primary memberships of y in the combined output set $\tilde{\tilde{B}}^1 + \tilde{\tilde{B}}^2 + \tilde{\tilde{B}}^3$; and in (b), the vertical axis shows memberships of y in the center-of-sums type-reduced set.

We show the calculations for two points here. First, consider the situation where the first, second and third consequent sets have heights equal to 0.9, 0.6, and 0.2, respectively. The corresponding point in the type-reduced set is calculated, using (5.17) and Fig. 5.13 (a), as $(1 \times 0.8 \times 1)/(\frac{0.9\times2+0.6\times3+0.2\times5}{0.9+0.6+0.2}) = 0.8/2.7059$. Next, observe that the point having maximum membership in the type-reduced set is calculated as $(1 \times 1 \times 1)/(\frac{0.9\times2+0.8\times3+0.2\times5}{0.9+0.8+0.2}) = 1/2.7368$, which agrees with the height defuzzified value for Example 5.4. The complete type-reduced set is depicted in Fig. 5.13 (b). Note that it ranges from 2.5625 to 2.7778, and is *not* centered at 2.7368.

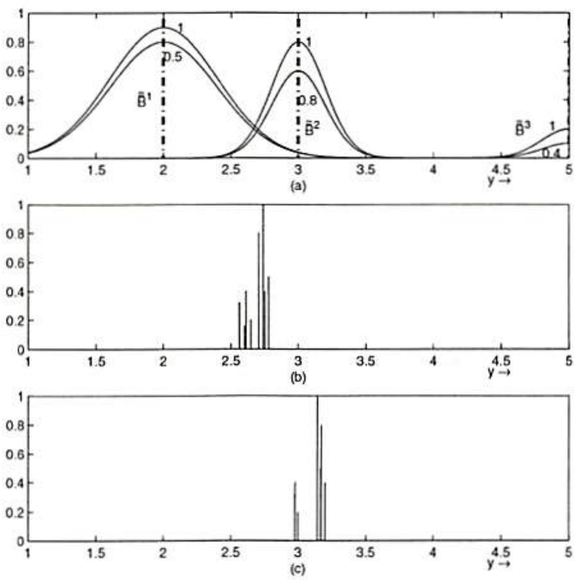


Figure 5.13: (a) The $\bar{y}^l$'s, and their output set memberships; (b) the height type-reduced set; and, (c) the modified height type-reduced set for Example 5.5. The point having unity membership in the height type-reduced set is equal to 2.7368 and its centroid is equal to 2.6985. The corresponding values for the modified height type-reduced set are 3.1429 and 3.1125, respectively. In (a), the vertical axis shows primary memberships of y in the output sets; and, in (b) and (c), the vertical axis shows memberships of y in the height type-reduced set.

4. **Modified height type-reducer** : The procedure for computing the modified height type-reduced set is exactly the same as that for computing the height type-reduced set, the only difference being that the modified height type-reducer also scales the output set membership of each $\bar{y}^l$ by $1/{\delta^l}^2$, where δ^l in this example is set equal to the standard deviation of the Gaussian with secondary membership equal to 1 for each of the output sets. For the sets centered at 2, 3 and 5, δ^l is equal to 0.4, 0.2 and 0.2, respectively. Figure 5.13 (c) shows the modified height type-reduced set.

5. **Center-of-sets type-reducer** : The center-of-sets type-reducer computes the centroid of each consequent set using (2.69), just like the centroid type-reducer computes the centroid of the combined output set. The type-reduced set is calculated using the centroids and the degrees of firing of each of the consequent sets, $\sqcap_{i=1}^{p}\tilde{\mu}_{\tilde{F}_i^l}(x_i)$. The centroids of the three consequent sets are 2, 3 and 4.8436 [the centroids are crisp because the consequent sets are type-1 : see Fig. 5.9 (a)]; and, their corresponding degrees of firing are $1/0.9 + 0.5/0.8$, $1/0.8 + 0.8/0.6$ and $1/0.2 + 0.4/0.1$, respectively. Figure 5.14 (b) shows the center-of-sets type-reduced set. The unity membership point in this set is equal to 2.7204. The slight

difference between this value and the center-of-sets defuzzified value in Example 5.4 is due to a slight difference in the discretization of the domain while computing the centroids of the consequent sets. The center-of-sets type-reduced set differs slightly from the height type-reduced set, because the centroid of the consequent set centered at 5 is different from its unity membership point [this set is not symmetric - see Fig. 5.9 (a)].

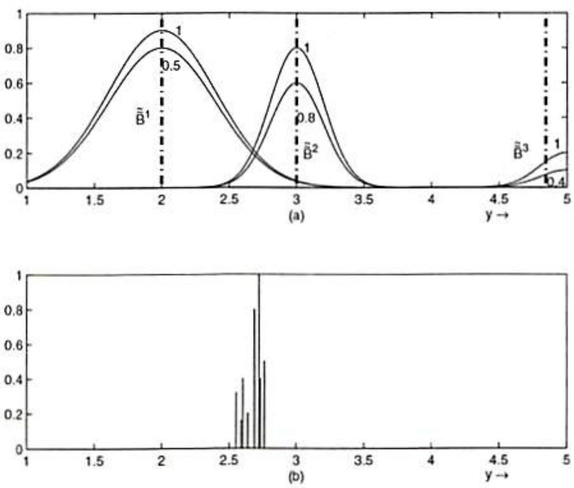


Figure 5.14: (a) The c^l's (which are crisp in this case, because the consequent sets are assumed to be type-1), and their output set memberships; and (b) corresponding center-of-sets type-reduced set for Example 5.5. The point having unity membership in the type-reduced set is equal to 2.7204 and the centroid of the type-reduced set is equal to 2.6832. In (a), the vertical axis shows primary memberships of y in the output sets; and in (b), the vertical axis shows memberships of y in the center-of-sets type-reduced set.

Table 5.1 summarizes the results of this example. Just as different defuzzification methods provide different results, different type-reduction methods also provide different results. Which type-reduction method to choose is an open issue, as is which defuzzification method to choose. □

Table 5.1: Results of Example 5.5.

Type-reduced set	Leftmost point	Rightmost point	Width	Unity height point	Centroid
Centroid	2.3403	2.5114	0.1711	2.4729	2.4321
Center-of-sums	2.3167	2.4511	0.1344	2.4172	2.3856
Height	2.5625	2.7778	0.2153	2.7368	2.6985
Modified height	2.9730	3.2000	0.2270	3.1429	3.1125
Center-of-sets	2.5527	2.7604	0.2077	2.7204	2.6832

5.4 Significance of the Type-Reduced Set

Here we take a closer look at the output of a type-reducer, the type-reduced set. Consider, again, a p-input single-output type-2 FLS , that uses singleton fuzzification, product or minimum inference, and has M rules of the form

$$R^l : \text{IF } x_1 \text{ is } \tilde{\text{F}}_1^l \text{ and } x_2 \text{ is } \tilde{\text{F}}_2^l \text{ and } \cdots \text{ and } x_p \text{ is } \tilde{\text{F}}_p^l, \text{ THEN } y \text{ is } \tilde{\text{G}}^l.$$

where $x_i \in X_i$ and $y \in Y$.

The type-reduced set of this FLS is a type-1 fuzzy subset of Y. Assuming centroid type-reduction [Eq. (5.7)], the type-reduced set is the centroid of the output type-2 set of the type-2 FLS; consequently, as explained in Section 2.5, each element of the type-reduced set is the centroid of some type-1 set embedded in the output set of the type-2 FLS. Each of these embedded sets can be thought of as an output set of some type-1 FLS, and, correspondingly, the type-2 FLS can be thought of as a collection of many different type-1 FLS's. Each of these type-1 FLS's is *embedded* in the type-2 FLS (just as their output sets are embedded in the output set of the type-2 FLS); so, *the type-reduced set is a collection of the outputs of all the type-1 FLS's embedded in the type-2 FLS* (see Fig. 5.15). In the continuous case, just as the number of embedded type-1 sets in a type-2 set is not countable, the number of embedded type-1 FLS's in a type-2 FLS is also not countable.

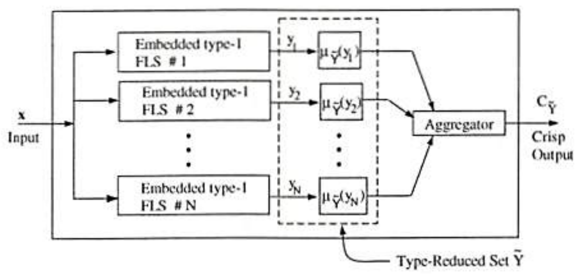


Figure 5.15: A type-2 FLS can be thought of as a collection of a large number of type-1 FLS's. The type-reduced set, $\tilde{Y}$ is a collection of the outputs of all these embedded type-1 FLS's. When the antecedent and consequent membership grades in the type-2 FLS have a continuous domain, the number of embedded type-1 FLS's is uncountable. This figure is drawn assuming that the membership grades have discrete (or discretized) domains. The memberships in the type-reduced set represent the level of uncertainty associated with each embedded type-1 FLS. A crisp output can be obtained by aggregating the outputs of all the embedded type-1 FLS's by, for example, finding the centroid of the type-reduced set (see Section 5.5).

Though we just discussed, and continue to discuss, the type-reduced set for a centroid type-reducer, these discussions are valid for any other kind of type-reducer. The type-reduced set for any type-2 FLS represents a collection of outputs of all the type-1 FLS's embedded in the type-2 FLS. It lets us represent the output of the type-2 FLS as a fuzzy set rather than as a crisp number, something *that can not be achieved with a type-1 FLS.*

The membership grade of a point $y_1 \in Y$ in the type-reduced set indicates the level of uncertainty associated with the type-1 FLS whose output is equal to y_1. This level of uncertainty is calculated from the secondary memberships associated with the specific type-1 set embedded in the type-2 output set. If more than one embedded type-1 FLS gives the same output, y_1, the membership corresponding to that point is taken to be the supremum of the levels of uncertainty associated with all the concerned embedded FLS's.

Observe, from (5.4), that if all the antecedent and consequent membership grades of the type-2 FLS are normal, and have only one point having unity secondary membership, then the output set membership grade of every $y \in Y$ will also be normal and will have only one point having unity secondary membership. Consequently, the type-reduced set [see (5.7)] will also be normal and will have only one point having unity membership. This point will correspond to the centroid of the principal membership function of the type-2 output set.

If all the type-2 uncertainties in the FLS were to collapse to type-1 uncertainties, i.e., if all the type-2 membership functions in the FLS were to collapse to their principal membership functions, the antecedent and consequent membership grades of each point would collapse to their unity membership points. This would cause the output type-2 set to collapse to its principal membership function, and the type-reduced set to collapse to a single point. This point would correspond to the centroid of the principal membership function of the output set. This shows that all our results are valid when all the type-2 uncertainties collapse to type-1 uncertainties.

We can think of a type-2 FLS as a "perturbed" version of a type-1 FLS, due to uncertainties in the membership functions. For example, when using Gaussian type-2 sets, we can think of the type-2 FLS as a perturbed form of a type-1 FLS whose membership functions are the principal membership functions of the type-2 FLS. The type-reduced set of the type-2 FLS can then be thought of as representing the uncertainty in the crisp output due to uncertainties in the membership functions. Some measure of the spread of the type-reduced set may then be taken to indicate the possible variation in the crisp output due to variations in the membership function parameters. This is analogous to using confidence intervals in a stochastic-uncertainty situation; however, what we have developed is for *linguistic uncertainties.*

One may argue that, in a type-1 FLS, the output set before defuzzification can be used in place of the type-reduced set, so we don't need type-2 FLS's; but, to do so is incorrect. The output set of a type-1 FLS just represents a combination of all the rule outputs of a single type-1 FLS, whereas the type-reduced set for a type-2 FLS represents a collection of outputs of a large number of type-1 FLS's, each having a level of uncertainty associated with it, equal to the membership in the type-reduced set.

5.5 Defuzzification

We *defuzzify* the type-reduced set to get a crisp output from the type-2 FLS. The most natural way of doing this seems to be by finding the centroid of the type-reduced set. Finding the centroid is equivalent to finding a weighted average of the outputs of all the type-1 FLS's embedded in the type-2 FLS, where the weights correspond to the memberships in the type-reduced set (see Fig. 5.15).

If the type-reduced set has only one point having unity membership, and if we wish to reduce computational complexity, we may think that a more straightforward choice for the defuzzified value is the unity membership point in the type-reduced set. Choosing the unity membership point, however, means that we are doing away with all the type-2 analysis, and are choosing the output corresponding to only the principal membership function type-1 FLS that is embedded in the type-2 FLS. Since it conveys no information about membership function uncertainties, it does not make sense to use the unity height point as the crisp output; unless, of course, the type-reduced set is convex and symmetric, in which case the unity height point is the same as the centroid.

Here we indicate the type-reduced set for an input $\mathbf{x}$ as $\tilde{\mathrm{Y}}(\mathbf{x})$. It can be obtained by any of our type-reduction methods. If $\tilde{\mathrm{Y}}(\mathbf{x})$ is discrete (or if we discretize it), the expression for the centroid of this set can be written as

$$C_{\tilde{\mathrm{Y}}}(\mathbf{x}) = \frac{\sum_{k=1}^{N} y_k \mu_{\tilde{\mathrm{Y}}}(y_k)}{\sum_{k=1}^{N} \mu_{\tilde{\mathrm{Y}}}(y_k)} \tag{5.41}$$

where y_k is a point of the type-reduced set and the total number of points in the (discretized) type-reduced set is N. Since the type-reduced set $\tilde{\mathrm{Y}}$ is a function of $\mathbf{x}$, its centroid, $C_{\tilde{\mathrm{Y}}}$, is also a function of $\mathbf{x}$. Figure 5.16 shows an example of an arbitrary type-reduced set discretized for centroid calculation. In general, for arbitrary membership functions, the type-reduced set is not

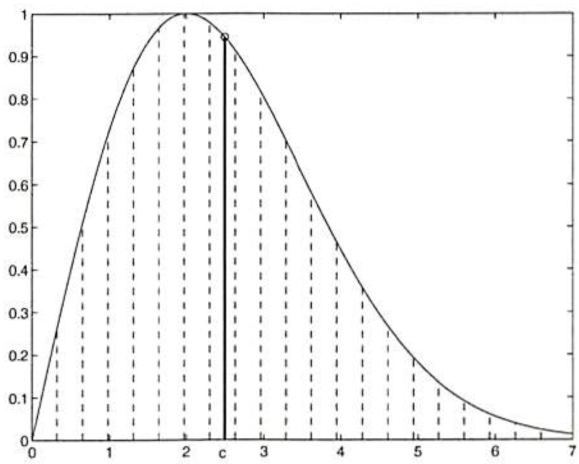


Figure 5.16: An arbitrary type-reduced set discretized to find the centroid. The centroid is indicated by point c in the figure.

symmetrical and the centroid location is different from the location of the unity height point, as demonstrated next, in :

Example 5.6 Consider the type-reduced sets for Example 5.5 shown in Figs. 5.11 - 5.14. A crisp output for each of these cases may be obtained by calculating the centroid of their type-reduced set. The crisp outputs for centroid, center-of-sums, height, modified height and center-of-sets type-reducers are listed in Table 5.1. Observe that these points are different than the points having unity membership in the type-reduced set; so, *antecedent and consequent uncertainties change the value of the defuzzified output of a type-2 FLS from that of a type-1 FLS*, and the type-reduced set establishes a band of values about the crisp value in much the same way that a confidence interval establishes a band about a point estimate. This band of values can be found in Table 5.1, under the column called "width". □

5.6 A Function Approximation Example of a Type-2 FLS

In this section, we present a function approximation example of a type-2 FLS which uses center-of-sets type-reduction. For simplicity, only the consequent sets are assumed to be type-2.

Example 5.7 The function to be approximated is $y = 100 - x^2$ for $x \in [-10, 10]$. Given are 10 realizations, each with 9 (x, y) pairs. Each of these pairs include values of y corrupted by additive noise which is uniformly distributed in $[-10, 10]$. For each applied input x^l $(l = 1, \ldots, 9)$, we find the minimum (y^l_{min}) and the maximum (y^l_{max}) of the 10 y values. The 9 $(x^l, [y^l_{min}, y^l_{max}])$ pairs are : $(x^1, [y^1_{min}, y^1_{max}]) = (-10, [-7.79, 6.49])$, $(x^2, [y^2_{min}, y^2_{max}]) = (-7.5, [34.72, 52.93])$, $(x^3, [y^3_{min}, y^3_{max}]) = (-5, [66.12, 84.1])$, $(x^4, [y^4_{min}, y^4_{max}]) = (-2.5, [84.93, 101.75])$, $(x^5, [y^5_{min}, y^5_{max}]) = (0, [93.09, 109.95])$, $(x^6, [y^6_{min}, y^6_{max}]) = (2.5, [88.02, 103.53])$, $(x^7, [y^7_{min}, y^7_{max}])$ $= (5, [65.37, 84.32])$, $(x^8, [y^8_{min}, y^8_{max}]) = (7.5, [34.14, 50.85])$, $(x^9, [y^9_{min}, y^9_{max}]) = (10, [-9.62, 9.62])$.

The FLS forms one rule from each pair. The rules are of the form,

$$\text{IF } x \text{ is } \tilde{\mathrm{A}}, \text{ THEN } y \text{ is } \tilde{\tilde{\mathrm{B}}}.$$

Observe that only the y_i's are uncertain in the given input-output pairs; therefore, we choose the antecedent sets as type-1 and the consequent sets as type-2. The antecedent sets are chosen to be type-1 Gaussian and the type-2 consequent sets are described as type-1 Gaussian fuzzy sets with uncertain means (as in Fig. 1.2). The FLS uses singleton fuzzification, *max* t-conorm, *product* t-norm, product inference and center-of-sets type-reduction.

The antecedent and consequent sets are depicted in Figs. 5.17 (a) and (b), respectively. Each antecedent set membership function is assigned a standard deviation equal to 1.25. Each consequent set membership function is obtained by perturbing the mean of a type-1 Gaussian fuzzy set. Every value of the mean is assumed to be equally uncertain and correspondingly, all the secondary memberships are set equal to 1, so that each consequent set is an interval type-2 set. For the lth rule, the mean of the consequent Gaussian is perturbed in the range $[y^l_{min}, y^l_{max}]$. In Example 2.5, we showed that the centroid of this type-2 set is a crisp set with domain approximately equal to $[y^l_{min}, y^l_{max}]$, when $(y^l_{max} - y^l_{min})$ is small compared to the standard deviation of the perturbed type-1 Gaussian set. In order to satisfy this condition, we have chosen this standard deviation equal to 40 [for this example, $\max_l(y^l_{max} - y^l_{min}) = 19.24$].

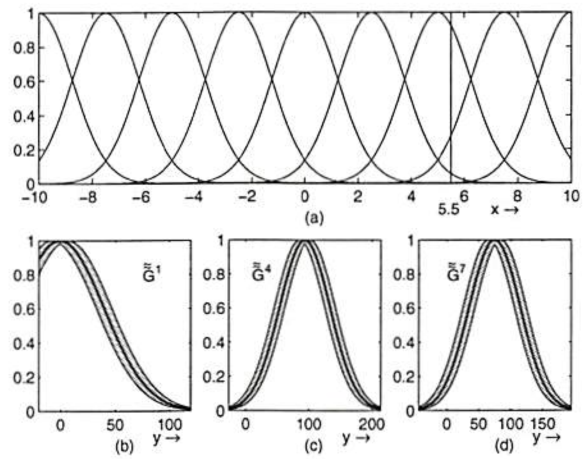


Figure 5.17: (a) Antecedent fuzzy sets for Example 5.7. Each antecedent set has the same principal membership standard deviation, equal to 1.25. Figures (b), (c) and (d) show the 1st, 4th and 7th consequent sets. All the consequent sets are interval type-2 sets and are obtained from type-1 Gaussian sets by letting the mean take values in an interval. Since the shaded regions of adjacent sets may overlap, it is not possible to show all the consequent sets on the same figure; therefore, we have arbitrarily shown three of them. The thick dark line, in each case, shows the Gaussian centered at $c^l = (y^l_{min} + y^l_{max})/2$. In (a), the position of a particular input point, $x = 5.5$, is also shown.

Figure 5.17 (a) shows the antecedent (type-1) membership grades for a particular input $x' = 5.5$. x' has non-zero memberships in three antecedent sets, 0.056135 in the set centered at 2.5 (x^6), 0.92312 in the set centered at 5 (x^7) and 0.27804 in the set centered at 7.5 (x^8); consequently, three rules are fired.

The center-of-sets type-reducer replaces each consequent set by its centroid and scales it (product inference) by the appropriate degree of firing. In this case, the centroids of the three fired consequents (associated with antecedent sets centered at x^6, x^7, and x^8, respectively) are crisp sets having domains $[88.02, 103.53]$, $[65.37, 84.32]$ and $[34.14, 50.85]$ respectively. The centroids and their corresponding degrees of firing are shown in Fig. 5.18 (a).

Calculations are simplified in this case because each $\tilde{C}^l$ is an interval type-1 set and the antecedent sets are type-1, so that the degrees of firing of each of the consequents are crisp. The center-of-sets type-reduced set depicted in Fig. 5.18 (b) is calculated using Theorem D.1. The type-reduced set is an interval type-1 set with domain $[59.39, 77.7]$. We can use the midpoint of this interval as the crisp output of the FLS. This value is equal to 68.55 and is shown with a dashed

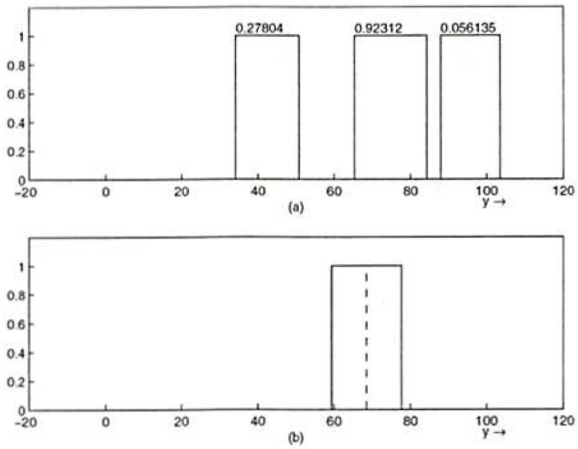


Figure 5.18: The input $x = 5.5$ has non-zero memberships in three antecedent sets. Corresponding to these three sets, three rules are fired. Figure (a) shows the centroids of the three fired sets and their corresponding degrees of firing. Figure (b) shows the corresponding center-of-sets type-reduced set and the crisp output value (shown by a dashed line).

line Fig. 5.18 (b). The actual function value at $x' = 5.5$ is 69.75. The two end-points of the domain of the type-reduced set indicate the lower and upper bounds on the crisp output value.

Figure 5.19 compares the true function value with the crisp output of the type-2 FLS and also shows the upper and lower bounds. For each value of x, these bounds correspond to the type-reduced set; they give a measure of the uncertainty in the approximation caused by noisy training values. □

Comment 1 : Next, we explain something that we observed in this example : *the output of the type-1 FLS, which uses the same antecedent sets as our type-2 FLS, type-1 consequent sets having centroids $c^l = (y^l_{min} + y^l_{max})/2$, and center-of-sets defuzzification, is the same as the crisp output of our type-2 FLS.*

Observe, from (5.30), that when only the consequents are type-2, $\tilde{Y}_{cos}(\mathbf{x})$ can be expressed as

$$\tilde{Y}_{cos}(\mathbf{x}) = \sum_{l=1}^{M} \frac{\mathcal{T}_{i=1}^{p} \mu_{\tilde{F}_i^l}(x_i)}{\sum_{l=1}^{M} \mathcal{T}_{i=1}^{p} \mu_{\tilde{F}_i^l}(x_i)} \tilde{C}_l \tag{5.42}$$

where both the summations denote algebraic sum.

Let $\phi^l(x) = [\mathcal{T}_{i=1}^{p} \mu_{\tilde{F}^l}(x)]/[\sum_{j=1}^{M} \mathcal{T}_{i=1}^{p} \mu_{\tilde{F}^j}(x)]$, then

$$\tilde{Y}_{cos}(x) = \sum_{l=1}^{M} \phi^l(x) \tilde{C}_l \tag{5.43}$$

Since $\tilde{C}_l = [y^l_{min}, y^l_{max}]$, $\phi^l(x)\tilde{C}_l = [\phi^l(x) y^l_{min}, \phi^l(x) y^l_{max}]$ (see Section 2.4.1), and, $\tilde{Y}_{cos}(x)$ is given as [see (2.64)]

$$\tilde{Y}_{cos}(x) = \left[\sum_{l=1}^{M} \phi^l(x) y^l_{min} \, , \, \sum_{l=1}^{M} \phi^l(x) y^l_{max} \right] \tag{5.44}$$

Since each $\tilde{C}_l$ is a crisp set, $\tilde{Y}_{cos}(x)$ is also a crisp set.

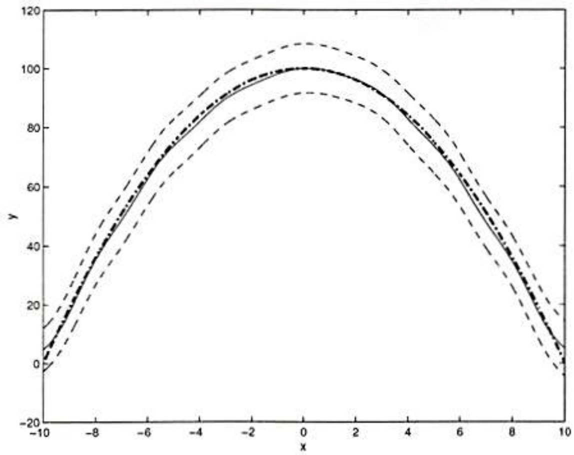


Figure 5.19: The solid line shows the crisp output of the type-2 FLS and the dashed lines indicate the upper and lower bounds. The true function value is shown by the thick dash-dotted line. The output of the type-1 FLS, which uses center-of-sets defuzzification and consequent sets having centroids $c^l = (y^l_{min} + y^l_{max})/2$, equals the crisp output of our type-2 FLS.

Using the fact that the centroid of an interval is its midpoint, the centroid of $\tilde{Y}_{cos}(x)$ can be expressed as

$$\begin{aligned} C_{\tilde{Y}_{cos}} &= \left[\sum_{l=1}^{M} \phi^l(x) y^l_{min} + \sum_{l=1}^{M} \phi^l(x) y^l_{max}\right] \Big/ 2 \\ &= \sum_{l=1}^{M} \phi^l(x) \left[\frac{y^l_{min} + y^l_{max}}{2}\right] \end{aligned} \tag{5.45}$$

which is the same as the output of a type-1 FLS that uses center-of-sets defuzzification and consequent sets having centroids $c^l = (y^l_{min} + y^l_{max})/2$. Observe that we did not have to use the fact that the number of antecedents, p, is equal to 1 in order to derive (5.45), which implies that this comment is applicable to multiple antecedent systems also. □

Comment 2 : Here we show that, for Example 5.7, the center-of-sets type-reducer is computationally more efficient than the height type-reducer.

Observe, from (5.44), that when using the center-of-sets type-reducer, $\tilde{Y}_{cos}(x)$ is an interval type-1 set; therefore, to obtain $\tilde{Y}_{cos}(x)$, the center-of-sets defuzzification calculation needs to be performed only twice, once for each end point of the interval.

For the height type-reduced set, the output set membership of each $\bar{y}^l$ is given as [(5.16)]

$$\tilde{\mu}_{\tilde{\tilde{B}}^l}(\bar{y}^l) = \mu_{\tilde{F}^l}(x) \sqcap \tilde{\mu}_{\tilde{\tilde{G}}^l}(\bar{y}^l) \tag{5.46}$$

Since the consequent set memberships of the $\bar{y}^l$'s [the $\tilde{\mu}_{\tilde{\tilde{G}}^l}(\bar{y}^l)$'s] are crisp sets, each $\tilde{\mu}_{\tilde{\tilde{B}}^l}(\bar{y}^l)$ is also a crisp set. Consequently, the expression for the height type-reduced set (5.17), in this case, becomes

$$\tilde{Y}_h(\mathbf{x}) = \int_{\theta_1} \int_{\theta_2} \cdots \int_{\theta_M} 1 \Big/ \frac{\sum_{l=1}^{M} \bar{y}^l \theta_l}{\sum_{l=1}^{M} \theta_l} \tag{5.47}$$

where each θ_l belongs to some interval in $[0,1]$, and we have made use of the fact that $\mathcal{T}_{l=1}^{M}\mu_{\tilde{\mathrm{D}}^l}(\theta_l) = 1$, because each $\tilde{\mathrm{D}}^l = \tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}^l}(\bar{y}^l)$ is a crisp set. Observe that the only simplification in (5.47) from (5.17) is that now we know that $\tilde{\mathrm{Y}}_h(\mathbf{x})$ is a crisp set. The domain points still have to be calculated. The computational procedure described in Appendix D.1 needs to be used for this purpose; consequently, calculation of the height type-reduced set in this example involves more computation than the center-of-sets type-reduced set. If, however, the intervals containing θ_l's are narrow, an approximate expression for (5.47) can be obtained (see Appendix D.2).

Observe that the number of antecedents need not necessarily be equal to 1 for this simplification; what is required is that all the antecedents be type-1. If the number of antecedents p is greater than 1, $\mu_{\tilde{\mathrm{F}}^l}(x)$ is replaced by $\mathcal{T}_{i=1}^{p}\mu_{\tilde{\mathrm{F}}_i^l}(x)$. Everything else remains the same. Observe, also, that in this example, the center-of-sets computation is simplified because we could approximate the centroid of each consequent set as $\tilde{\mathrm{C}}^l = [y_{min}^l, y_{max}^l]$. If the conditions required for such an approximation are not satisfied, we have to compute the consequent centroids using the computational procedure described in Appendix D.1; and, then, the center-of-sets method may turn out to be computationally more expensive than the height method. □

5.7 Type-Reduction for Gaussian Type-2 Fuzzy Logic Systems

In this section, we focus on type-2 FLS's whose antecedent and consequent membership functions are all Gaussian type-2 sets (recall that a Gaussian type-2 set is one whose secondary membership functions are all Gaussians). We refer to such FLS's as "Gaussian type-2 FLS's".

All the type-1 defuzzification methods described in Section 5.3 involve some kind of a weighted summation. In this section, we make use of Theorem 2.5 to find an approximate expression for the extension of a weighted summation to the case where all the quantities involved are type-1 Gaussian fuzzy numbers. This lets us find an approximate expression for the type-reduced set when certain conditions are satisfied, without having to perform the otherwise computationally intensive type-reduction calculations. In all the following cases, we assume that the condition (2.78) is satisfied; however, even if it is not satisfied, we will continue to use Theorem 2.5 to obtain an approximate expression for the type-reduced set because of the tremendous savings in computation that it offers. Whenever we need to compute the centroid of a Gaussian type-2 set having a continuous domain, we will assume that the t-norm used is minimum and use Corollary 2.2.

Consider a p-input single-output type-2 FLS , that uses singleton fuzzification, product t-norm and product inference, and has M rules of the form

$$R^l : \text{IF } x_1 \text{ is } \tilde{\tilde{\mathrm{F}}}_1^l \text{ and } x_2 \text{ is } \tilde{\tilde{\mathrm{F}}}_2^l \text{ and } \cdots \text{ and } x_p \text{ is } \tilde{\tilde{\mathrm{F}}}_p^l, \text{ THEN } y \text{ is } \tilde{\tilde{\mathrm{G}}}^l.$$

where $x_i \in X_i$, $y \in Y$, and each $\tilde{\tilde{\mathrm{F}}}_i^l$ and $\tilde{\tilde{\mathrm{G}}}^l$ (for $l = 1, \ldots, M$ and $i = 1, \ldots, n$) is a Gaussian type-2 set.

Let the mean and standard deviation of $\tilde{\mu}_{\tilde{\tilde{\mathrm{F}}}_i^l}(x_i)$ be $m_i^l(x_i)$ and $\sigma_i^l(x_i)$ for $x_i \in X_i$ ($l = 1, \ldots, M$ and $i = 1, \ldots, n$), and let the mean and standard deviation of $\tilde{\mu}_{\tilde{\tilde{\mathrm{G}}}^l}(y)$ be $m_G^l(y)$ and $\sigma_G^l(y)$ for each $y \in Y$ ($l = 1, \ldots, M$). Then, from (5.4) and (2.61), we see that the output set for every rule is also approximately a Gaussian type-2 set.

Under minimum t-norm, the degree of firing $\sqcap_{i=1}^{p}\tilde{\mu}_{\tilde{\tilde{\mathrm{F}}}_i^l}(x_i)$ may not remain a Gaussian; therefore, we consider only product t-norm here. Note that, if perchance the output set of a type-2 FLS using minimum t-norm is a Gaussian type-2 set, then Theorem 2.5 may be used.

1. **Centroid type-reduction** : The centroid type-reducer combines output sets for different rules by finding their union, so that the membership function of the combined output set,

$\tilde{\tilde{\mathrm{B}}}$, is as given in (5.6). Observe, from Theorem 2.3, that under maximum t-conorm, the *join* between two Gaussian type-1 sets may not remain a Gaussian. Consequently, the combined output set, $\tilde{\tilde{\mathrm{B}}}$, may not remain a Gaussian type-2 set; however, since the output set is always discretized for the centroid calculation, if all of the secondary membership functions in the discretized output set are Gaussian (i.e., if the points at which secondary membership functions are non-Gaussian, are not considered for centroid calculation), we can apply Corollary 2.2 to get an approximate expression for the centroid. It should be kept in mind though that, since we are excluding some points from the centroid calculation, at which the secondary membership functions are non-Gaussian, in some cases, the approximation may be poor.

If the combined output set is discretized into N points, $y_1, \ldots, y_N$, and if $m(y_i)$ and $\sigma(y_i)$ are the mean and standard deviation of the output set membership of y_i, (2.79) and (2.80) can be used to find an approximate expression for the type-reduced set. We repeat these equations here for convenience :

$$\mathcal{M}(\tilde{C}_{\tilde{\tilde{\mathrm{B}}}}) = \frac{\sum_{i=1}^{N} y_i m(y_i)}{\sum_{i=1}^{N} m(y_i)} \tag{5.48}$$

$$\Sigma(\tilde{C}_{\tilde{\tilde{\mathrm{B}}}}) = \frac{\sum_{i=1}^{N} |y_i - \mathcal{M}(\tilde{C}_{\tilde{\tilde{\mathrm{B}}}})| \sigma(y_i)}{\sum_{i=1}^{N} m(y_i)} \tag{5.49}$$

2. **Center-of-sums type-reduction** : The center-of-sums type-reducer combines output sets for different rules by summing them, so that the membership function of the combined output set, $\tilde{\tilde{\mathrm{B}}}$, is as given in (5.13). From Theorem 2.4 (Chapter 2), we know that the sum of Gaussian fuzzy numbers is also a Gaussian fuzzy number. Since each $\tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}^l}(y)$ is a Gaussian type-1 set, each $\tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}}(y)$ is also a Gaussian type-1 set and consequently the output set, $\tilde{\tilde{\mathrm{B}}}$, is a Gaussian type-2 set.

   Once the combined output set $\tilde{\tilde{\mathrm{B}}}$ is obtained, the procedure for finding the centroid (which is approximately a Gaussian type-1 set - see Corollary 2.2) is exactly the same as in the case of the centroid type-reducer. The centroid is a type-1 Gaussian with mean and standard deviation as given in (5.48) and (5.49), respectively.

   Observe that, in this case, unlike the centroid type-reducer, there is no need of any further approximation, such as excluding the points with non-Gaussian secondary membership functions.

3. **Height type-reduction** : The expression for the height type-reduced set [from (5.17)] is given as

$$\tilde{\mathrm{Y}}_h(\mathbf{x}) = \int_{\theta_1}\int_{\theta_2}\cdots\int_{\theta_M} \prod_{l=1}^{M} \mu_{\tilde{\mathrm{D}}^l}(\theta_l) \Big/ \frac{\sum_{l=1}^{M} \bar{y}^l \theta_l}{\sum_{l=1}^{M} \theta_l} \tag{5.50}$$

where $\theta_l \in \tilde{\mathrm{D}}^l = \tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}^l}(\bar{y}^l) = \tilde{\mu}_{\tilde{\tilde{\mathrm{G}}}^l}(\bar{y}^l) \sqcap [\sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{\tilde{\mathrm{F}}}_i^l}(x_i)]$ for $l = 1, \ldots, M$.

Using the Gaussian *meet* approximation in (2.61), we see that $\tilde{\mathrm{D}}^l$ is approximately a Gaussian type-1 set with mean $\mathcal{M}(\tilde{\mathrm{D}}^l)$ and standard deviation $\Sigma(\tilde{\mathrm{D}}^l)$, where

$$\mathcal{M}(\tilde{\mathrm{D}}^l) = m_G^l(\bar{y}^l) \prod_{i=1}^{p} m_i^l(x_i) \tag{5.51}$$

and

$$\Sigma(\tilde{\mathrm{D}}^l) = \Big[[\sigma_G^l(\bar{y}^l)]^2 \prod_i [m_i^l(x_i)]^2 + [\sigma_1^l(x_1)]^2[m_G^l(\bar{y}^l)]^2 \prod_{i;i\neq 1} [m_i^l(x_i)]^2 + \cdots$$
$$\cdots + [\sigma_p^l(x_p)]^2[m_G^l(\bar{y}^l)]^2 \prod_{i;i\neq p} [m_i^l(x_i)]^2\Big]^{\frac{1}{2}} \tag{5.52}$$

Using Theorem 2.5, we see that $\tilde{\mathrm{Y}}_h(\mathbf{x})$ is also approximately a Gaussian type-1 set with mean $\mathcal{M}_h(\mathbf{x})$ and standard deviation $\Sigma_h(\mathbf{x})$, where

$$\mathcal{M}_h(\mathbf{x}) = \frac{\sum_{l=1}^M \bar{y}^l \mathcal{M}(\tilde{\mathrm{D}}^l)}{\sum_{l=1}^M \mathcal{M}(\tilde{\mathrm{D}}^l)} \tag{5.53}$$

and

$$\Sigma_h(\mathbf{x}) = \frac{\sqrt{\sum_{l=1}^M \left[\bar{y}^l - \mathcal{M}_h(\mathbf{x})\right]^2 \Sigma^2(\tilde{\mathrm{D}}^l)}}{\sum_{l=1}^M \mathcal{M}(\tilde{\mathrm{D}}^l)} \tag{5.54}$$

The sequence of computations for height type-reduction is as follows : For given values of $x_1, \ldots, x_p$, we determine the mean, $m_i^l(x_i)$, and standard deviation, $\sigma_i^l(x_i)$, of each given antecedent membership grade, $\tilde{\mu}_{\tilde{\mathrm{F}}_i^l}(x_i)$, ($l = 1, 2, \ldots, M$ and $i = 1, 2, \ldots, p$). We are also given the mean, $m_G^l(\bar{y}^l)$, and standard deviation, $\sigma_G^l(\bar{y}^l)$, of the consequent membership grade, $\tilde{\mu}_{\tilde{\mathrm{G}}^l}(\bar{y}^l)$, for each $\bar{y}^l$. Then,

(a) For each l ($l = 1, 2, \ldots, M$), calculate the mean and standard deviation of $\tilde{\mu}_{\tilde{\mathrm{B}}^l}$ using (5.51) and (5.52); and,

(b) Compute the mean and standard deviation of the type-reduced set using (5.53) and (5.54).

4. **Modified height type-reduction** : The expression for the modified height type-reduced set [from (5.20)] is given as

$$\tilde{\mathrm{Y}}_{mh}(\mathbf{x}) = \int_{\theta_1} \int_{\theta_2} \cdots \int_{\theta_M} \prod_{l=1}^M \mu_{\tilde{\mathrm{D}}^l}(\theta_l) \Bigg/ \frac{\sum_{l=1}^M \bar{y}^l \theta_l / {\delta^l}^2}{\sum_{l=1}^M \theta_l / {\delta^l}^2} \tag{5.55}$$

where δ^l is some measure of spread (e.g., the standard deviation of the principal membership function) of the lth consequent set and all other symbols have the same meaning as in (5.50).

Proceeding as in case of the height type-reducer, we find that $\tilde{\mathrm{Y}}_{mh}(\mathbf{x})$ is approximately a Gaussian type-1 set with mean $\mathcal{M}_{mh}(\mathbf{x})$ and standard deviation $\Sigma_{mh}(\mathbf{x})$, where

$$\mathcal{M}_{mh}(\mathbf{x}) = \frac{\sum_{l=1}^M \bar{y}^l \mathcal{M}(\tilde{\mathrm{D}}^l) / {\delta^l}^2}{\sum_{l=1}^M \mathcal{M}(\tilde{\mathrm{D}}^l) / {\delta^l}^2}, \tag{5.56}$$

$$\Sigma_{mh}(\mathbf{x}) = \frac{\sqrt{\sum_{l=1}^M \left[\bar{y}^l - \mathcal{M}_{mh}(\mathbf{x})\right]^2 \Sigma^2(\tilde{\mathrm{D}}^l) / {\delta^l}^4}}{\sum_{l=1}^M \mathcal{M}(\tilde{\mathrm{D}}^l) / {\delta^l}^2}, \tag{5.57}$$

and $\mathcal{M}(\tilde{\mathrm{D}}^l)$ and $\Sigma(\tilde{\mathrm{D}}^l)$ are computed using (5.51) and (5.52), respectively.

The sequence of computations to be performed is the same as in the case of height type-reduction.

5. **Center-of-sets type-reduction** : The center-of-sets type-reduced set is given as [from (5.22)]

$$\tilde{Y}_{cos}(\mathbf{x}) = \int_{d_1} \cdots \int_{d_M} \int_{e_1} \cdots \int_{e_M} \prod_{l=1}^{M} \tilde{\mu}_{\tilde{C}_l}(d_l) \prod_{l=1}^{M} \tilde{\mu}_{\tilde{E}_l}(e_l) \Big/ \frac{\sum_{l=1}^{M} d_l e_l}{\sum_{l=1}^{M} e_l} \tag{5.58}$$

where $d_l \in \tilde{C}_l = \tilde{C}_{\tilde{\tilde{G}}^l}$ is the centroid of the lth consequent set, and $e_l \in \tilde{E}_l = \sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{\tilde{F}}_i^l}(x_i)$ is the degree of firing for the lth consequent set, for $l = 1, \ldots, M$.

From Corollary 2.2, we know that $\tilde{C}_{\tilde{\tilde{G}}^l}$ is approximately a Gaussian type-1 set. Let its mean and standard deviation be $\mathcal{M}(\tilde{C}_{\tilde{\tilde{G}}^l})$ [see (2.79)] and $\Sigma(\tilde{C}_{\tilde{\tilde{G}}^l})$ [see (2.80)], respectively. Using the Gaussian *meet* approximation in (2.61), we also have that the degree of firing $\tilde{E}_l$ is approximately a Gaussian type-1 set with mean $\mathcal{M}(\tilde{E}_l)$ and standard deviation $\Sigma(\tilde{E}_l)$, where

$$\mathcal{M}(\tilde{E}_l) = \prod_{i=1}^{p} m_i^l(x_i) \tag{5.59}$$

and

$$\Sigma(\tilde{E}_l) = \left[[\sigma_1^l(x_1)]^2 \prod_{i;i\neq 1} [m_i^l(x_i)]^2 + \cdots + [\sigma_p^l(x_p)]^2 \prod_{i;i\neq p} [m_i^l(x_i)]^2 \right]^{\frac{1}{2}} \tag{5.60}$$

Then, using Theorem 2.5, we have that $\tilde{Y}_{cos}(\mathbf{x})$ is approximately a Gaussian type-1 set with mean $\mathcal{M}_{cos}(\mathbf{x})$ and standard deviation $\Sigma_{cos}(\mathbf{x})$, where

$$\mathcal{M}_{cos}(\mathbf{x}) = \frac{\sum_{l=1}^{M} \mathcal{M}(\tilde{C}_{\tilde{\tilde{G}}^l}) \mathcal{M}(\tilde{E}_l)}{\sum_{l=1}^{M} \mathcal{M}(\tilde{E}_l)} \tag{5.61}$$

and

$$\Sigma_{cos}(\mathbf{x}) = \frac{\sqrt{\sum_{l=1}^{M} \left[\mathcal{M}^2(\tilde{E}_l) \Sigma^2(\tilde{C}_{\tilde{\tilde{G}}^l}) + \left[\mathcal{M}(\tilde{C}_{\tilde{\tilde{G}}^l}) - \mathcal{M}_{cos}(\mathbf{x})\right]^2 \Sigma^2(\tilde{E}_l) \right]}}{\sum_{l=1}^{M} \mathcal{M}(\tilde{E}_l)} \tag{5.62}$$

The sequence of computations that needs to be performed for center-of-sets type-reduction is as follows : For given values of $x_1, \ldots, x_p$, we determine the mean, $m_i^l(x_i)$, and standard deviation, $\sigma_i^l(x_i)$, of each given antecedent membership grade, $\tilde{\mu}_{\tilde{\tilde{F}}_i^l}(x_i)$, ($l = 1, 2, \ldots, M$ and $i = 1, 2, \ldots, p$). We are also given the mean, $m_G^l(y)$, and standard deviation, $\sigma_G^l(y)$, of the consequent membership grade, $\tilde{\mu}_{\tilde{\tilde{G}}^l}(y)$, for each $y \in Y$. Then,

(a) For each consequent set, compute $\mathcal{M}(\tilde{C}_{\tilde{\tilde{G}}^l})$ and $\Sigma(\tilde{C}_{\tilde{\tilde{G}}^l})$ using (2.79) and (2.80), respectively;

(b) For the degree of firing $[\tilde{E}_l = \sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{\tilde{F}}_i^l}(x_i)]$ corresponding to each consequent set, compute $\mathcal{M}(\tilde{E}_l)$ and $\Sigma(\tilde{E}_l)$ using (5.59) and (5.60), respectively; and,

(c) Calculate the mean and standard deviation of the type-reduced set using (5.61) and (5.62), respectively.

Example 5.8 In this example, we illustrate the use of the just-described type-reduction approximation methods for Gaussian type-2 FLS's. We consider a single input - single output type-2 FLS using product t-norm and product inference, which has rules of the form :

$$R^l : \text{IF } x \text{ is } \tilde{\tilde{F}}^l, \text{ THEN } y \text{ is } \tilde{\tilde{G}}^l.$$

where $x, y \in [0, 10]$.

Figures 5.20 (a) and (b) show the antecedent and consequent sets. Each of these sets is Gaussian type-2 with a Gaussian principal membership function. The principal membership functions of the antecedent sets, $\tilde{\tilde{F}}^1$, $\tilde{\tilde{F}}^2$ and $\tilde{\tilde{F}}^3$, are centered at 2, 5 and 8, and have standard deviations 1.2, 1.1 and 1, respectively. The standard deviations of secondary membership functions are proportional to the means of the secondary membership functions; the constant of proportionality for each of $\tilde{\tilde{F}}^1$, $\tilde{\tilde{F}}^2$ and $\tilde{\tilde{F}}^3$ is 0.2. The principal membership functions of the consequent sets, $\tilde{\tilde{G}}^1$, $\tilde{\tilde{G}}^2$ and $\tilde{\tilde{G}}^3$, are centered at 6, 2 and 9, and have standard deviations 1, 1.2 and 1, respectively. The standard deviations of secondary membership functions are again proportional to the means of the secondary membership functions; the constant of proportionality for each of $\tilde{\tilde{G}}^1$, $\tilde{\tilde{G}}^2$ and $\tilde{\tilde{G}}^3$ is 0.3.

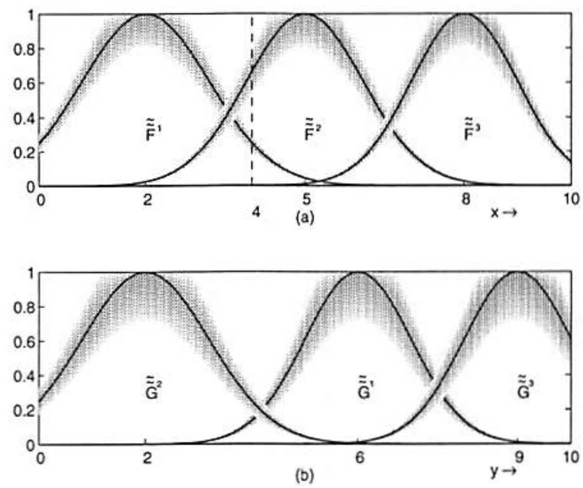


Figure 5.20: (a) Antecedent sets (the vertical axis shows the primary memberships of x in the antecedent sets) and (b) consequent sets (the vertical axis shows the primary memberships of y in the consequent sets) for Example 5.8. The applied input ($x = 4$) is shown in Fig. (a).

The applied input is $x = 4$ [shown in Fig. 5.20 (a)]. It has non-zero memberships in two antecedents $\tilde{\tilde{F}}_1$ and $\tilde{\tilde{F}}_2$. Figures 5.21 and 5.22 depict the output sets and approximate type-reduced sets [obtained using (5.48) and (5.49)] for the centroid and center-of-sums type-reducers, respectively. Figure 5.23 depicts the $\bar{y}^l$'s, their Gaussian-approximated output set memberships [$\tilde{\mu}_{\tilde{\tilde{B}}^l}(\bar{y}^l)$'s - computed using (5.51) and (5.52)] and the approximate type-reduced sets (obtained using the procedure described above in paragraph 3) for the height and modified height type-reducers. For the modified height type-reducer, the δ^l's were set equal to standard deviations of the principal membership functions of the consequent sets : $\delta^1 = 1$, $\delta^2 = 1.2$ and $\delta^3 = 1$. Figure 5.24 depicts the centroids of the consequent sets, their degrees of firing and the approximate type-reduced set (using the procedure described above in paragraph 5) for the center-of-sets type-reducer.

Observe that since the approximation to the type-reduced set in each case is a Gaussian type-1 set, it's centroid is equal to its mean which is also the unity height point in the set. As explained in Section 5.4, this point corresponds to the output of a type-1 FLS which uses the principal membership functions of this Gaussian type-2 FLS. Since all the principal membership functions for the consequent sets in this example are normal, convex and symmetric, the outputs of the height and center-of-sets defuzzifiers for the "principal" type-1 FLS are, as explained in Section 5.3.5, the same; therefore, the means of the height and center-of-sets type-reduced sets are also the same. Their spreads, however, are different, as can be seen from Eqs. (5.54) and (5.62).

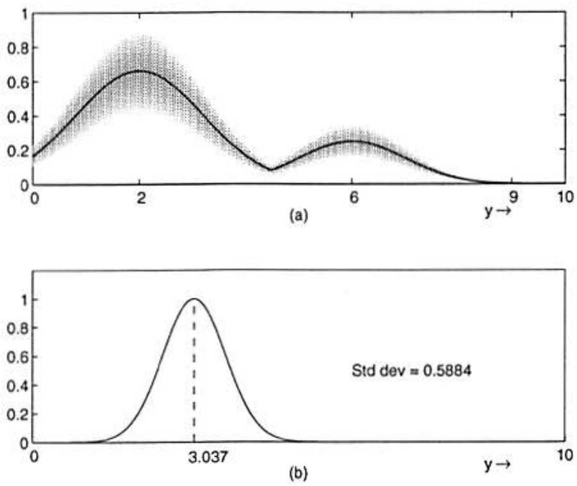


Figure 5.21: (a) The combined output set for the centroid type-reducer (the vertical axis shows the primary memberships of y in the combined output set) and (b) the approximate centroid type-reduced set [obtained using (5.48) and (5.49)] for Example 5.8.

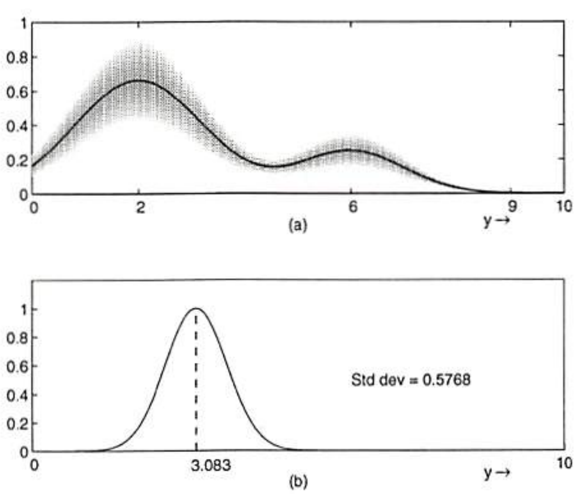


Figure 5.22: (a) The combined output set for the center-of-sums type-reducer (the vertical axis shows the primary memberships of y in the combined output set) and (b) the approximate center-of-sums type-reduced set [obtained using (5.48) and (5.49)] for Example 5.8.

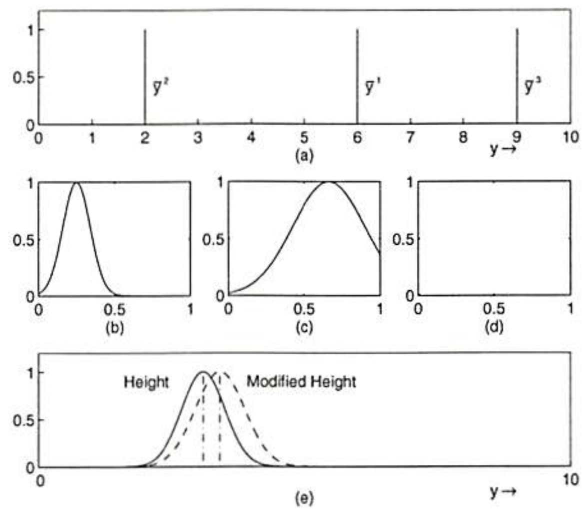


Figure 5.23: (a) The $\bar{y}^l$'s, and their output set memberships : (b) $\tilde{\mu}_{\tilde{\tilde{B}}^1}(\bar{y}^1)$, (c) $\tilde{\mu}_{\tilde{\tilde{B}}^2}(\bar{y}^2)$, and (d) $\tilde{\mu}_{\tilde{\tilde{B}}^3}(\bar{y}^3)$; and, (e) the approximate height (solid line) and modified height (dashed line) type-reduced sets [obtained using (5.53), (5.54), (5.56) and (5.57)] for Example 5.8. Figures (b), (c) and (d) show plots of primary memberships (horizontal axis) versus secondary memberships (vertical axis). Figure (d) is empty, because $\tilde{\mu}_{\tilde{\tilde{B}}^3}(\bar{y}^3)$ is zero (1/0). The means of the type-reduced sets are 3.097 (height) and 3.41 (modified height); and their standard deviations are 0.4056 (height) and 0.4651 (modified height). For the modified height type-reducer, $\delta^1 = 1$, $\delta^2 = 1.2$ and $\delta^3 = 1$.

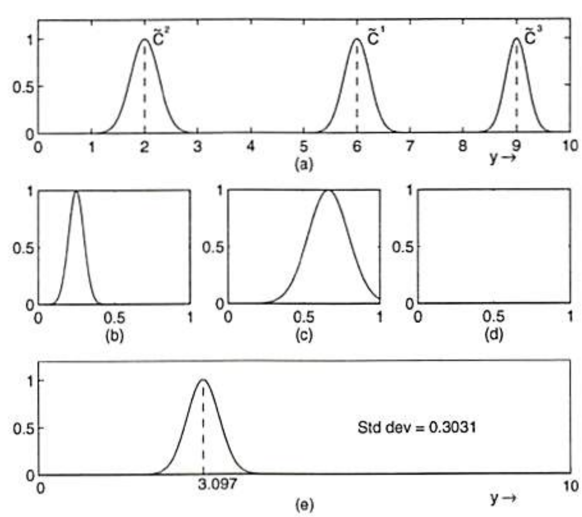


Figure 5.24: The centroids of consequent sets, $\tilde{C}^l$'s, are depicted in Fig. (a) and their respective degrees of firing, $\tilde{\mu}_{\tilde{\tilde{F}}^l}(x)$'s, for $l = 1$, 2 and 3 are depicted in Figs. (b), (c), and (d), respectively. Figure (e) shows the approximate center-of-sets type-reduced set [obtained using (5.61) and (5.62)] for Example 5.8. Figures (b), (c) and (d) show plots of primary memberships (horizontal axis) versus secondary memberships (vertical axis). Figure (d) is empty, because $\tilde{\mu}_{\tilde{\tilde{F}}^3}(x)$ is zero (1/0).

Table 5.2 summarizes the results of this example. The table also shows centroids of the true type-reduced set in each case, found by computing the centroid of the type-2 output set numerically, as discussed in Sections 5.3.1 - 5.3.5. We did, however, make use of the Gaussian *meet* approximation while obtaining the type-2 output set in all the cases [see, for example, (5.51) and (5.52)].

The number of computations required to compute the centroid of a type-2 set having a continuous domain is astronomical; therefore, even the computational results for centroid, center-of-sums and center-of-sets type-reducers are only approximate. Not more than 10 domain points were considered while computing the centroid of any type-2 set with a continuous domain. Also, as mentioned in Example 2.7, it is not easy to predict the nature of the true type-reduced set; therefore, though the mean of the Gaussian approximation for the center-of-sets, centroid and center-of-sums type-reducers appears to be much closer to the true centroid of the type-reduced set, at present *we cannot make any general statements about the approximation error between a particular type-reduced set and its Gaussian approximation.* Appendix C.11 demonstrates how to find bounds on the domain of the type-reduced set. These bounds may be used to bound the difference between the centroids of the true and the approximate type-reduced sets (see the comments towards the end of the Appendix). It may be possible to obtain tighter bounds, or an expression for the approximation error, but we leave this for future research.

Table 5.2: Results of Example 5.8.

Type-reduced set	Mean	Standard deviation	Centroid of the true type-reduced set
Centroid	3.037	0.5884	3.072
Center-of-sums	3.083	0.5768	3.113
Height	3.097	0.4056	3.268
Modified height	3.41	0.4651	3.530
Center-of-sets	3.097	0.3031	3.090

□

5.8 Limitation of the Gaussian Approximations

We wish to emphasize the fact that the results in Section 5.7 are only approximate (see Corollary 2.2 which was used to obtain these results). If the type-2 uncertainty on the antecedent and consequent sets is small, the approximation is close to the true answer and then the centroid of the approximate type-reduced set is close to the centroid of the actual type-reduced set; but, if the type-2 uncertainty on the antecedents and the consequents is not small enough, the approximation may not be very close to the true answer and the true crisp output of the type-2 FLS may differ significantly from the approximate crisp output.

Note also that when we use this approximation, the centroid of the type-reduced set is the same as the unity height point of that set [e.g., see Figs. 5.21 (b) - 5.24 (b)]. As explained in Section 5.4, this point is equal to the output of the type-1 FLS which uses the principal membership functions of the type-2 FLS (we henceforth call this type-1 FLS the "principal" type-1 FLS). *If one is only interested in using the crisp defuzzified output of the type-2 FLS*, using the Gaussian approximation, therefore, does not serve any great purpose. In this case, we recommend actually computing the type-reduced set using the methods described in Section 5.3 and then finding its centroid, though this process can be computationally very very expensive. Doing this will give the centroid of the *true* type-reduced set (see Table 5.2). The approximation introduced in Section 5.7 is useful if one wants to estimate the *spread* of the type-reduced set quickly without going through all the computations. The same can be said about the approximations for triangular and interval type-2 FLS's.

5.9 Interval Type-2 Fuzzy Logic Systems

Comparable results for interval type-2 FLS's, i.e., for FLS's using interval type-2 sets, can be developed by using the results given in Chapter 2 and Appendix D. Even the actual type-reduced set for interval type-2 FLS's can be obtained relatively easily by following the computational procedure described in Appendix D.1.

5.10 Triangular Type-2 Fuzzy Logic Systems

Triangular type-2 FLS's are FLS's which use triangular type-2 sets. Type-reduction results for these FLS's can be developed in a manner very very similar to that in Section 5.7 by using the results in [8].

5.11 Fuzzy Basis Functions in a Type-2 FLS

In a type-1 FLS using height or modified height defuzzification, the defuzzified output can be expressed as a fuzzy basis function (fbf) expansion [15]

$$y(\mathbf{x}) = \sum_{l=1}^{M} \bar{y}^l \phi^l(\mathbf{x}) \tag{5.63}$$

where $\phi^l(\mathbf{x}) = \left[\mathcal{T}_{i=1}^{p}\mu_{\mathrm{F}_i^l}(\mathbf{x})\right] / \left[\sum_{j=1}^{M} \mathcal{T}_{i=1}^{p}\mu_{\mathrm{F}_i^j}(\mathbf{x})\right]$, provided that the consequent sets are <u>normal</u>. A center-of-sets defuzzified output [see (5.21)] can also be expressed in this form with the consequent set centroids, c^l's, replacing $\bar{y}^l$'s, without requiring the consequent sets to be normal.

Observe that these type-1 fbf expansions are weighted sums of the form

$$y(z_1, \ldots, z_M, w_1, \ldots, w_M) = \frac{\sum_{l=1}^{M} w_l z_l}{\sum_{l=1}^{M} w_l} \tag{5.64}$$

where the weights w_l [e.g., $\mathcal{T}_{i=1}^{p}\mu_{\mathrm{F}_i^l}(x_i)$] depend only on the antecedent memberships and the parameters z_l (e.g., $\bar{y}^l$) depend only on the consequents. In this section we attempt to see if there is a parallel to the fbf expansion in the type-2 case.

The extension of (5.64) to the type-2 case allows each w_l to be replaced by a type-1 set $\tilde{\mathrm{W}}_l$ and each z_l by a type-1 set $\tilde{\mathrm{Z}}_l$. According to our convention, this extension is

$$\tilde{\mathrm{Y}}(\tilde{\mathrm{Z}}_1, \ldots, \tilde{\mathrm{Z}}_M, \tilde{\mathrm{W}}_1, \ldots, \tilde{\mathrm{W}}_M) = \int_{z_1} \cdots \int_{z_M} \int_{w_1} \cdots \int_{w_M} \mathcal{T}_{l=1}^{M}\mu_{\tilde{\mathrm{Z}}_l}(z_l) \star \mathcal{T}_{l=1}^{M}\mu_{\tilde{\mathrm{W}}_l}(w_l) \Bigg/ \frac{\sum_{l=1}^{M} w_l z_l}{\sum_{l=1}^{M} w_l} \tag{5.65}$$

where $\mathcal{T}$ and $\star$ both indicate the t-norm used ... product or minimum, $w_l \in \tilde{\mathrm{W}}_l$ and $z_l \in \tilde{\mathrm{Z}}_l$ for $l = 1, \ldots, M$.

In a height [(5.17)] or modified height [(5.20)] type-reducer, the $\tilde{\mathrm{W}}^l$'s are the output set memberships of the $\bar{y}^l$'s, namely the $\tilde{\mu}_{\tilde{\mathrm{B}}^l}(\bar{y}^l)$'s. Since $\tilde{\mu}_{\tilde{\mathrm{B}}^l}(\bar{y}^l)$ depends on both the antecedent and the consequent set memberships of the $\bar{y}^l$'s [see (5.16)], it is not possible to write the term to the right of the slash in (5.65) as a weighted sum in which the weights depend only on the antecedent memberships and the parameters z_l depend only on the consequents.

In the case of a center-of-sets type-reducer [(5.22)], however, each term to the right of the slash in (5.65) can be interpreted as a valid fbf expansion, because the $\tilde{W}^l$'s now correspond to the degrees of firing, $\sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{F}_i^l}(x_i)$'s, which depend solely on the antecedents and the $\tilde{Z}^l$'s are centroids of the consequent sets (which, of course, depend solely on the consequents). So, every point in the domain of a center-of-sets type-reduced set [the term to the right of the slash in (5.65)] can be expressed as a fbf expansion similar to the one in (5.63) [see (5.64) - a restatement of (5.63)]. If we view the type-reduced set as a collection of a large number of type-1 FLS's (see Fig. 5.15), when using center-of-sets type-reduction, the output of each embedded type-1 FLS can be expressed as a fbf expansion. This is another reason to use center-of-sets type-reduction.

When all the type-2 uncertainties collapse to type-1 uncertainties, the type-reduced set collapses to a single point (see Section 5.4), which corresponds to the correct type-1 fbf expansion for the resulting type-1 FLS.

The expression in (5.65), in its most general form, can not be expressed as a weighted sum of type-1 sets. If, however, only the z_l's are fuzzy and all the w_l's are crisp (this corresponds to the case where all the antecedent sets are type-1 in a center-of-sets type-reducer, as in our function approximation example, in Section 5.6), $\tilde{Y}$ can be written as

$$\begin{aligned} \tilde{Y}(\tilde{Z}_1, \ldots, \tilde{Z}_M, w_1, \ldots, w_M) &= \frac{\sum_{l=1}^{M} w_l \tilde{Z}_l}{\sum_{l=1}^{M} w_l} \\ &= \sum_{l=1}^{M} \phi_l \tilde{Z}_l \end{aligned} \tag{5.66}$$

In a center-of-sets type-reducer, $w_l = \prod_{i=1}^{p} \mu_{\tilde{F}_i^l}(x_i)$ and $\tilde{Z}_l$ is the centroid of the lth consequent set, so that the basis functions ϕ_l in (5.66) depend solely on the antecedent memberships and $\tilde{Z}_l$'s depend solely on the consequents.

If the type-reduced set, $\tilde{Y}$, of a type-2 FLS is discretized into N points, $y_1, \ldots, y_N$, the centroid of the type-reduced set is given as

$$C_{\tilde{Y}} = \frac{\sum_{i=1}^{N} y_i \mu_{\tilde{Y}}(y_i)}{\sum_{i=1}^{N} \mu_{\tilde{Y}}(y_i)} \tag{5.67}$$

where each y_i itself can be expressed as a weighted sum [see the term to the right of the slash in (5.65)]. $\tilde{Y}$ can be obtained from any of the type-reduction methods discussed earlier (see Sections 5.3.1 - 5.3.5 for the expressions for specific type-reduced sets).

Note that though (5.67) is a weighted average, it is not a fbf expansion, because each y_i and $\mu_{\tilde{Y}}(y_i)$ depend on both the antecedent memberships and the consequents. When all the type-2 uncertainties collapse to type-1 uncertainties, however, the type-reduced set collapses to a single point, say y', so that $\mu_{\tilde{Y}}(y') = 1$ and $\mu_{\tilde{Y}}(y) = 0$ for $y \neq y'$. In this case, $C_{\tilde{Y}} = y'$ and y' corresponds to the correct fbf expansion of the form (5.64) in the type-1 case for height, modified height and center-of-sets defuzzifiers (assuming that the height and modified height defuzzifiers use normal consequent sets).

Chapter 6

Examples of Type-2 Fuzzy Logic Systems

In this chapter, we describe two examples of type-2 FLS's that will give the reader an idea of the circumstances under which a type-2 FLS can be used and the kind of output one can expect from a type-2 FLS. Applications for type-2 FLS's are by no means limited just to the situations described here and will continue to be a topic of future research.

In Section 6.1, we consider the problem of designing a FLS from rules collected by surveying multiple experts. We show how linguistic uncertainty about membership functions of the FLS, as well as rule uncertainty from multiple experts, each of whom may give different answers to the same question, can be handled in the type-2 framework. In Section 6.2, we demonstrate how information associated with numerical uncertainty in the training data for a type-1 FLS can be interpreted in a type-2 framework, so as to obtain bounds on the type-1 FLS output. We do this for the problem of forecasting the Mackey-Glass chaotic time series [14].

6.1 Collecting Rules by Means of a Survey

In this example, we consider the problem of designing a FLS based on rules collected from multiple experts. Let us suppose that we are designing a FLS to approximate a mapping $f : [0,10] \times [0,10] \to [0,10]$, and that the domain of each input as well as output are divided into three fuzzy sets $\tilde{F}^1$, $\tilde{F}^2$ and $\tilde{F}^3$. To form the rule-base, surveys are collected from several experts. The surveys ask questions of the form :

$$\text{IF } x_1 \text{ is } \tilde{F}^i \text{ and } x_2 \text{ is } \tilde{F}^j\text{, THEN what is } y\text{ ?}$$

where $i, j = 1, 2, 3$. Each question gives one rule, and there are 9 such rules.

Different experts may answer the same question differently. Table 6.1 shows the information collected from (hypothetical) surveys. For each rule, the numbers under $\tilde{F}^i$ $(i = 1, 2, 3)$ show the fraction of experts who answered "y is $\tilde{F}^i$" to that particular question.

Note that, in this example, each of the antecedents as well as the consequent use the same three fuzzy sets; therefore, any linguistic uncertainty in the membership functions of these sets appears in the antecedents as well as the consequents.

We consider two cases : 1) there is no uncertainty associated with the membership functions of the three fuzzy sets (i.e., the three sets are type-1); and 2) the membership functions are uncertain (i.e., the three sets are type-2). In each case, we consider two different approaches to handling the uncertainty introduced in the rules due to the different responses from different experts.

6.1.1 Type-1 Membership Functions

Figure 6.1 (a) shows the membership functions for the three sets $\tilde{F}^1$, $\tilde{F}^2$ and $\tilde{F}^3$, which are Gaussian type-1 sets. Now, we consider two ways of handling the uncertainty due to multiple responses.

Table 6.1: Information collected from (hypothetical) surveys. The integer values for x_1 and x_2 are the indices for sets $\tilde{F}^i$ $(i = 1, 2, 3)$. For each rule, the numbers under $\tilde{F}^i$ $(i = 1, 2, 3)$ show the fraction of experts who answered "y is $\tilde{F}^i$" to that particular question. We use these numbers as the weights w_i^l.

Rule no.	x_1	x_2	y		
			$\tilde{F}^1$	$\tilde{F}^2$	$\tilde{F}^3$
1	1	1	0.8	0.2	0
2	1	2	0.25	0.7	0.05
3	1	3	0.05	0.25	0.7
4	2	1	0.3	0.6	0.1
5	2	2	0.15	0.8	0.05
6	2	3	0	0.25	0.75
7	3	1	0.2	0.5	0.3
8	3	2	0	0.1	0.9
9	3	3	0	0	1

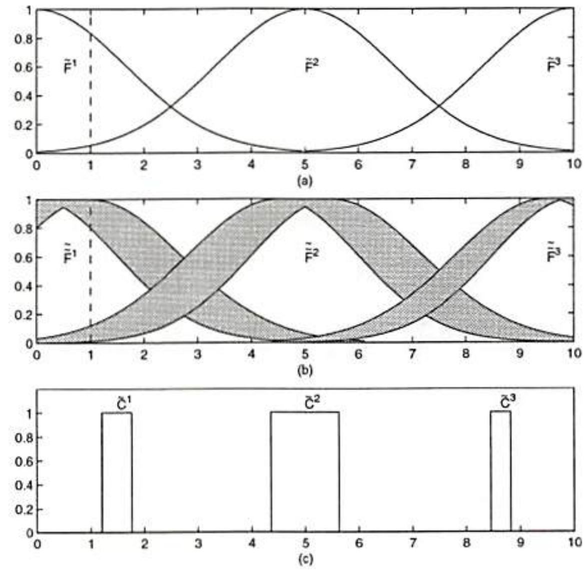


Figure 6.1: Membership functions for the survey example : (a) type-1 membership functions - case 1; (b) type-2 membership functions - case 2; (c) centroids of the type-2 membership functions in Fig. (b). The centroids of $\tilde{F}^1$, $\tilde{F}^2$ and $\tilde{F}^3$ in (a) are $c^1 = 1.3213$, $c^2 = 5$ and $c^3 = 8.6787$, respectively.

6.1.1.1 First Approach : Keeping One Response

In our first approach to forming a rule base, we choose a single consequent for each rule. We accomplish this in two different ways : (1) by keeping the response with the largest weight; and, (2) by averaging the centroids of all the responses for each rule and using this average in place of the rule consequent centroid.

Method 1 : Keeping the Response with the Largest Weight

In this approach to forming a rule-base, we choose the set for each rule that has the largest number of experts in favor of it, in which case the rule-base is as shown in Table 6.2 and the FLS is type-1. For this system, we use center-of-sets defuzzification. A plot of y_{cos} versus (x_1, x_2) is shown in Fig. 6.2.

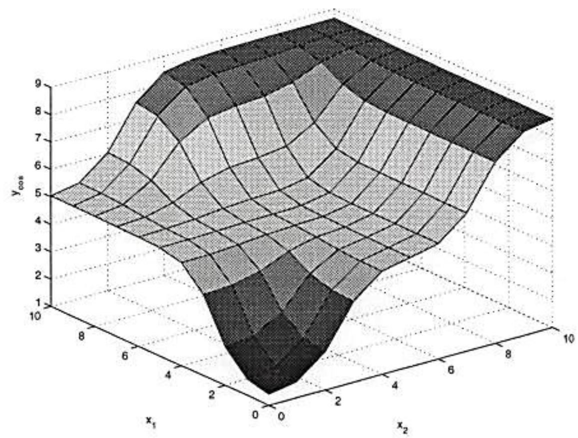


Figure 6.2: The plot of y_{cos} versus (x_1, x_2) for the type-1 FLS using rules from Table 6.2.

So that the reader can understand how we obtained these results, we next show the calculation of the output corresponding to $(x_1, x_2) = (1, 3)$. The memberships of $x_1 = 1$ in $\tilde{\mathrm{F}}^1$, $\tilde{\mathrm{F}}^2$ and $\tilde{\mathrm{F}}^3$ are [see the dashed line in Fig. 6.1 (a)] 0.8341, 0.0548 and 0.00, respectively; and those of $x_2 = 3$ are 0.1953, 0.4839 and 0.00, respectively. For these values of x_1 and x_2, four rules (rule numbers 1, 2, 4 and 5, with consequents $\tilde{\mathrm{F}}^1$, $\tilde{\mathrm{F}}^2$, $\tilde{\mathrm{F}}^2$, and $\tilde{\mathrm{F}}^2$, respectively) are fired; and the corresponding degrees of firing are 0.1629 (e.g., 0.8341×0.1953), 0.4036, 0.0107 and 0.0265, respectively. The centroids of the fired consequent sets, $\tilde{\mathrm{F}}^1$ and $\tilde{\mathrm{F}}^2$, are 1.3213 and 5, respectively [see Fig. 6.1 (a)]; and the output is [see (5.21)]

$$\begin{aligned} y_{cos}(1,3) &= \frac{0.1629 \times 1.3213 + (0.4036 + 0.0107 + 0.0265) \times 5}{0.1629 + 0.4036 + 0.0107 + 0.0265} \\ &= 4.0084 \end{aligned} \tag{6.1}$$

Method 2 : Averaging the Responses

In this approach to forming a rule-base, we find a weighted average of the rule consequents for each rule, where the weights used are obtained from the responses of the experts (see Table 6.1); and, then we use this average in place of the rule consequent centroid, i.e., we treat the FLS as a type-1 TSK FLS (see Section 5.3.7), having rules of the form

$$R^l : \text{IF } x_1 \text{ is } \tilde{\mathrm{F}}^i \text{ and } x_2 \text{ is } \tilde{\mathrm{F}}^j, \text{ THEN } y = c^l_{avg}$$

where

$$c^l_{avg} = \frac{w^l_1 c^1 + w^l_2 c^2 + w^l_3 c^3}{w^l_1 + w^l_2 + w^l_3} \tag{6.2}$$

Table 6.2: Keeping the response with the largest weight

Rule no.	x_1	x_2	y
1	1	1	$\tilde{F}^1$
2	1	2	$\tilde{F}^2$
3	1	3	$\tilde{F}^3$
4	2	1	$\tilde{F}^2$
5	2	2	$\tilde{F}^2$
6	2	3	$\tilde{F}^3$
7	3	1	$\tilde{F}^2$
8	3	2	$\tilde{F}^3$
9	3	3	$\tilde{F}^3$

in which w_i^l is the weight associated with the ith consequent for the lth rule ($i = 1, 2, 3; l = 1, \ldots, 9$); and, c^i is the centroid of the ith consequent set ($i = 1, 2, 3$). The centroids of the three sets are $c^1 = 1.3213$, $c^2 = 5$ and $c^3 = 8.6787$ (see Fig. 6.1). The complete rule-base is shown in Table 6.3. A plot of y_{tsk} versus (x_1, x_2) is shown in Fig. 6.3.

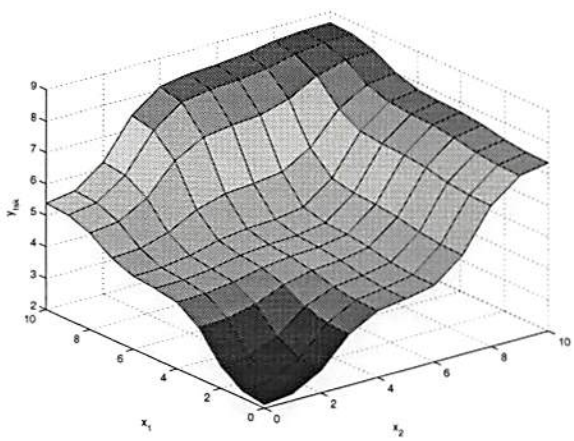


Figure 6.3: The plot of y_{tsk} versus (x_1, x_2) for the type-1 FLS using rules from Table 6.3.

We show a sample calculation for the output corresponding to $(x_1, x_2) = (1, 3)$. The memberships of x_1 and x_2 as well as the degrees of firing of the four rules (rule numbers 1, 2, 4 and 5) are the same as computed in Method 1. In this case, however, the consequents of all the four rules are different (see Table 6.3). The output, therefore, is [see (5.34)]

$$\begin{aligned} y_{tsk}(1,3) &= \frac{0.1629 \times 2.0570 + 0.4036 \times 4.2643 + 0.0107 \times 4.2643 + 0.0265 \times 4.6321}{0.1629 + 0.4036 + 0.0107 + 0.0265} \\ &= 3.6856 \end{aligned} \tag{6.3}$$

6.1.1.2 Second Approach : Preserving All the Responses

In our second approach to forming a rule base, we seek to preserve the distribution of the expert-responses for each rule. Observe, from Table 6.1, that each possible response to every question can

Table 6.3: Averaging the responses, when the membership functions are type-1. For each rule, c^l_{avg}'s are weighted averages of consequent centroids, c^1, c^2 and c^3, where the weights are obtained from Table 6.1.

Rule no. (l)	x_1	x_2	c^l_{avg}
1	1	1	2.0570
2	1	2	4.2643
3	1	3	7.3911
4	2	1	4.2643
5	2	2	4.6321
6	2	3	7.7590
7	3	1	5.3679
8	3	2	8.3108
9	3	3	8.6787

be considered to form one rule of a FLS; and, since one rule in a FLS can have only one consequent, different responses to the same question can be considered as rules from different FLS's. So, all the expert-responses taken together can be viewed as a *collection of many different type-1 FLS's*, each corresponding to one combination of expert-responses. For the results shown in Table 6.1, there are 3^9 possible type-1 FLS's. Each possible consequent for every rule can be assigned a weight that is equal to the percentage of experts who voted in favor of it; hence, each rule in every one of the 3^9 FLS's can be assigned the weight of its consequent; and, each FLS can be assigned a weight equal to the t-norm of the weights of its 9 rules (in this example, we will use only product t-norm). So, one survey can be represented as a collection of many different type-1 FLS's, each having this weight associated with it.

Next, we demonstrate that *for every applied input, the output of this system can be represented as a fuzzy set, whose elements are the outputs of all these different type-1 FLS's, and whose memberships are the weights associated with the different FLS's.* In our example, each rule can have three possible consequents, corresponding to three different possible responses to a question ($\tilde{F}^1$, $\tilde{F}^2$ or $\tilde{F}^3$); hence, we shall consider every possible combination of rules, in which any of the rules can have any of the three consequents. Let $w^l_{i_l}$ be the weight and $c^l_{i_l}$ be the consequent centroid associated with the lth rule, if its consequent is $\tilde{F}^{i_l}$ ($l = 1, \ldots, 9$ and $i_l = 1, 2, 3$). Assume that the weights are normalized, so that the maximum weight for each rule is 1, i.e., $\max_{i_l} w^l_{i_l} = 1$. Also, let $\phi^1(\mathbf{x}), \phi^2(\mathbf{x}), \ldots, \phi^9(\mathbf{x})$ be the fuzzy basis functions associated with the 9 rules for some applied input $\mathbf{x}$. Though each rule has different consequents in different FLS's, it has the same antecedents; therefore, each of the 3^9 type-1 FLS's possesses the same basis functions, $\phi^1(\mathbf{x}), \phi^2(\mathbf{x}), \ldots, \phi^9(\mathbf{x})$. Assuming that each of these type-1 FLS's uses center-of-sets defuzzification [see (5.21)], the fuzzy set consisting of the outputs of all 3^9 different FLS's (as discussed in the previous paragraph) can be represented as

$$\tilde{Y}(\mathbf{x}) = \sum_{i_1=1}^{3} \sum_{i_2=1}^{3} \cdots \sum_{i_9=1}^{3} \left[\mathcal{T}_{l=1}^{9} w^l_{i_l} \Big/ \sum_{l=1}^{9} \phi^l(\mathbf{x}) c^l_{i_l} \right] \tag{6.4}$$

where the summations outside the square bracket indicate logical union and the one inside indicates algebraic sum. $\mathcal{T}$ indicates the chosen t-norm. In this example, we will use only the product t-norm.

Next, we show that (6.4) can also be represented as

$$\tilde{Y}(\mathbf{x}) = \sum_{l=1}^{9} \phi^l(\mathbf{x}) \left[w^l_1/c^l_1 + w^l_2/c^l_2 + w^l_3/c^l_3 \right] \tag{6.5}$$

where the summation outside the square brackets indicates algebraic sum and those inside indicate logical union. To do this, observe that Eq. (6.5) can be rewritten as [see (2.63); remember, also, that $\phi^l(\mathbf{x})$ can be represented as $1/\phi^l(\mathbf{x})$]

$$\begin{aligned}\tilde{\mathrm{Y}}(\mathbf{x}) &= \sum_{l=1}^{9}\left[w_1^l \Big/ \phi^l(\mathbf{x})c_1^l + w_2^l \Big/ \phi^l(\mathbf{x})c_2^l + w_3^l \Big/ \phi^l(\mathbf{x})c_3^l\right] \\ &= \sum_{l=1}^{9}\left[\sum_{i_l=1}^{3} w_{i_l}^l \Big/ \phi^l(\mathbf{x})c_{i_l}^l\right] \\ &= \sum_{i_1=1}^{3} w_{i_1}^1 \Big/ \phi^1(\mathbf{x})c_{i_1}^1 + \sum_{i_2=1}^{3} w_{i_2}^2 \Big/ \phi^2(\mathbf{x})c_{i_2}^2 + \cdots + \sum_{i_9=1}^{3} w_{i_9}^9 \Big/ \phi^9(\mathbf{x})c_{i_9}^9 \end{aligned} \quad (6.6)$$

where all the summations indicate logical union and the "+" signs indicate algebraic sum. Equation (6.6) indicates that $\tilde{\mathrm{Y}}(\mathbf{x})$ is the algebraic sum of 9 type-1 fuzzy sets, each of which has 3 terms in it. Let

$$\tilde{\mathrm{A}}_l(\mathbf{x}) = \sum_{i_l=1}^{3} w_{i_l}^l \Big/ \phi^l(\mathbf{x})c_{i_l}^l \quad ; \quad l = 1, 2, \ldots, 9 \quad (6.7)$$

where the summation denotes logical union. Using (2.64), we can add $\tilde{\mathrm{A}}_1(\mathbf{x})$ and $\tilde{\mathrm{A}}_2(\mathbf{x})$ algebraically as follows :

$$\begin{aligned}\tilde{\mathrm{A}}_1(\mathbf{x}) + \tilde{\mathrm{A}}_2(\mathbf{x}) &= \sum_{i_1=1}^{3} w_{i_1}^1 \Big/ \phi^1(\mathbf{x})c_{i_1}^1 + \sum_{i_2=1}^{3} w_{i_2}^2 \Big/ \phi^2(\mathbf{x})c_{i_2}^2 \\ &= \sum_{i_1=1}^{3}\sum_{i_2=1}^{3} (w_{i_1}^1 \star w_{i_2}^2) \Big/ \left[\phi^1(\mathbf{x})c_{i_1}^1 + \phi^2(\mathbf{x})c_{i_2}^2\right] \\ &= \sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\left[\mathcal{T}_{l=1}^{2} w_{i_l}^l \Big/ \sum_{l=1}^{2}\phi^l(\mathbf{x})c_{i_l}^l\right] \end{aligned} \quad (6.8)$$

Adding $\tilde{\mathrm{A}}_3(\mathbf{x})$ to (6.8), we have [using (2.64) again]

$$\begin{aligned}\tilde{\mathrm{A}}_1(\mathbf{x}) + \tilde{\mathrm{A}}_2(\mathbf{x}) + \tilde{\mathrm{A}}_3(\mathbf{x}) &= \sum_{i_1=1}^{3}\sum_{i_2=1}^{3} (w_{i_1}^1 \star w_{i_2}^2) \Big/ \left[\phi^1(\mathbf{x})c_{i_1}^1 + \phi^2(\mathbf{x})c_{i_2}^2\right] + \sum_{i_3=1}^{3} w_{i_3}^3 \Big/ \phi^3(\mathbf{x})c_{i_3}^3 \\ &= \sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\sum_{i_3=1}^{3} (w_{i_1}^1 \star w_{i_2}^2 \star w_{i_3}^3) \Big/ \left[\phi^1(\mathbf{x})c_{i_1}^1 + \phi^2(\mathbf{x})c_{i_2}^2 + \phi^3(\mathbf{x})c_{i_3}^3\right] \\ &= \sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\sum_{i_3=1}^{3}\left[\mathcal{T}_{l=1}^{3} w_{i_l}^l \Big/ \sum_{l=1}^{3}\phi^l(\mathbf{x})c_{i_l}^l\right] \end{aligned} \quad (6.9)$$

Continuing in this fashion, it is easy to see that the RHS of (6.6) is equal to (6.4).

Next, we show that $\tilde{\mathrm{Y}}(\mathbf{x})$ in (6.5) can be interpreted as the type-reduced set of a certain type-2 FLS when center-of-sets type-reduction is used. Observe, from (5.43), that when only the consequent sets in a FLS are type-2, the center-of-sets type-reduced set can be expressed as

$$\tilde{\mathrm{Y}}_{cos}(\mathbf{x}) = \sum_{l=1}^{M}\phi^l(\mathbf{x})\tilde{\mathrm{C}}_l \quad (6.10)$$

where $\tilde{C}_l$ is the centroid of the lth consequent set and $\phi^l(x) = [\mathcal{T}_{i=1}^p \mu_{\tilde{F}^l}(x)] \ / \ [\sum_{j=1}^M \mathcal{T}_{i=1}^p \mu_{\tilde{F}^j}(x)]$. Comparing (6.10) and (6.5), *we see that (6.5) is the same as the expression for a center-of-sets type-reduced set of a 9 rule type-2 FLS, when all antecedents are type-1 and the consequent centroid for the lth rule is equal to* $(w_1^l/c_1^l + w_2^l/c_2^l + w_3^l/c_3^l)$.

In Section 5.4 we saw that the type-reduced set of a type-2 FLS represents a collection of all the outputs of all the type-1 FLS's embedded in that type-2 FLS. In this example, we followed the reverse path, i.e., we started with a collection of many different type-1 FLS's and showed that they can be represented as embedded FLS's in a single type-2 FLS.

The number of these possible embedded type-1 FLS's can be computed as follows : each rule can have three different consequents, corresponding to three different possible responses to a question ($\tilde{F}^1$, $\tilde{F}^2$ or $\tilde{F}^3$), and, there are 9 such rules; so, the total number of possible combinations is $3 \times 3 \times \cdots$ (9 times) $= 3^9$! [Remember that (6.5) is another way to represent (6.4), and the 3^9 terms are clearly present in (6.4).] This is a large number; however, actual computations are simplified by the fact that every input fires only a small number of rules (i.e., the degree of firing for every rule, computed by multiplying the antecedent memberships of x, is non-zero for only a small number of rules), and every consequent centroid consists of only 3 points, i.e., $(w_1^l/c_1^l + w_2^l/c_2^l + w_3^l/c_3^l)$.

Note that our type-2 representation does not depend on the number of inputs or on the number of antecedent sets. All that is required is that the antecedents be type-1 and that every FLS in the collection have the same number of rules and exactly the same antecedent structure for the rules. Note also that, since there exist no physical consequent sets with the centroids of the form $(w_1^l/c_1^l + w_2^l/c_2^l + w_3^l/c_3^l)$, it is better to think of this type-2 FLS as a type-2 TSK FLS which uses type-1 fuzzy numbers of the form $(w_1^l/c_1^l + w_2^l/c_2^l + w_3^l/c_3^l)$ in the consequent for the lth rule (see Section 5.3.7 for a comparison between a Mamdani FLS using center-of-sets type-reduction and a type-2 TSK FLS).

In Theorem 6.1 given next, *equivalence* between the collection of type-1 FLS's and a type-2 TSK FLS means that, for any given input, the elements of the type-reduced set of the type-2 TSK FLS are the outputs of all the possible type-1 FLS's, and their memberships are the weights associated with each of the type-1 FLS's. These weights are computed by multiplying the weights of individual rule consequents in each FLS. If more than one type-1 FLS gives the same output, say y', the membership for y' is taken as the maximum of the weights of all the type-1 FLS's giving output equal to y'.

Theorem 6.1 *Suppose that M rules for a FLS are collected by surveying multiple experts. If the survey has M questions and each question has N possible answers (corresponding to N possible consequents* $\tilde{G}_1^l, \tilde{G}_2^l, \ldots, \tilde{G}_N^l$*, for the lth rule), the data collected from the survey, in general, represents a collection of* N^M *possible type-1 FLS's. Assume further that weights are assigned to the responses from experts, such that* $w_i^l \in [0,1]$ *is the weight associated with the consequent* $\tilde{G}_i^l$*. Then representing the survey as a collection of these* N^M *type-1 FLS's using center-of-sets defuzzification, is equivalent to modeling the survey as a type-2 TSK FLS that has the same number of rules and the same antecedent structure as the survey, but whose lth consequent is a type-1 fuzzy set equal to* $\sum_{i=1}^N w_i^l/c_i^l$*, where the sum indicates logical union, and* c_i^l *is the centroid of* $\tilde{G}_i^l$.

Proof : The proof uses the same argument presented above when we showed the equivalence between the type-1 and the type-2 systems [see (6.4) - (6.10)]. □

Comment : Though normalization of the weights (i.e., $\max_i w_i^l = 1$ for $l = 1, \ldots, M$) is not required to achieve equivalence between the collection of type-1 FLS's and the type-2 TSK FLS described in Theorem 6.1, it is highly recommended for the reasons explained next.

The expression for the type-reduced set of the type-2 TSK FLS can be written as [see (6.5)]

$$\tilde{Y}_{TSK}(\mathbf{x}) = \sum_{l=1}^M \phi^l(\mathbf{x}) \Big(\sum_{i=1}^N w_i^l/c_i^l \Big) \qquad (6.11)$$

where the summation outside the large parentheses indicates algebraic sum and the one inside indicates logical union. To show the contribution of the kth rule, (6.11) can be rewritten as

$$\begin{aligned} \tilde{\mathrm{Y}}_{TSK}(\mathbf{x}) &= \sum_{\substack{l=1 \\ l\neq k}}^{M} \phi^l(\mathbf{x})\Big(\sum_{i=1}^{N} w_i^l/c_i^l\Big) + \phi^k(\mathbf{x})\Big(\sum_{i=1}^{N} w_i^l/c_i^l\Big) \\ &= \sum_{\substack{l=1 \\ l\neq k}}^{M} \phi^l(\mathbf{x})\Big(\sum_{i=1}^{N} w_i^l/c_i^l\Big) + \Big[\sum_{i=1}^{N} w_i^l \Big/ \phi^k(\mathbf{x})c_i^l\Big] \end{aligned} \quad (6.12)$$

Now suppose that, for a particular input $\mathbf{x}'$, $\phi^k(\mathbf{x}') = 0$, i. e., the kth rule has not fired. The contribution of the kth term in (6.11), then, is equal to zero, and one would expect the type-reduced set to be equal to [see (6.12)]

$$\tilde{\mathrm{Y}}_{TSK}(\mathbf{x}') = \sum_{\substack{l=1 \\ l\neq k}}^{M} \phi^l(\mathbf{x})\Big(\sum_{i=1}^{N} w_i^l/c_i^l\Big) \quad (6.13)$$

Equation (6.12), however, gives us (assuming maximum t-conorm)

$$\begin{aligned} \tilde{\mathrm{Y}}_{TSK}(\mathbf{x}') &= \sum_{\substack{l=1 \\ l\neq k}}^{M} \phi^l(\mathbf{x}')\Big(\sum_{i=1}^{N} w_i^l/c_i^l\Big) + \Big[\sum_{i=1}^{N} w_i^l \Big/ 0\Big] \\ &= \sum_{\substack{l=1 \\ l\neq k}}^{M} \phi^l(\mathbf{x})\Big(\sum_{i=1}^{N} w_i^l/c_i^l\Big) + \Big(\max_i w_i^l\Big)\Big/0 \end{aligned} \quad (6.14)$$

Note that if $\max_i w_i^l \neq 1$, (6.13) and (6.14) are, in general, not equal. This can be explained as follows.

Consider a type-1 set $\tilde{\mathrm{F}} = \int_v f(v)/v$. Adding $0 = 1/0$ to this set gives us

$$\tilde{\mathrm{F}} + 1/0 = \int_v f(v)/v + 1/0 = \int_v [f(v) \star 1]/(v+0) = \int_v f(v)/v = \tilde{\mathrm{F}}\,; \quad (6.15)$$

whereas, adding $w/0$, for some $w \in (0,1)$, gives us

$$\tilde{\mathrm{F}} + w/0 = \int_v f(v)/v + w/0 = \int_v [f(v) \star w]/(v+0) = \int_v [f(v) \star w]/v\,, \quad (6.16)$$

which shows that, in general, $\tilde{\mathrm{F}} + 1/0 \neq \tilde{\mathrm{F}} + w/0$.

The above analysis shows that if the weights are not normalized, we cannot compute the type-reduced set, $\tilde{\mathrm{Y}}_{TSK}(\mathbf{x}')$, using (6.13), and, instead we have to use (6.14) to include the effect of the kth rule, even though it did not fire. This certainly does not seem intuitive; but, the way to avoid this problem is to normalize all the weights, so that $\max_i w_i^l = 1$ for $l = 1, \ldots, M$. A similar argument can be presented for the case when more than one $\phi^l(\mathbf{x})$ are zero. □

Using Theorem 6.1, we can now treat the FLS obtained from the survey as a type-2 TSK FLS, whose rules are of the form (see Section 5.3.7)

$$R^l : \text{IF } x_1 \text{ is } \tilde{\mathrm{F}}^i \text{ and } x_2 \text{ is } \tilde{\mathrm{F}}^j, \text{ then } y = \tilde{\mathrm{K}}^l.$$

where elements of the fuzzy set $\tilde{K}^l$ consist of the centroids of the three consequent sets $\tilde{F}^1$, $\tilde{F}^2$ and $\tilde{F}^3$ (we now denote the centroids as c^1, c^2 and c^3 for convenience); and their memberships are assigned in proportion to the number of experts who voted in favor of them. We normalize the memberships for each rule by dividing by the maximum weight, e.g., the three memberships for the first rule are found as $0.8/0.8 = 1$, $0.2/0.8 = 0.25$ and $0/0.8 = 0$, respectively. The rule-base, in this case, is shown in Table 6.4. For each rule, due to normalization, the number having unity membership in the consequent is the same as the centroid of the consequent set of the corresponding rule in Table 6.2.

Table 6.4: Preserving all the responses. c^i $(i = 1, 2, 3)$ indicates the centroid of $\tilde{F}^i$ - see Fig. 6.1 (a). The memberships for each $\tilde{K}^l$ are normalized by dividing by the largest weight (see Table 6.1).

Rule no. (l)	x_1	x_2	$\tilde{K}^l$
1	1	1	$1/c^1 + 0.25/c^2$
2	1	2	$0.3571/c^1 + 1/c^2 + 0.0714/c^3$
3	1	3	$0.0714/c^1 + 0.3571/c^2 + 1/c^3$
4	2	1	$0.5/c^1 + 1/c^2 + 0.1667/c^3$
5	2	2	$0.1875/c^1 + 1/c^2 + 0.0625/c^3$
6	2	3	$0.3333/c^2 + 1/c^3$
7	3	1	$0.4/c^1 + 1/c^2 + 0.6/c^3$
8	3	2	$0.1111/c^2 + 1/c^3$
9	3	3	$1/c^3$

The output of the system is computed, as [see (6.10) and also Section 5.3.7]

$$\tilde{Y}(\mathbf{x}) = \sum_{l=1}^{9} \frac{\mu_{\tilde{F}_1^l}(x_1)\mu_{\tilde{F}_2^l}(x_2)}{\sum_{j=1}^{9} \mu_{\tilde{F}_1^j}(x_1)\mu_{\tilde{F}_2^j}(x_2)} \tilde{K}^l \tag{6.17}$$

where the summations denote algebraic sum.

It is difficult to depict the type-reduced set of this FLS as a plot of $\tilde{Y}$ versus (x_1, x_2), because there is more than one possible output for each input. We show the outputs (i.e., the type-reduced sets) for 5 different inputs : $(0, 0)$, $(1, 3)$, $(5, 5)$, $(7, 7.5)$ and $(10, 10)$ in Fig. 6.4.

Observe that each $\tilde{K}^l$ in Table 6.4 has only one point having unity membership; and, since the memberships of the $\tilde{K}^l$'s are normalized by dividing the weights in Table 6.1 by the largest weight for each rule, the unity membership point of each $\tilde{K}^l$ $(l = 1, \ldots, M)$ corresponds to the centroid of the consequent set having the largest weight for each rule (compare Tables 6.2 and 6.4). The type-reduced set computed using (6.17), therefore, also has only one point having unity membership (see Section 2.4 for more discussion about algebraic operations on type-1 sets), and this unity membership point is the same as $y_{cos}(\mathbf{x})$, where $y_{cos}(\mathbf{x})$ is the output of the type-1 FLS in the first approach - Method 1.

In order to explain how the results in Fig. 6.4 were obtained, we show the calculation of the type-reduced set for $(x_1, x_2) = (1, 3)$. The antecedent memberships and degrees of firing are the same as those described at the end of Section 6.1.1.1. The 4 consequents, however, are now type-1 fuzzy numbers $\tilde{K}^1 = 1/c^1 + 0.25/c^2$, $\tilde{K}^2 = 0.3571/c^1 + 1/c^2 + 0.0714/c^3$, $\tilde{K}^4 = 0.5/c^1 + 1/c^2 + 0.1667/c^3$ and $\tilde{K}^5 = 0.1875/c^1 + 1/c^2 + 0.0625/c^3$, where $c^1 = 1.3213$, $c^2 = 5$ and $c^3 = 8.6787$. The type-reduced set is obtained as [see (6.17) and also Section 2.4]

$$\tilde{Y}_{TSK}(1,3) = \frac{0.1629 \times \tilde{K}^1 + 0.4036 \times \tilde{K}^2 + 0.0107 \times \tilde{K}^4 + 0.0265 \times \tilde{K}^5}{0.1629 + 0.4036 + 0.0107 + 0.0265}$$

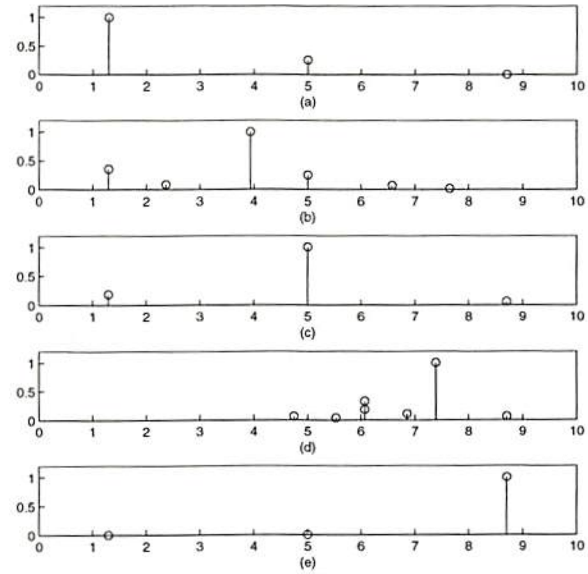


Figure 6.4: Five outputs of the type-2 TSK FLS using rules from Table 6.4 (when the membership functions are all assumed type-1) for the inputs : (a) $(0,0)$, (b) $(1,3)$, (c) $(5,5)$, (d) $(7,7.5)$, and (e) $(10,10)$. In Fig. (d), two domain-points of the type-reduced set are very close to each other (near $y=6$); therefore, it appears as if the two circles (indicating membership grades in the type-reduced set) are at the same domain-point.

$$\begin{aligned}
&= 0.2698 \times \tilde{\mathrm{K}}^1 + 0.6685 \times \tilde{\mathrm{K}}^2 + 0.0177 \times \tilde{\mathrm{K}}^4 + 0.0439 \times \tilde{\mathrm{K}}^5 \\
&= 0.2698 \times [1/c^1 + 0.25/c^2] + 0.6685 \times [0.3571/c^1 + 1/c^2 + 0.0714/c^3] \\
&\quad +0.0177 \times [0.5/c^1 + 1/c^2 + 0.1667/c^3] \\
&\quad +0.0439 \times [0.1875/c^1 + 1/c^2 + 0.0625/c^3] \\
&= \left[1/(0.2698 \times c^1) + 0.25/(0.2698 \times c^2)\right] \\
&\quad + \left[0.3571/(0.6685 \times c^1) + 1/(0.6685 \times c^2) + 0.0714/(0.6685 \times c^3)\right] \\
&\quad + \left[0.5/(0.0177 \times c^1) + 1/(0.0177 \times c^2) + 0.1667/(0.0177 \times c^3)\right] \\
&\quad + \left[0.1875/(0.0439 \times c^1) + 1/(0.0439 \times c^2) + 0.0625/(0.0439 \times c^3)\right] \\
&= [1/0.3565 + 0.25/1.349] \\
&\quad +[0.3571/0.8833 + 1/3.3425 + 0.0714/5.802] \\
&\quad +[0.5/0.0234 + 1/0.0885 + 0.1667/0.1536] \\
&\quad +[0.1875/0.0580 + 1/0.2195 + 0.0625/0.3810] \qquad (6.18)
\end{aligned}$$

where summations inside the square brackets indicate logical union and those outside indicate algebraic sum. The RHS of (6.18) is an algebraic sum of 4 type-1 sets. To simplify it, we need to use (2.64). The resulting type-1 set has $2 \times 3 \times 3 \times 3 = 54$ terms. If more than one of these 54 terms give the same domain point [the quantity to the right of the slash in (6.18)] in the resulting type-1 set, we choose the maximum membership grade [the quantity to the left of the slash in (6.18)] for that domain point. We performed this addition on a computer. Neglecting terms with very small memberships, the result obtained is

$$\begin{aligned}
\tilde{\mathrm{Y}}_{TSK}(1,3) &\approx 0.3571/1.3213 + 0.0893/2.3792 + 1/3.9421 + 0.25/5 \\
&\quad +0.0714/6.5630 + 0.0179/7.6208 \qquad (6.19)
\end{aligned}$$

The output is depicted in Fig. 6.4 (b). The unity membership point does not appear to be exactly equal to the output of the type-1 FLS in Section 6.1.1.1 - Method 1, because the computer program for the present case ignores all the degrees of firing less than 0.05, in order to speed up computations.

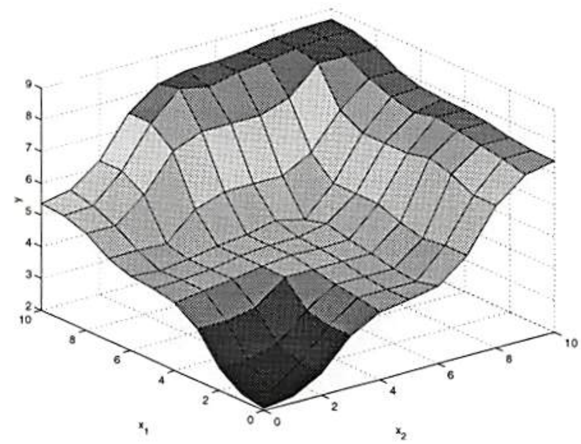


Figure 6.5: The plot of the crisp output of the type-2 FLS versus (x_1, x_2) using rules from Table 6.4.

Figure 6.5 shows a plot of the crisp output of this type-2 FLS versus (x_1, x_2). The crisp output is obtained by finding the centroid of the type-reduced set in each case (see Section 5.5 for more discussion on defuzzification).

6.1.1.3 A Comparison of the Two Approaches

The first approach (Section 6.1.1.1) only showed the output of the type-1 FLS associated with either the largest number of experts (Method 1) or the averaged responses (Method 2); whereas, the second approach (Section 6.1.1.2) showed the outputs of all the possible type-1 FLS's that could be designed from the given data (along with the appropriate weights of the FLS's). The crisp output in the second approach is the average of the outputs of all the type-1 FLS's embedded in the type-2 TSK FLS (see Section 5.4).

Note that the second approach first computes the outputs of all the embedded FLS's and then computes their average, whereas Method 2 in the first approach first averages the responses and then computes the output of the resulting type-1 FLS.

Figure 6.6 shows the difference between the crisp output in the second approach and the crisp output in the first approach when only the response with the largest weight is preserved. The MSE between the two crisp outputs is 0.3544 [the MSE is found by squaring all the ϵ values, i.e., values on the independent axis, in Fig. 6.6, summing them, and dividing the sum by the total number of (x_1, x_2) pairs]. The difference between these two crisp outputs is larger when there is a substantial disagreement among the experts, i.e., when the histogram of responses is more spread out. Observe, for example, from Table 6.1, that the histogram of responses for rule number 3 (when x_1 is close to zero and x_2 is close to 10) is more spread out than the histogram of responses for rule number 8 (when x_1 is close to 10 and x_2 is close to 5); and correspondingly, in Fig. 6.6, the error near $(x_1, x_2) = (0, 10)$ is much larger than that near $(x_1, x_2) = (10, 5)$. Figure 6.7 shows the difference between the crisp output in the second approach and the crisp output in the first approach when the responses are averaged. The MSE in this case is 0.0159. The differences in this case are much smaller than those in Fig. 6.6.

If we just consider the crisp outputs, the second method in the first approach (where the responses are averaged) is in some sense closer to the second approach. This observation suggests that *if one's aim is to obtain only the crisp outputs from all the collected surveys, the second method*

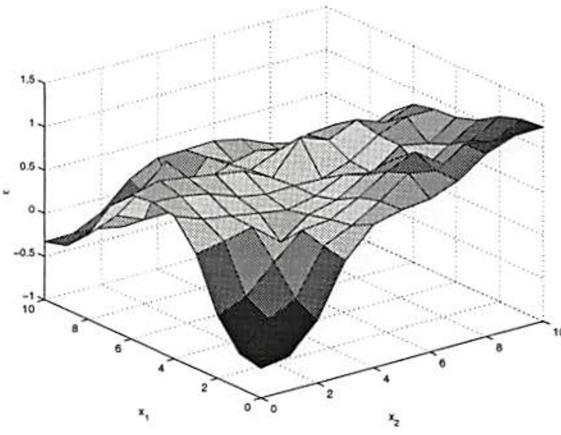


Figure 6.6: The plot of the difference between the crisp outputs for the two approaches versus (x_1, x_2) : keeping the response with the largest weight and preserving all the responses, in the type-1 case.

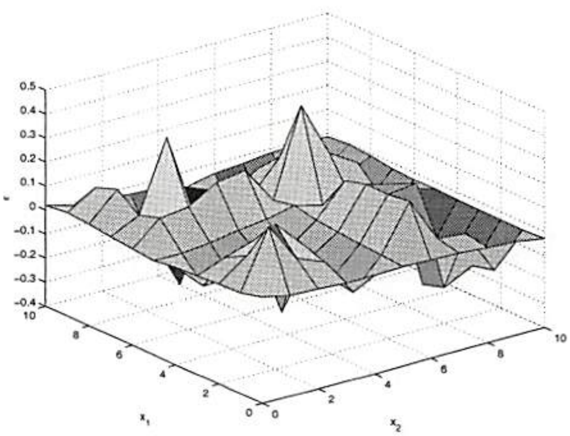


Figure 6.7: The plot of the difference between the crisp outputs for the two approaches versus (x_1, x_2) : averaging the responses and preserving all the responses, in the type-1 case.

in the first approach is the better method to choose. The first method in the first approach ignores all but the responses having a maximum weight. The computational complexity of the second approach is much greater than that for either of the methods in the first approach. The second approach, however, provides us with some information that neither of the two methods in the first approach does, namely, for every input, it gives us the outputs of *all* the possible FLS's that can be designed from the survey, along with their appropriate weights. If one desires to make use of this information in some way (e.g., during a decision-making process), the second approach should be adopted.

If a linguistic output, i.e., the name of the set to which the output belongs ($\tilde{F}^1$, $\tilde{F}^2$ or $\tilde{F}^3$), is desired from the type-2 FLS, it can be given as the set in which the crisp output has its maximum membership.

6.1.2 Type-2 Membership Functions

In Section 6.1.1, we fixed the parameters of the antecedent and consequent membership functions (means and standard deviations of the Gaussians). If we were to obtain these parameters by surveying experts, most likely we would obtain different answers from them; hence, in this case, we should also let the membership functions be uncertain. We do this by letting both the means and the standard deviations be uncertain, with the following values : the means of $\tilde{\tilde{F}}^1$, $\tilde{\tilde{F}}^2$, and $\tilde{\tilde{F}}^3$ are assumed to vary in $[0, 1]$, $[4.5, 5.5]$ and $[9.5, 10]$, respectively, and, the standard deviation of each Gaussian is assumed to be uncertain in $[1.5, 1.7]$. All these values are assumed to be uniformly uncertain, which means that the resulting type-2 sets are interval type-2 sets. Although we have chosen these values arbitrarily, they would usually be obtained by a survey as mentioned above.

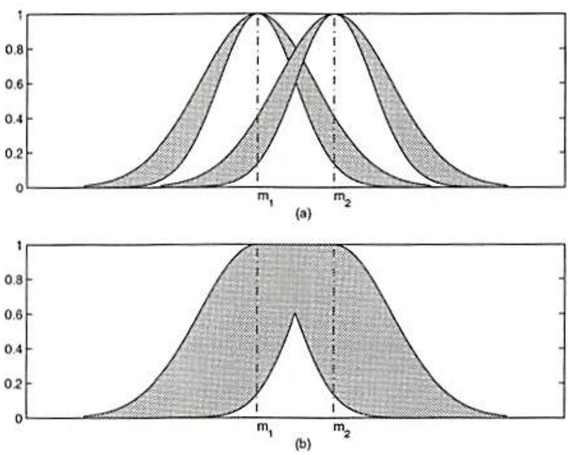


Figure 6.8: Figures depicting the procedure to obtain a type-2 set which is generated by a Gaussian with both uncertain mean and standard deviation. If the mean is uncertain in $[m_1, m_2]$, we draw the uncertain standard deviation curve centered at every point in interval $[m_1, m_2]$. Figure (a) shows the uncertain standard deviation curves at m_1 and m_2. Figure (b) shows the resulting type-2 set, obtained by combining all the curves similar to the one in (a).

Figure 6.1 (b) shows the membership functions for the three sets. In Chapter 1, we described the procedure for obtaining the membership function of a type-2 set obtained from a Gaussian type-1 set having either an uncertain mean or an uncertain standard deviation. The sets in Fig. 6.1 (b) have both of these uncertainties; however, since all the values of means and standard deviations are assumed equally uncertain, obtaining their type-2 membership function is a fairly easy task. See

Fig. 6.8 for an explanation. The centroids of $\tilde{\tilde{F}}^1$, $\tilde{\tilde{F}}^2$ and $\tilde{\tilde{F}}^3$ are computed, using the computational procedure described in Appendix D.1. These centroids are depicted in Fig. 6.1 (c).

Again, we consider two different approaches to handling the uncertainty due to multiple responses.

6.1.2.1 First Approach : Keeping One Response

Just as in Section 6.1.1.1, we consider two methods of choosing a single response : (1) by keeping the response with the largest weight; and, (2) by averaging the centroids of all the responses for each rule and using this average in place of the rule consequent centroid.

Method 1 : Keeping the Response with the Largest Weight

In this case, we use the rule-base in Table 6.2. The outputs (type-reduced sets) for five inputs, obtained using center-of-sets type-reduction are depicted in Fig. 6.9. Each of the sets in this case is an interval.

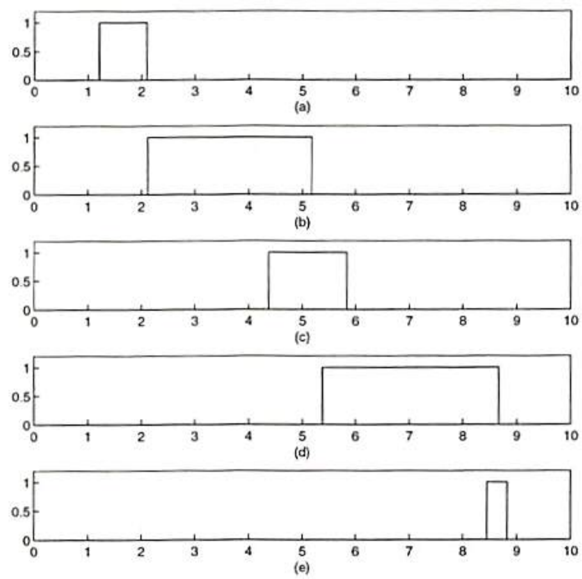


Figure 6.9: Five outputs of the type-2 FLS using rules from Table 6.2 (when the membership functions are assumed type-2) for the inputs : (a) $(0,0)$, (b) $(1,3)$, (c) $(5,5)$, (d) $(7,7.5)$, and (e) $(10,10)$.

We show the calculation of the type-reduced set for $(x_1, x_2) = (1,3)$. The comparable calculation, when the membership functions are type-1, is given in Section 6.1.1.1. The memberships of x_1 in the three antecedent sets $\tilde{\tilde{F}}^1$, $\tilde{\tilde{F}}^2$ and $\tilde{\tilde{F}}^3$ are the intervals $[0.8007, 1]$, $[0.0111, 0.1201]$ and 0, respectively. These membership grades are computed by drawing a vertical line passing through x_1 and noting its intersections with the shaded portions [see Fig. 6.1 (b)]. The membership grades of x_2 in $\tilde{\tilde{F}}^1$, $\tilde{\tilde{F}}^2$ and $\tilde{\tilde{F}}^3$ are $[0.1353, 0.5006]$, $[0.2494, 0.6775]$ and 0, respectively. Again the four rules, 1,2,4 and 5, are fired and their respective degrees of firing, which are computed by finding the *meet* of their antecedent membership grades, are the intervals $[0.1084, 0.5006]$, $[0.1997, 0.6775]$, $[0.0015, 0.0601]$ and $[0.0028, 0.0814]$, e.g., the degree of firing for rule 1 is computed as [see (2.32)]

$$\begin{aligned}\tilde{\mu}_{\tilde{\tilde{F}}^1}(x_1) \sqcap \tilde{\mu}_{\tilde{\tilde{F}}^1}(x_2) &= [0.8007, 1] \sqcap [0.1353, 0.5006] \\ &= [0.8007 \times 0.1353\,,\, 1 \times 0.5006] \\ &= [0.1084, 0.5006] \end{aligned} \quad (6.20)$$

The consequent of rule 1 is $\tilde{\tilde{F}}^1$ and that of rules 2, 4 and 5 is $\tilde{\tilde{F}}^2$; therefore, the degrees of firing for $\tilde{\tilde{F}}^1$ and $\tilde{\tilde{F}}^2$ are $[0.1084, 0.5006]$ and $[0.204, 0.819]$, respectively. The degree of firing for $\tilde{\tilde{F}}^2$ is obtained by adding the degrees of firing of rules 2, 4 and 5 [see (2.66)], as

$$\begin{aligned}
&[0.1997, 0.6775] + [0.0015, 0.0601] + [0.0028, 0.0814] \\
&= [0.1997 + 0.0015 + 0.0028\,,\ 0.6775 + 0.0601 + 0.0814] \\
&= [0.204, 0.819] \qquad (6.21)
\end{aligned}$$

The centroids of the fired consequents, $\tilde{\tilde{F}}^1$ and $\tilde{\tilde{F}}^2$ are [see Fig. 6.1 (c)] $\tilde{C}^1 = [1.2058, 1.7617]$ and $\tilde{C}^2 = [4.3562, 5.6288]$, respectively.

The output is computed using (5.22), as follows

$$\tilde{Y}_{cos}(1,3) = \int_{d_1\in[1.2058,1.7617]}\int_{d_2\in[4.3562,5.6288]}\int_{e_1\in[0.1084,0.5006]}\int_{e_2\in[0.204,0.819]} 1 \Big/ \frac{d_1e_1 + d_2e_2}{e_1+e_2} \qquad (6.22)$$

where, the integrals denote logical union. Note that, since $\tilde{C}^1$, $\tilde{C}^2$, $\tilde{E}^1$ and $\tilde{E}^2$ are all intervals, $\tilde{Y}_{cos}(1,3)$ is also an interval. Equation (6.22) is evaluated using the computational procedure described in Section D.1. $\tilde{Y}_{cos}(1,3)$ is depicted in Fig. 6.9 (b).

Recall, from Section 6.1.1.1, that when there is no uncertainty associated with the membership functions of the three fuzzy sets, $y_{cos}(1,3) = 4.0084$. Figure 6.9 (b) clearly reveals the effect of membership function uncertainty on this value. Note that the band of uncertainty is *not* symmetric about the point 4.0084, which means that the crisp output of the type-2 FLS is different than 4.0084. A plot of the crisp outputs for this type-2 FLS is depicted in Fig. 6.10. The difference plot between this crisp output and the output of the type-1 FLS keeping the maximum response (in Fig. 6.2) is shown in Fig. 6.11. It reveals the effects of antecedent and consequent uncertainties. The MSE for the results in Fig. 6.11 is 0.0735.

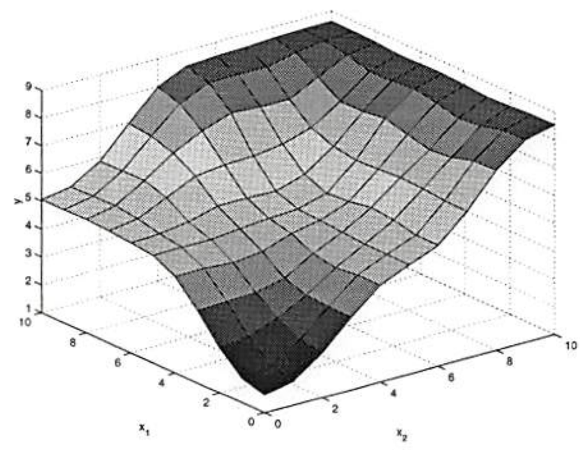


Figure 6.10: The plot of the crisp output of the type-2 FLS versus (x_1, x_2) using rules from Table 6.2, when the membership functions are assumed type-2.

Method 2 : Averaging the Responses

Here, we adopt the same strategy as in Section 6.1.2.1 Method 2, i.e., we create a type-2 TSK FLS (see Section 5.3.7), which has rules of the form :

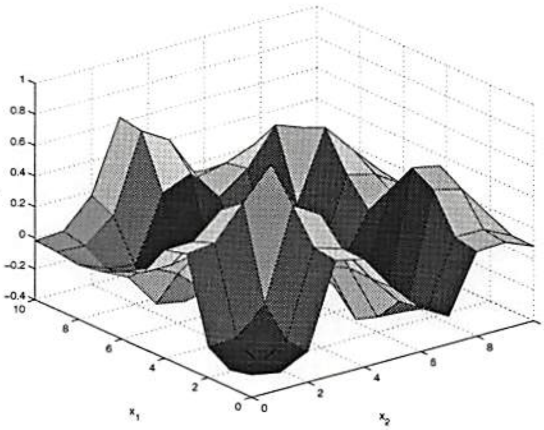


Figure 6.11: The plot of the difference between the crisp outputs for the type-1 and type-2 cases versus (x_1, x_2), when only the reponse with the maximum weight is chosen for each rule.

$$R^l : \text{IF } x_1 \text{ is } \tilde{\tilde{\mathrm{F}}}^i \text{ and IF } x_2 \text{ is } \tilde{\tilde{\mathrm{F}}}^j\text{, THEN } y = \tilde{\mathrm{C}}^l_{avg}$$

where

$$\tilde{\mathrm{C}}^l_{avg} = \frac{w_1^l \tilde{\mathrm{C}}^1 + w_2^l \tilde{\mathrm{C}}^2 + w_3^l \tilde{\mathrm{C}}^3}{w_1^l + w_2^l + w_3^l} \tag{6.23}$$

where, the "+" signs denote algebraic sum, $\tilde{\mathrm{C}}^i$ is the centroid of $\tilde{\tilde{\mathrm{F}}}^i$ $(i = 1, 2, 3)$, and w_i^l are the weights obtained from Table 6.1. The rules for this FLS are listed in Table 6.5. The type-reduced sets for five different inputs are depicted in Fig. 6.12. Each of the type-reduced sets in this case also is an interval. Comparing Figs. 6.9 and 6.12, we see the difference between keeping the responses with a maximum weight and averaging the responses. The difference is larger when there is larger disagreement among the experts. Observe, for example, that the type-reduced sets for $(x_1, x_2) = (10, 10)$, when there is no disagreement among the experts (see Table 6.1), are the same in Figs. 6.9 and 6.12; but, those for $(x_1, x_2) = (0, 0)$, when there is some disagreement among the experts, are different.

Table 6.5: Averaging the responses, when the membership functions are type-2. For each rule, $\tilde{\mathrm{C}}^l_{avg}$'s are weighted averages of consequent centroids, c^1, c^2 and c^3, where the weights are obtained from the Table 6.1. In each case, $\tilde{\mathrm{C}}^l_{avg}$ is an interval set.

Rule no. (l)	x_1	x_2	$\tilde{\mathrm{C}}^l_{avg}$
1	1	1	[1.8359, 2.5351]
2	1	2	[3.7736, 4.8216]
3	1	3	[7.0683, 7.6702]
4	2	1	[3.8210, 4.7879]
5	2	2	[4.0886, 5.2084]
6	2	3	[7.4308, 8.0232]
7	3	1	[4.9559, 5.8131]
8	3	2	[8.0457, 8.5021]
9	3	3	[8.4556, 8.8213]

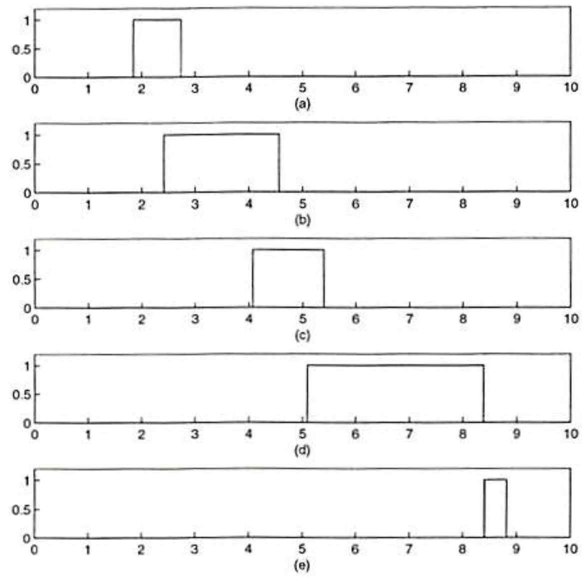


Figure 6.12: Five outputs of the type-2 FLS using rules from Table 6.5 (when the membership functions are assumed type-2) for the inputs : (a) $(0,0)$, (b) $(1,3)$, (c) $(5,5)$, (d) $(7,7.5)$, and (e) $(10,10)$.

We show the calculation for the type-reduced set corresponding to the input $(x_1, x_2) = (1,3)$. The membership grades of x_1 and x_2 in the antecedent fuzzy sets and the degrees of firing for the rules are the same as calculated in Method 1. In this case, the rule consequents for rule numbers 1, 2, 4 and 5 are $\tilde{C}^1_{avg} = [1.8359, 2.5351]$, $\tilde{C}^2_{avg} = [3.7736, 4.8216]$, $\tilde{C}^4_{avg} = [3.8210, 4.7879]$ and $\tilde{C}^5_{avg} = [4.0886, 5.2084]$, respectively. The consequent for the first rule, for example, is calculated as [see (6.23) and Theorem D.1]

$$\begin{aligned}\tilde{C}^1_{avg} &= 0.8 \times \tilde{C}^1 + 0.2 \times \tilde{C}^2 \\ &= 0.8 \times [1.2058, 1.7617] + 0.2 \times [4.3562, 5.6288] \\ &= [0.8 \times 1.2058 + 0.2 \times 4.3562\,,\, 0.8 \times 1.7617 + 0.2 \times 5.6288] \\ &= [1.8359, 2.5351] \end{aligned} \tag{6.24}$$

The type-reduced set is computed using (5.22), as

$$\tilde{Y}_{TSK}(1,3) = \int_{k^1}\int_{k^2}\int_{k^4}\int_{k^5}\int_{e^1}\int_{e^2}\int_{e^4}\int_{e^5} 1 \Bigg/ \frac{k^1e^1 + k^2e^2 + k^4e^4 + k^5e^5}{e^1 + e^2 + e^4 + e^5} \tag{6.25}$$

where : the integrals denote logical union; $k^l \in \tilde{C}^l_{avg}$ for $l = 1,2,4,5$, where $\tilde{C}^l_{avg}$ are as mentioned in the previous paragraph; and, $e^l \in \tilde{E}^l$ for $l = 1,2,4,5$, where $\tilde{E}^1, \tilde{E}^2, \tilde{E}^4, \tilde{E}^5$, the degrees of firing associated with the 4 rules (see Method 1), are the intervals $[0.1084, 0.5006]$, $[0.1997, 0.6775]$, $[0.0015, 0.0601]$ and $[0.0028, 0.0814]$, respectively. The quantity to the left of the slash in (6.25) is unity because all the sets involved are crisp. $\tilde{Y}_{TSK}(1,3)$ is obtained using the computational procedure described in Section D.1.

In this case, since we have averaged the responses, each rule has a different consequent; therefore, the term to the right of the slash on the RHS of (6.25) has four terms each in the numerator and the denominator. In the first method, three of the four fired rules (rule numbers 2, 4 and 5) had the same consequent, $\tilde{\tilde{F}}^2$; and therefore, the term to the right of the slash on the RHS of (6.22) contains only two terms each in the numerator and the denominator.

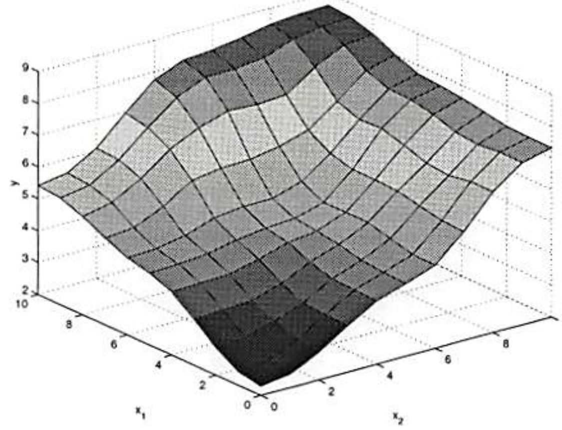


Figure 6.13: The plot of the crisp output of the type-2 FLS versus (x_1, x_2) using rules from Table 6.5, when the membership functions are assumed type-2.

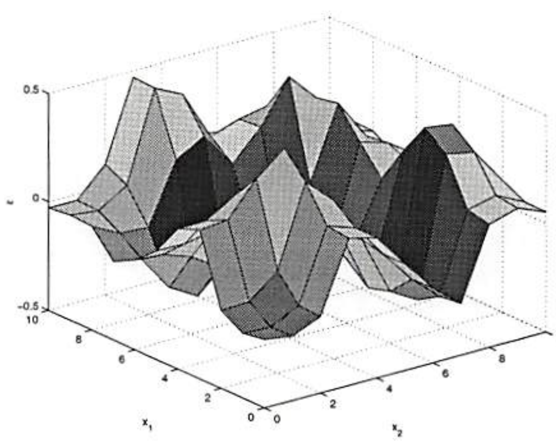


Figure 6.14: The plot of the difference between the crisp outputs for the type-1 and type-2 cases versus (x_1, x_2), when the reponses are averaged for each rule.

Crisp output for this type-2 TSK FLS is obtained by computing the centroid of the type-reduced set and is depicted in Fig. 6.13. The difference plot between this crisp output and the output of the type-1 FLS using averaged responses (Fig. 6.3) is shown in Fig. 6.14. It shows the effect of antecedent and consequent uncertainties on the crisp output of the FLS. The MSE for the results in Fig. 6.14 is 0.0402.

6.1.2.2 Second Approach : Preserving All the Responses

In this case, we use the rule-base in Table 6.4, with the type-1 fuzzy centroids $\tilde{C}^1$, $\tilde{C}^2$ and $\tilde{C}^3$ replacing c^1, c^2 and c^3, respectively. We also use the membership grades in the $\tilde{K}^l$'s to scale the membership functions of $\tilde{C}^1$, $\tilde{C}^2$ and $\tilde{C}^3$, e.g., if the domains of $\tilde{C}^1$ and $\tilde{C}^2$ are $[c_1^1, c_2^1]$ and $[c_1^2, c_2^2]$, respectively, then $\tilde{K}^1$ (see Table 6.4) will be equal to $1/[c_1^1, c_2^1] + 0.25/[c_1^2, c_2^2]$, where the + sign indicates logical union. Using center-of-sets type-reduction, the outputs (type-reduced sets) for the inputs $(0,0)$, $(1,3)$, $(5,5)$, $(7,7.5)$ and $(10,10)$ are shown in Fig. 6.15. The type-reduced sets in this case may not be very accurate, because the discretization performed for the actual computation of (6.30) was somewhat coarse.

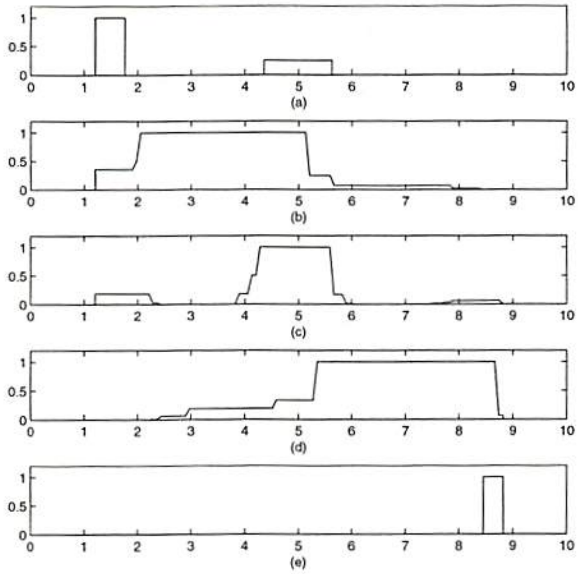


Figure 6.15: Five outputs of the type-2 TSK FLS using rules from Table 6.4 (when the membership functions are assumed type-2) for the inputs : (a) $(0,0)$, (b) $(1,3)$, (c) $(5,5)$, (d) $(7,7.5)$, and (e) $(10,10)$.

The type-reduced set corresponding to the input $(x_1, x_2) = (1,3)$ was computed as described next. The antecedent memberships and the degrees of firing of the rules are the same as described at the end of Section 6.1.2.1. Recall that the consequents of $\tilde{\tilde{\mathrm{F}}}^1$, $\tilde{\tilde{\mathrm{F}}}^2$ and $\tilde{\tilde{\mathrm{F}}}^3$ are $\tilde{\mathrm{C}}^1 = [1.2058, 1.7617]$, $\tilde{\mathrm{C}}^2 = [4.3562, 5.6288]$ and $\tilde{\mathrm{C}}^3 = [8.4556, 8.8213]$, respectively [see Fig. 6.1 (c)]; therefore, the consequents in this case are

$$\begin{aligned}
\tilde{\mathrm{K}}^1 &= 1/\tilde{\mathrm{C}}^1 + 0.25/\tilde{\mathrm{C}}^2 \\
&= 1/[1.2058, 1.7617] + 0.25/[4.3562, 5.6288] \qquad (6.26)\\
\tilde{\mathrm{K}}^2 &= 0.3571/\tilde{\mathrm{C}}^1 + 1/\tilde{\mathrm{C}}^2 + 0.0714/\tilde{\mathrm{C}}^3 \\
&= 0.3571/[1.2058, 1.7617] + 1/[4.3562, 5.6288] + 0.0714/[8.4556, 8.8213] \qquad (6.27)\\
\tilde{\mathrm{K}}^4 &= 0.5/\tilde{\mathrm{C}}^1 + 1/\tilde{\mathrm{C}}^2 + 0.1667/\tilde{\mathrm{C}}^3 \\
&= 0.5/[1.2058, 1.7617] + 1/[4.3562, 5.6288] + 0.1667/[8.4556, 8.8213] \qquad (6.28)\\
\tilde{\mathrm{K}}^5 &= 0.1875/\tilde{\mathrm{C}}^1 + 1/\tilde{\mathrm{C}}^2 + 0.0625/\tilde{\mathrm{C}}^3 \\
&= 0.1875/[1.2058, 1.7617] + 1/[4.3562, 5.6288] + 0.0625/[8.4556, 8.8213] \qquad (6.29)
\end{aligned}$$

where, in each case, the number to the left of the slash indicates the membership grade of each point in the interval to the right of the slash, and the "+" sign indicates logical union. Figure 6.16 depicts $\tilde{\mathrm{K}}^1$, $\tilde{\mathrm{K}}^2$, $\tilde{\mathrm{K}}^4$ and $\tilde{\mathrm{K}}^5$.

The output is computed using (5.22), as

$$\tilde{\mathrm{Y}}_{TSK}(1,3) = \int_{k^1}\int_{k^2}\int_{k^4}\int_{k^5}\int_{e^1}\int_{e^2}\int_{e^4}\int_{e^5} \prod_{\substack{l=1 \\ l\neq 3}}^{5} \mu_{\tilde{\mathrm{K}}^l}(k^l) \Bigg/ \frac{k^1e^1 + k^2e^2 + k^4e^4 + k^5e^5}{e^1 + e^2 + e^4 + e^5} \qquad (6.30)$$

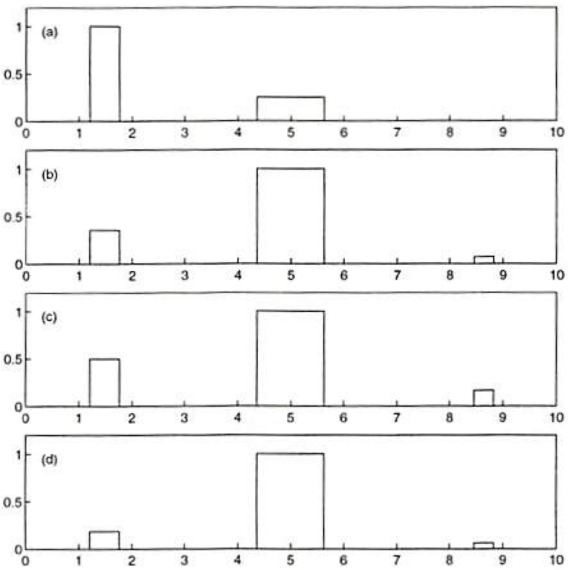


Figure 6.16: The fired rule consequents, $\tilde{K}^1$ [Fig. (a)], $\tilde{K}^2$ [Fig. (b)], $\tilde{K}^4$ [Fig. (c)], and $\tilde{K}^5$ [Fig. (d)], corresponding to the input $(x_1, x_2) = (1, 3)$, of the type-2 TSK FLS when the membership functions are type-2 and all the responses are preserved.

where : the integrals denote logical union; $k^l \in \tilde{K}^l$ for $l = 1, 2, 4, 5$, where the $\tilde{K}^l$'s are given by Eqs. (6.26) - (6.29); and, $e^l \in \tilde{E}^l$ for $l = 1, 2, 4, 5$, where $\tilde{E}^1, \tilde{E}^2, \tilde{E}^4, \tilde{E}^5$, the degrees of firing associated with the 4 rules (see Section 6.1.2.1), are the intervals $[0.1084, 0.5006]$, $[0.1997, 0.6775]$, $[0.0015, 0.0601]$ and $[0.0028, 0.0814]$, respectively. Note that the term $\prod_{l=1}^{4} \mu_{\tilde{E}^l}(e^l)$, which appears in (5.22), does not appear to the left of the slash on the RHS of (6.30) because it is equal to 1, since all the degrees of firing are intervals. $\tilde{Y}_{TSK}(1, 3)$ is computed as described in Section 5.3.5, and is depicted in Fig. 6.15 (b). A crisp output can be obtained for this FLS by finding the centroid of the type-reduced set. Figure 6.17 shows a plot of the crisp output versus (x_1, x_2).

Figures 6.4 and 6.15 both show all the possible outputs due to different answers obtained from the survey. Figure 6.15, however, also shows the variation in the output due to the uncertain nature of the membership functions. It is also interesting to compare the figures in Fig. 6.15 with their counterparts in Figs. 6.9 and 6.12. Doing this reveals the effect of preserving all the responses versus keeping the response with the largest weight.

Figure 6.18 shows the difference in the crisp outputs for the second approach in the type-1 and type-2 cases (Figs. 6.5 and 6.17). This shows the effect of antecedent and consequent uncertainties on the crisp output for the second approach. The MSE for the results in Fig. 6.18 is 0.1930.

An interesting observation, that can be made comparing the MSE's for Figs. 6.11 (MSE = 0.0735), 6.14 (MSE = 0.0402) and 6.18 (MSE = 0.1930), is that neglecting antecedent and consequent uncertainties causes the largest MSE's for the second approach and the smallest MSE's for the second method (averaging the responses) in the first approach.

If a linguistic output is desired from the type-2 FLS, it may be obtained by finding the crisp output and finding the fuzzy set in which it has maximum membership.

6.1.2.3 A Comparison of the Two Approaches

Our observations and recommendations in the type-2 case are quite similar to those in the type-1 case. The computational complexity of the first approach is significantly lower than that of the second approach. This is because of the fact that all the sets involved in the first approach are

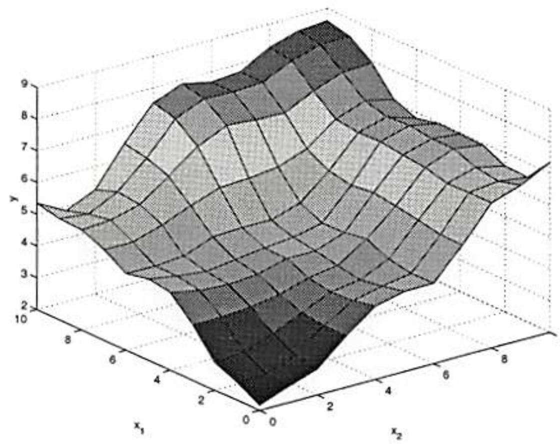


Figure 6.17: The plot of the crisp output of the type-2 FLS versus (x_1, x_2) using the second approach, when the membership functions are assumed type-2.

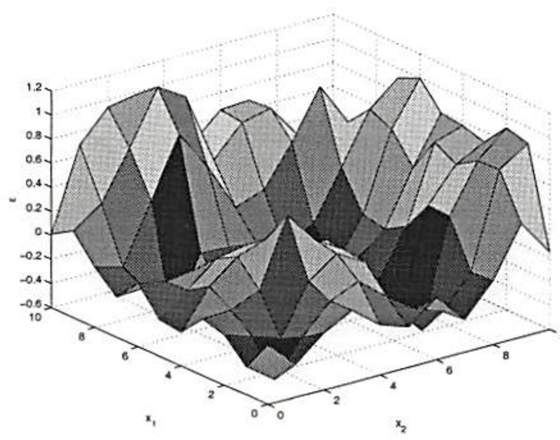


Figure 6.18: The plot of the difference between the crisp outputs for the type-1 and type-2 cases versus (x_1, x_2), when the second approach is used.

interval sets and we have an efficient algorithm for type-reduction computations of interval type-2 FLS's. The FLS in the second approach is not an interval type-2 FLS, because the consequents of this TSK FLS are not interval type-1 sets [see (6.26) - (6.29)]; and therefore, the type-reduction computations are very expensive.

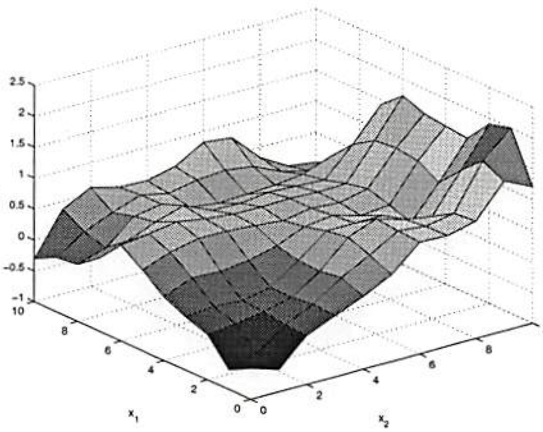


Figure 6.19: The plot of the difference between the crisp outputs for the two approaches versus (x_1, x_2) : keeping the response with the largest weight and preserving all the responses, in the type-2 case.

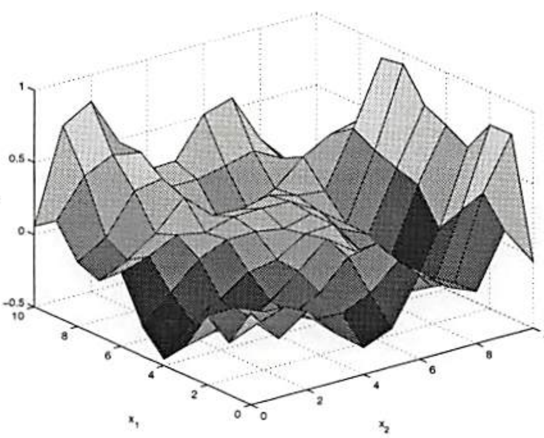


Figure 6.20: The plot of the difference between the crisp outputs for the two approaches versus (x_1, x_2) : averaging the responses and preserving all the responses, in the type-2 case.

Figure 6.19 depicts the difference between the crisp output of the second approach and the crisp output of the first method in the first approach. Figure 6.20 presents comparable results for the crisp output of the second approach and the crisp output of the second method in the first approach. Note that the MSE's associated with Figures 6.19 and 6.20 are 0.4813 and 0.0857, respectively. Based on MSE's, if one is interested only in the crisp output, we recommend using the first approach with averaging of the responses. Because of its high computational complexity, we recommend using the second approach only when the additional information obtained from it (i.e., the uncertainty in the output due to multiple responses) is necessary.

6.1.3 Conclusions

From the results of Sections 6.1.1 and 6.1.2, we conclude that in both type-1 and type-2 cases, *if one is interested only in the crisp output of the FLS, averaging of responses is the most efficient method.* The additional complexity of preserving all the responses makes this practical only when the additional information obtained from it is absolutely necessary.

We also observed that linguistic uncertainties in the membership functions, which can be incorporated in a type-2 FLS, do affect the crisp output of the FLS, i.e., a type-2 FLS gives us more information about the output than a type-1 FLS, when the antecedents and/or consequents are uncertain.

6.2 Time-Series Prediction

In this example, we demonstrate how available information about the training data can be incorporated into a type-2 FLS to obtain more information about its output than can be obtained using a type-1 FLS. We consider the following problem : A FLS is trained with noisy data. If we know the noise strength, how can we obtain bounds on the output of the FLS within which the true value of the output is likely to lie ?

We demonstrate this with the example of a 4 input - 1 output, one step predictor for the Mackey-Glass chaotic time-series [14]. If $x(k)$ $[k = 1, 2, 3, \ldots]$ is the time-series, given $x(k-3)$, $x(k-2)$, $x(k-1)$ and $x(k)$, we predict $x(k+1)$. It has already been shown that type-1 fuzzy logic systems do a good job at modelling this chaotic time-series [19, 26]. We consider the situation when the available training data is noisy, and see how we can get more information about the true value of the output if we have some information about the noise.

The Mackey-Glass time-series is generated using the following delay differential equation [14]:

$$\frac{dx(t)}{dt} = \frac{0.2x(t-\tau)}{1 + x^{10}(t-\tau)} - 0.1x(t) \tag{6.31}$$

We use $\tau = 30$ (for which it is known that the Mackey-Glass time series is chaotic) and initial condition $x_0 = 0.1$. The FLS is trained with (i.e., the rules are formed from) 500 input-output pairs using $x(1001)$, $x(1002)$, ..., $x(1504)$ [the first input-output pair is $\{[x(1001), x(1002), x(1003), x(1004)], x(1005)\}$; the next one is $\{[x(1002), x(1003), x(1004), x(1005)], x(1006)\}$; and so on]. The training data is corrupted by zero-mean uniform noise. Three different signal-to-noise ratios (SNRs) are considered : 0, 10 and 20 dB. In each case, the FLS is tested on 200 points from $x(1505)$ to $x(1708)$.

6.2.1 Using Interval Type-2 Sets

Here we make use of interval type-2 sets. The following approach is adopted : we design a type-1 FLS using the available data, and then create a type-2 FLS from this type-1 FLS by incorporating information that is available about the noise. Then, given any input, we use this type-2 FLS to obtain the range of values within which the true output is likely to lie.

Note that both the type-1 and type-2 FLS's are designed based on *a single available training realization.* If we are given a different noisy realization of the same data set, we would normally choose a different set of type-1 and type-2 FLS parameters to obtain the best predictions. This is generally what happens in practice, where only a single realization is available. *We, therefore, provide only single realization results here.*

6.2.1.1 Designing the Type-1 FLS

We consider only numerical information, and design a type-1 FLS from the given input-output pairs using the simple one-pass algorithm described in [26]. The FLS uses Gaussian type-1 fuzzy sets, singleton fuzzification, product t-norm, product inference and center-of-sets defuzzification. In this case, the domain of each of the 4 inputs and the range are identical. The interval between the minimum and maximum values of the training set is divided into 31 fuzzy sets.

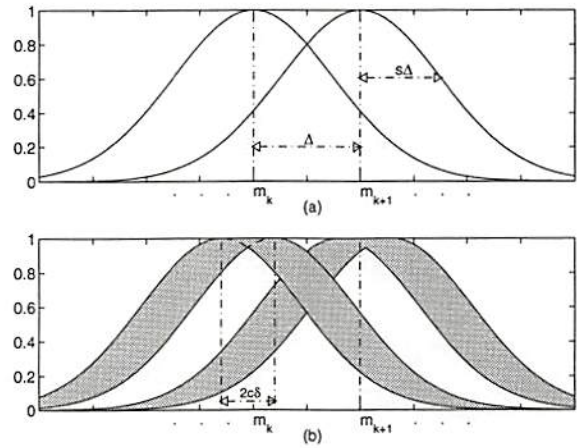


Figure 6.21: (a) Gaussian fuzzy sets used by the type-1 FLS in the time-series forecasting example. The sets are equispaced, the distance between two consecutive centers being Δ. The standard deviation of each set is equal to $s\Delta$, where the multplier s is determined experimentally. (b) Two type-2 sets created by displacing the means of the two type-1 sets shown in Fig. (a). These sets are used to design a type-2 FLS for the time-series forecasting example. The mean of each Gaussian in the type-2 FLS is assumed to be uncertain in an interval of length $2c\delta$. Note that this interval of length $2c\delta$ is not necessarily symmetrical about the mean of the corresponding set in the type-1 FLS.

Figure 6.21 (a) shows two of the 31 fuzzy sets used by this type-1 FLS. The centers of these sets are equispaced (the distance between two consecutive centers is Δ), and remain fixed during the design procedure. Their values depend on the training sample. All the sets have the same standard deviation, which is set equal to $s\Delta$, where the multiplier s is chosen experimentally, so that, for the given training realization, the type-1 FLS gives the lowest mean-square error (MSE) with the true output (note that, in practice, it is not necessary that s be the same for all the sets in the FLS). Subsequent training involves formation of the rule-base (using the algorithm described in [26]) and determining the multiplier s. See Section 6.2.1.3 for numerical values of the parameters of this type-1 FLS for different training samples. Note that we do not claim to have designed an optimum type-1 FLS, i.e., it may be possible to obtain better predictions using some other design methodology.

6.2.1.2 Designing the Type-2 FLS

Once the type-1 FLS is designed, we create a type-2 FLS from it by incorporating information about the noise. Similar to the type-1 FLS, the type-2 FLS uses singleton fuzzification, product t-norm, product inference, and center-of-sets type-reduction. It also uses the same number of fuzzy sets and the same rules as the type-1 FLS. The only difference now is that the antecedent and consequent sets are type-2. To design these type-2 sets, we reason as described next.

Our aim is to model the uncertainty introduced in the FLS due to the noisy training data by using type-2 fuzzy sets. If we had many different realizations of the training data and if we followed the same design procedure for the type-1 FLS as explained above, the means of the 31 sets would not remain the same from realization to realization. They would vary in some range depending upon the strength of the noise.

Recall that a zero-mean, uniform random variable with standard deviation σ_{noise}, can take values in $[-\delta, \delta]$, where $\delta = \sqrt{3}\sigma_{noise}$. (The variance of a uniform random variable in $[-\delta, \delta]$ is $\delta^2/3$.) The possible range within which each mean can vary, therefore, is $\pm\delta$, where $\delta = \sqrt{3}\sigma_{noise}$. So, we model each set in the type-2 FLS as a type-2 set associated with a Gaussian type-1 set, where the latter has an uncertain mean [see Example 1.2, and also Fig. 6.21 (b)]. For simplicity, we assume that there is no uncertainty in the standard deviations of the type-1 Gaussians. In every type-2 set, all the possible values of the mean are assumed to be equally uncertain, so that we set all the secondary memberships equal to unity, in which case, the type-2 set becomes an interval type-2 set (see Section 1.3). The interval, within which the mean of each Gaussian can vary, is set equal to $2c\delta$ [see Fig. 6.21 (b)], where the multiplier c is a design parameter. Changing the value of c changes the width of the type-reduced set; however, it is not very easy to visualize the effect of changing c on the output, because the antecedents and consequents both are type-2 in this case. A suitable value for c can, therefore, be found by experimenting with different values to get tight bounds on the output.

Note that the major strength of a type-2 FLS lies in predicting the uncertainty in the output given the uncertainty in the data used to design the system. This is something that is not possible using a type-1 FLS. The crisp output of the type-2 FLS, however, may also be used to obtain a better prediction than obtained by a type-1 FLS, as explained next.

Assuming that the type-2 FLS uses the same standard deviations for its uncertain-mean Gaussians as used by the type-1 FLS, after the type-1 FLS has been completely designed, the parameter c in the type-2 FLS still needs to be determined. Just as it isn't necessary to choose the parameter s in the type-1 FLS to be the same for every type-1 set, it isn't necessary to choose the parameter c to be the same for all the sets. The range of uncertainty for a type-1 mean m_1 is also not always required to be $[m_1 - c\delta, m_1 + c\delta]$; it can be, for example, $[m_1 - 2c\delta, m_1]$ or $[m_1, m_1 + 2c\delta]$ or anywhere in between these two extremes. The length of this interval, however, is always $2c\delta$ [see Fig. 6.21 (b)]. So, there are more parameters to tune in a type-2 FLS than a type-1 FLS; and, given any realization of the training data, we can tune the additional type-2 parameters to obtain a better performance over the type-1 FLS.

6.2.1.3 Results

Figures 6.22 - 6.24 show the outputs of the type-1 and type-2 FLS's, for *single realizations* of noisy training data. The parameters of the type-2 FLS in each case were tuned, as explained in the previous paragraph, so that they gave better predictions than the type-1 FLS. For each case, we see that the true time-series values lie almost completely within the bounds provided by the type-2 FLS. Since we reduce the amount of type-2 uncertainty (i.e., the interval of uncertainty of each mean) as SNR increases (the range of uncertainty of each mean is proportional to the standard deviation of the noise), the bands grow narrower as SNR increases. The MSE's between the type-1 FLS and the true output in the three cases ($0dB$, $10dB$ and $20dB$) are 1.97×10^{-2}, 4.72×10^{-3} and 1.63×10^{-3}, respectively; and those between the crisp output of the type-2 FLS and the true output are 1.34×10^{-2}, 4.38×10^{-3} and 1.59×10^{-3}, respectively. The most significant improvement in MSE's is at the low SNR values.

To see how many points lie outside the upper and lower bounds, we computed a "bound error" as follows : for points lying above the upper bound, we computed the squared error with the upper bound; for points lying below the lower bound, we computed the squared error with the lower bound; and, then we found the mean of all these squared errors with the bounds (this includes points which don't exceed the bounds). The bound errors for $0dB$ and $10dB$ are 0, and the bound error for $20dB$ is 3.33×10^{-5}; so, the bounds provided by center-of-sets type-reduction are quite good. As the SNR increases, the distance between the upper and lower bounds grows smaller; therefore, more points lie outside the bounds.

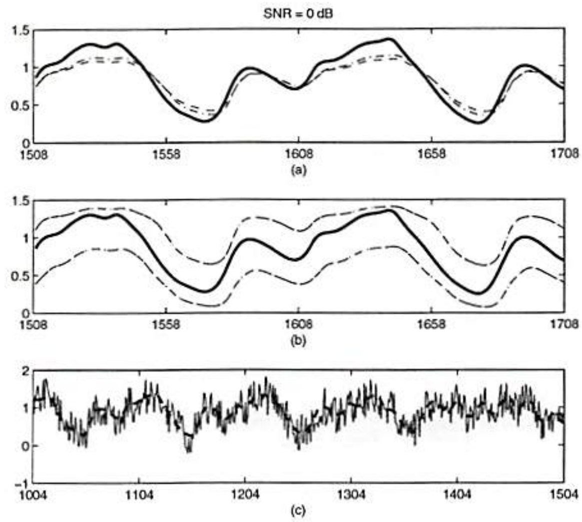


Figure 6.22: Mackey-Glass chaotic time-series prediction in the presence of uniform noise, when SNR = $0dB$. The thick solid line in (a) and (b) indicates the true time-series. In (a), the dashed line indicates the type-1 output and the dash-dotted line indicates the type-2 crisp output. Figure (b) shows the upper and lower bands obtained with the type-2 FLS. In Fig. (c), the solid line shows the noisy data used for training and the thick dashed line shows the noisefree data. Training is performed with the 500 input-output pairs in $x(1001), x(1002), \ldots, x(1504)$ and testing is done with 200 input-output pairs in $x(1505), x(1506), \ldots, x(1708)$. The type-2 FLS uses interval type-2 sets.

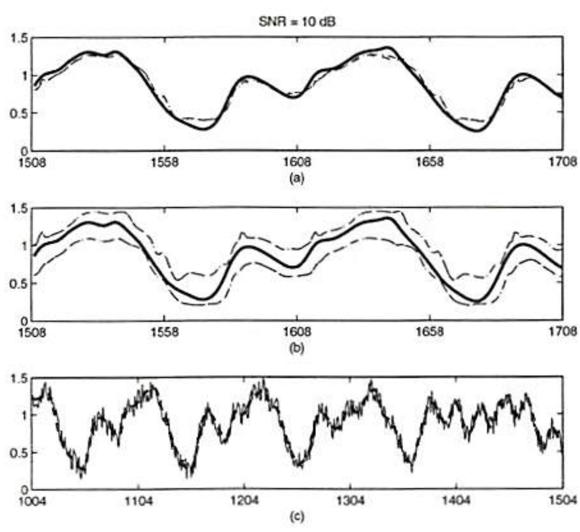


Figure 6.23: Mackey-Glass chaotic time-series prediction in the presence of uniform noise, when SNR = $10dB$. The rest of the details are the same as in the caption to Fig. 6.22.

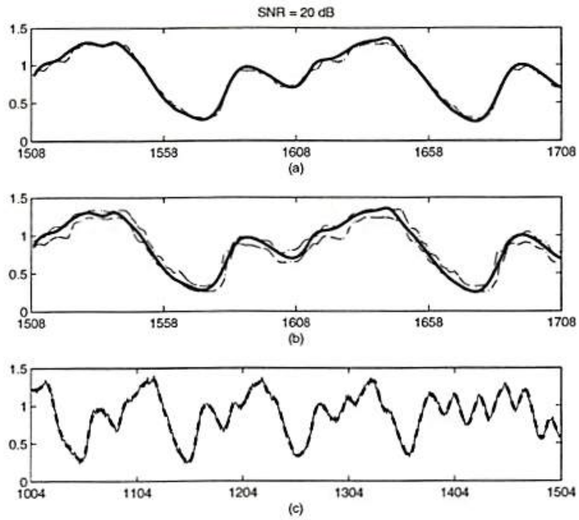


Figure 6.24: Mackey-Glass chaotic time-series prediction in the presence of uniform noise, when SNR $= 20dB$. The rest of the details are the same as in the caption to Fig. 6.22.

The type-1 and type-2 parameters chosen for the three examples shown in Figures 6.22 - 6.24, are summarized in Table 6.6. Observe, from this table, that the range within which the noisy time-series values lie, $(m_{31} - m_1)$, decreases as the SNR increases.

The calculations for this example were performed as follows (see, also, Section D.1.2). The centroids of the consequent sets were computed using the Appendix D.1 computational procedure, and were stored for further use; for every applied input, the degree of firing for each rule $[\sqcap_{i=1}^{4}\mu_{\tilde{F}_i^l}(x_i)]$ was computed by finding the *meet* of the antecedent memberships for that rule (see Example 2.3); and finally, the type-reduced set was computed using the Appendix D.1 computational procedure again.

6.2.1.4 Comments

In the examples considered above, the FLS's were hand tuned, i.e., all the parameters were manually adjusted. Even better results might be obtained by using, for example, a steepest descent procedure.

In practice, if we do not know the actual strength of the noise, we can assume a reasonable value for it. If the noise is not uniform, we can still choose a reasonable interval of uncertainty for any input (e.g., $\pm 2\sigma$ points for Gaussian noise) and use uncertain-mean Gaussian sets (like the ones used in this section) in the type-2 FLS.

6.2.2 Using Gaussian Type-2 Sets

In this section, we show how Gaussian type-2 sets can be used for the time-series prediction example. We adopt an approach very similar to the one that we adopted when interval type-2 sets were used. We first design a type-1 FLS using the available noisy data; and, then we create a Gaussian type-2 FLS by incorporating our knowledge about the noise into this type-1 FLS. Again, *we provide only single realization results.* Because, at present, we do not have any fast agorithms to compute the actual type-reduced set for a Gaussian type-2 FLS, we make use of the Gaussian approximation (see Section 5.7); but, since the approxmation can be used only when the type-2 uncertainty is small [see Theorem (2.5)], we can use this type-2 FLS only for high SNR's. This limits the use of this approach considerably.

Table 6.6: Parameters of the type-1 and type-2 FLS's for the forecasting example. $m_1, \ldots, m_{31}$ denote the means of the 31 type-1 sets and the 31 type-2 sets are denoted as $\tilde{\tilde{\mathrm{F}}}_1, \ldots, \tilde{\tilde{\mathrm{F}}}_{31}$. Note that, when the SNR is $20dB$, the c values are different for different sets.

SNR	0dB	10dB	20dB
m_1	-0.2038	0.1437	0.2382
m_{31}	1.7989	1.4749	1.3968
$\Delta = \sqrt{3}\sigma_{noise}$	0.0668	0.0444	0.0386
s	6.75	2.75	1.75
δ	0.4928	0.1558	0.0493
c	0.4	0.5	0.75 for $\tilde{\tilde{\mathrm{F}}}_1, \ldots, \tilde{\tilde{\mathrm{F}}}_6$, and $\tilde{\tilde{\mathrm{F}}}_{26}, \ldots, \tilde{\tilde{\mathrm{F}}}_{31}$; and, 0.6 for $\tilde{\tilde{\mathrm{F}}}_7, \ldots, \tilde{\tilde{\mathrm{F}}}_{25}$
Range of uncertainty for the means	$[m_i - c\delta, m_i]$ for $i = 1, \ldots, 15$; $[m_i - c\delta/2, m_i + c\delta/2]$ for $i = 16$; and, $[m_i, m_i + c\delta]$ for $i = 17, \ldots, 31$	$[m_i - c\delta, m_i]$ for $i = 1, \ldots, 7$; $[m_i - c\delta/2, m_i + c\delta/2]$ for $i = 8, \ldots, 24$; and, $[m_i, m_i + c\delta]$ for $i = 25, \ldots, 31$	$[m_i - c\delta, m_i]$ for $i = 1, \ldots, 6$; $[m_i - c\delta/2, m_i + c\delta/2]$ for $i = 7, \ldots, 25$; and, $[m_i, m_i + c\delta]$ for $i = 26, \ldots, 31$

6.2.2.1 Designing the Type-1 FLS

The type-1 FLS is designed in exactly the same manner as described in Section 6.2.1.1.

6.2.2.2 Designing the Type-2 FLS

We create a type-2 FLS from this type-1 FLS by replacing the Gaussian type-1 sets by Gaussian type-2 sets. The Gaussian type-2 sets are designed as described next.

Using the same rationale as in the case of interval type-2 sets, we can say that if we had many different realizations of the training data, the means of the Gaussians in different type-1 FLS's, designed based on these different realizations, would vary in some range. If we assume that the noise added is Gaussian, we should ideally design every set in the type-2 FLS as a type-2 set resulting from a Gaussian type-1 set having a mean that is also uncertain, and that is described by a Gaussian membership function (see Appendix A - Example A.2 for an example of such a set); however, since the equations for this kind of type-2 sets are very complicated, we use Gaussian type-2 sets (see Section 1.3) instead. Since our type-2 FLS uses Gaussian type-1 sets, we use Gaussian type-2 sets having Gaussian principal membership functions. We also assume that the standard deviations of the secondary membership functions are proportional to the means of the secondary membership functions, so that the 2-D representation of the type-2 set looks like the one shown in Fig. 1.10.

The Gaussian type-2 set described in the previous paragraph can be described using three parameters : the mean, M, and standard deviation, Σ, of the principal membership function, and the constant of proportionality c, for the mean and standard deviation of the secondary membership functions [for any x, $\sigma(x) = cm(x)$, where $m(x)$ and $\sigma(x)$ are, respectively, the mean and standard deviation of the membership grade of x]. Of these three parameters, we fixed M and Σ to be equal to the mean and standard deviation of the respective type-1 sets in the type-1 FLS; and, we set all the type-1 sets to have the same standard deviation.

To decide the value of parameter c, we proceed as follows. Recall that, since M is the mean of the principal membership function of the Gaussian type-2 set [i.e., $m(M) = 1$], the secondary membership function corresponding to M is a Gaussian centered at 1. Consequently, $\sigma(M) = c$, i.e., c is the standard deviation of the membership grade of M. We find an appropriate value for this standard deviation next.

We first design a type-2 set by letting the mean of a Gaussian with standard deviation Σ be a Gaussian type-1 set with mean M and standard deviation σ_{noise}, where σ_{noise} is the noise standard deviation (see Fig. A.2). Then, we find the secondary membership function corresponding to M in this type-2 set and approximate it with a Gaussian. Finally, using the standard deviation of this Gaussian as the parameter c, we design the Gaussian type-2 sets that we actually use for the computations.

The procedure is not as complicated as it sounds. Let M be the mean of the Gaussian type-1 set and, let $[m_1, m_2]$ be its range of uncertainty. In our case, we set $m_1 \triangleq M - 2\sigma_{noise}$ and $m_2 \triangleq M + 2\sigma_{noise}$. From (A.6) [(A.5) gives the same result], (A.9) and (A.10), we see that the secondary membership function corresponding to a domain point x is

$$\mu_2(x,\mu_1) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x+\Sigma\sqrt{-2\ln(\mu_1)}-M}{\sigma_{noise}}\right)^2} & ; \ \mu_1 \in [\mu_1^2, \mu_1^1] \\ \max\left\{e^{-\frac{1}{2}\left(\frac{x+\Sigma\sqrt{-2\ln(\mu_1)}-M}{\sigma_{noise}}\right)^2}, e^{-\frac{1}{2}\left(\frac{x-\Sigma\sqrt{-2\ln(\mu_1)}-M}{\sigma_{noise}}\right)^2}\right\} & ; \ \mu_1 \in [\mu_1^1, 1] \\ 0 & ; \ \text{otherwise} \end{cases} \tag{6.32}$$

where

$$\mu_1^1 = e^{-\frac{1}{2}\left(\frac{x-m_1}{\sigma_{noise}}\right)^2}, \tag{6.33}$$

$$\mu_1^2 = e^{-\frac{1}{2}\left(\frac{x-m_2}{\sigma_{noise}}\right)^2} \tag{6.34}$$

and, the Gaussian type-1 membership function, with mean M and standard deviation σ_{noise}, for the mean of the Gaussian type-1 set is assumed to be contained in $[m_1, m_2]$, where $[m_1, m_2]$ is the range of uncertainty of the mean. From (6.32), (6.33) and (6.34), we find that the secondary membership function corresponding to $x = M = (m_1 + m_2)/2$ is

$$\mu_2(M,\mu_1) = \begin{cases} \mu_1^{(\Sigma^2/\sigma_{noise}^2)} & ; \ \mu_1 \in \left[e^{-\frac{1}{2}\left(\frac{m_2-m_1}{2\sigma_{noise}}\right)^2}, 1\right] \\ 0 & ; \ \text{otherwise} \end{cases} \tag{6.35}$$

Observe, from (6.35), that

$$\mu_2(M, 1) = 1 \tag{6.36}$$

Next, we approximate $\mu_2(M,\mu_1)$ with a Gaussian. To do this, we find the value of μ_1, μ_1', where $\mu_2(M,\mu_1') = \exp\{-1/2\}$ (corresponding to a point one standard deviation away from the mean of a Gaussian). This value can be obtained as

$$(\mu_1')^{\left(\frac{\Sigma}{\sigma_{noise}}\right)^2} = e^{-\frac{1}{2}} \quad \Longleftrightarrow \quad \mu_1' = e^{-\frac{1}{2}\left(\frac{\sigma_{noise}}{\Sigma}\right)^2} \tag{6.37}$$

Using (6.36) and (6.37), we can roughly approximate $\mu_2(M,\mu_1)$ with a Gaussian having mean 1 and standard deviation $(1 - \mu_1') = (1 - \exp\{-(1/2)(\sigma_{noise}/\Sigma)^2\})$; therefore, we choose the parameter c for our Gaussian type-2 sets to be

$$c = c_0\left(1 - e^{-\frac{1}{2}\left(\frac{\sigma_{noise}}{\Sigma}\right)^2}\right) \tag{6.38}$$

where c_0 is some constant, which can be determined experimentally. This choice seems reasonable. If there is no noise, noise standard deviation $\sigma_{noise} = 0$, and correspondingly $c = 0$, indicating that there is no type-2 uncertainty required; and, when σ_{noise} increases, c grows larger.

The expression for the centroid of this type of a Gaussian type-2 set can be obtained (approximately) as shown in Example 2.8. We restate the result here.

Consider a Gaussian type-2 set $\tilde{\tilde{A}} \subset X$. If the principal membership function of $\tilde{\tilde{A}}$ is a Gaussian type-1 set with mean M and standard deviation Σ, and, if for every $x \in X$, the standard deviation, $\sigma(x)$, of the membership grade is $\sigma(x) = cm(x)$, where c is a constant, then the centroid of $\tilde{\tilde{A}}$ is approximately a Gaussian type-1 set with mean M and standard deviation $\sqrt{2/\pi}c\Sigma$, provided that $c \ll 1/k$, where k is the number of standard deviations of a Gaussian considered significant.

In order for this centroid approximation to be good, we require that $c \ll 0.5$ (assuming $k = 2$). Condition (2.78) on the Gaussian approximation also requires that c be small; therefore, we can use this approach only for those cases where this condition on c is satisfied.

We assume that $c_0 = 1.5$ for all the type-2 sets in this example. (Note, however, that in practice, c_0 need not be the same for all the fuzzy sets.)

6.2.2.3 Results

As mentioned earlier, since this approach can be used only for high SNRs, we just show one example (see Fig. 6.25), where the training data is corrupted with zero mean additive Gaussian noise, and the SNR is $20dB$. We use center-of-sets type-reduction and the Gaussian approximation (see Section 5.7) to compute the type-reduced set. In this case $c = 0.1619$ [calculated using (6.38), with $c_0 = 1.5$]. The MSE of the type-1 FLS output with the true time-series is 1.94×10^{-3}. The crisp output of the type-2 FLS is the same as the output of the type-1 FLS (see next paragraph); and, the bound error (see Section 6.2.1.3) of the type-2 FLS is 5.45×10^{-3}.

Recall that the crisp output of a Gaussian type-2 FLS using the Gaussian approximation, is the same as the output of its principal type-1 FLS (see Section 5.8). Observe, also, that in this example, the type-1 FLS trained with noisy data is the principal type-1 FLS of the Gaussian type-2 FLS; therefore, the crisp output of the type-2 FLS in this case, is identical to the output of the type-1 FLS. If one is interested only in the crisp output of the FLS, the actual type-reduced set of this type-2 FLS should be computed numerically. Type-reduction, however, is computationally very expensive and at present, we have no fast algorithms to perform this operation for Gaussian type-2 sets .

6.2.3 A Comparison of the Two Approaches

In Section 6.2.1, we showed that an interval type-2 FLS can be used to obtain bounds on the output, as well as a better crisp prediction, by tuning its parameters. This was possible because we have an efficient algorithm for type-reduction computations of interval type-2 FLS's.

The Gaussian type-2 FLS in Section 6.2.2 also gave bounds on the output; but, because we used the Gaussian approximation, the crisp output of this type-2 FLS is the same as the type-1 FLS output. Also, since the approximation can only be used when the amount of type-2 uncertainty is small, we can only use Gaussian type-2 sets for very high SNRs. So, at this point, we can recommend using a Gaussian type-2 FLS only for very high SNRs and if one is interested only in obtaining bounds on the output; however, if fast algorithms for Gaussian type-reduction can be developed, a Gaussian type-2 FLS can then also be used for lower SNRs.

Note that, although we used uniform noise while working with interval type-2 sets, and Gaussian noise while working with Gaussian type-2 sets, this need not necessarily be so. In practice, we rarely, if ever, know the pdf of noise, so that either approach can be used for any kind of noise.

Note, also, that, in this example, we first designed a type-1 FLS and then created a type-2 FLS from it by using information available about the noise in the training data. There can, however,

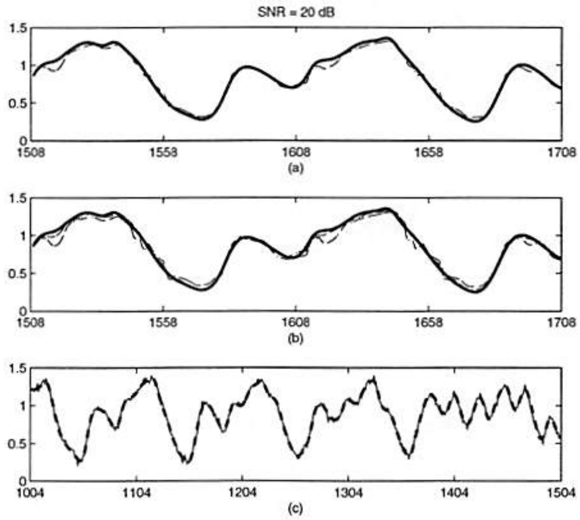


Figure 6.25: Mackey-Glass chaotic time-series prediction in the presence of Gaussian noise, when SNR $= 20dB$. The type-2 FLS uses Gaussian type-2 sets. The thick solid line in (a) and (b) indicates the true time-series. In (a), the dashed line indicates the type-1 output and the dash-dotted line indicates the type-2 crisp output. Figure (b) shows the upper and lower bands obtained with the type-2 FLS. In Fig. (c), the solid line shows the noisy data used for training and the thick dashed line shows the noisefree data. Training is performed with the 500 input-output pairs in $x(1001), x(1002), \ldots, x(1504)$ and testing is done with 200 input-output pairs in $x(1505), x(1506), \ldots, x(1708)$.

be other design approaches. For example, we could start with a type-2 system, and tune it's parameters directly using the training data. We leave these other design possibilities as directions for future research.

Chapter 7

Conclusions

7.1 Completed Work

In this report, we have developed a type-2 fuzzy logic system. Our work can be broadly divided into three parts :

1. Casting the nuts and bolts
Since the results in the existing type-2 FL literature were far from sufficient for our work, we started from the very basic operations of unions, intersections and complements of type-2 sets, and developed the results that we would need in order to implement a type-2 FLS. In all the cases, we considered two t-norms : minimum and product. When it was not possible to give simple results for general membership functions, we concentrated on Gaussian and interval type-2 sets and provided results that simplify working with these two kinds of membership functions (see [8] for results for triangular type-2 sets). We studied, in detail, the following topics.
(a) Set-theoretic operations on type-2 sets : We examined the operations of unions, intersections and complements of type-2 sets in great detail, and, developed easily implementable algorithms for performing these operations. When actual results were difficult to generalize, we developed practical approximations.
(b) Algebraic operations on type-1 and type-2 sets : We examined the algebraic operations of addition and multiplication on the membership grades of type-2 sets (which themselves are type-1 sets). We also defined the "centroid of a type-2 set" using Zadeh's Extension Principle, and studied aspects related to its computation.
(c) Properties of membership grades of type-2 sets : We thoroughly investigated properties of membership grades of type-2 sets, e.g., *commutativity* and *associativity*.
(d) Type-2 relations and their compositions : We showed the validity of the "extended" sup-star composition in the type-2 case and studied type-2 fuzzy relations and their compositions.

2. Putting the system together
Using the results in **(1)** above, we implemented the operations of inference, type-reduction and defuzzification in a type-2 FLS. "Type-reduction" is a term that we have coined for an extended version of a type-1 defuzzification operation. We studied the following important type-reduction methods (corresponding to different type-1 defuzzification methods) in detail : centroid, center-of-sums, height, modified height and center-of-sets. We introduced "center-of-sets" type-reduction to overcome some shortcomings of height type-reduction. In general, type-reduction tends to be computationally quite expensive; therefore, we provided approximations that considerably simplify type-reduction for Gaussian, and interval type-2 FLS's (see [8] for triangular type-2 FLS's). For interval type-2 FLS's, the exact type-reduced result can be obtained relatively easily by using a computational procedure that we have developed.

3. Showing that it works
We demonstrated the use of a type-2 FLS and its advantages over a type-1 FLS by considering two real applications : managing rules collected by means of a survey, and time-series prediction. In the survey example, we showed how the linguistic uncertainty about membership functions of the FLS, as well as rule uncertainty from multiple experts, each of whom may give different answers to the same question, can be handled in the type-2 framework. In the time-series example, we considered a type-1 FLS trained with noisy data and showed how information about the noise in the training data can be incorporated in a type-2 FLS to obtain bounds on the output and also better predictions.

7.2 Future Research Directions

Many directions for future research are possible, both in theory and applications of type-2 FLS's. We give a few examples next.

Theory of type-2 FLS's
1. All of our work dealt with minimum and product *t*-norms; and, most of it concentrated on Gaussian, interval and triangular membership functions. Similar results can be developed for other choices of membership functions, *t*-norms or other parameters.
2. We have developed efficient computational procedures that can obtain actual (i.e., not approximate) results for interval type-2 FLS's. It may be possible to develop similar procedures and/or analytical results for Gaussian, triangular or other type-2 membership functions.
3. We considered only singleton fuzzification in our work. A "Non-singleton" type-2 FLS can be developed, possibly using the extended sup-star composition results that we have provided.

Applications of type-2 FLS's
1. Design and training procedures for type-2 FLS's (i.e., procedures for choosing and tuning the parameters of the FLS's from available data) need to be developed.
2. New areas of application, where the additional information provided by a type-2 FLS (e.g., the type-reduced set) will be useful, need to be explored.

7.3 Final Conclusion

A type-2 FLS takes into account rule uncertainties in a FLS. Just as variance provides a measure of dispersion about the mean in statistical-based designs, the type-reduced set of a type-2 FLS provides a measure of dispersion about the crisp output of the FLS. We believe that this additional information provided by a type-2 FLS makes it indispensable as a modeling tool in the presence of linguistic and/or numerical uncertainties.

Appendix A

Examples for Chapter 1

In this Appendix, we provide 3-D representations for Examples 1.1 and 1.2, assuming that the standard deviation of the Gaussian in Example 1.1 and the mean of the Gaussian in Example 1.2 are Gaussian type-1 sets. We shall see that obtaining the 3-D representations in this case is fairly more complicated than obtaining the ones in Figs. 1.4 and 1.5, where the uncertain standard deviation and mean were assumed to be crisp sets.

Example A.1 Consider Example 1.1. Suppose that the standard deviation of this Gaussian is a type-1 fuzzy set with domain $[\sigma_1, \sigma_2]$ that is characterized by a Gaussian membership function with mean $M_\sigma = \frac{\sigma_1+\sigma_2}{2}$ and standard deviation $\Sigma_\sigma = \frac{\sigma_2-\sigma_1}{4}$. These values for M_σ and Σ_σ were chosen for illustration purposes only. The membership grade for each x still has the same domain as it had when all the values of the standard deviation were equally uncertain, but now, we assign secondary memberships as follows. For any x (e.g., $x = 0.65$ in Fig. 1.1), if a primary membership $\mu_1 \in [0,1]$ is such that $\mu_1 = \exp\{-\frac{1}{2}\left(\frac{x-m}{\sigma'}\right)^2\}$ for some $\sigma' \in [\sigma_1, \sigma_2]$ ($\sigma_1 = 0.1$ and $\sigma_2 = 0.2$ in Fig. 1.1), then we set the secondary membership corresponding to this x and μ_1, $\mu_2(x, \mu_1)$, equal to the membership of σ' in the fuzzy set σ, i.e., we set

$$\mu_2(x,\mu_1) = e^{-\frac{1}{2}\left(\frac{\sigma'-M_\sigma}{\Sigma_\sigma}\right)^2} \quad \text{where} \quad \mu_1 = e^{-\frac{1}{2}\left(\frac{x-m}{\sigma'}\right)^2} \tag{A.1}$$

In Fig. 1.1, for $x = 0.65$, this occurs for $\mu_1 \in [0.3247, 0.7548]$. If a primary membership $\mu_1 \in [0,1]$ is such that no $\sigma' \in [\sigma_1, \sigma_2]$ satisfies $\mu_1 = \exp\{-\frac{1}{2}\left(\frac{x-m}{\sigma'}\right)^2\}$, we set $\mu_2(x,\mu_1) = 0$. In Fig. 1.1, for $x = 0.65$, this occurs for $\mu_1 \notin [0.3247, 0.7548]$. Note that the above choice of $\mu_2(x, \mu_1)$ was quite arbitrary. One may choose $\mu_2(x, \mu_1)$ to be any suitable function of σ'.

Note that $\mu_1 = \exp\left\{-\frac{1}{2}\left(\frac{x-m}{\sigma'}\right)^2\right\} \Rightarrow \sigma' = |x-m|/\sqrt{-2\ln(\mu_1)}$, where we have made use of the fact that σ', being the standard deviation of a Gaussian, is positive. Consequently, we can rewrite (A.1) explicitly in terms of x and μ_1 as follows : When $x \neq m$,

$$\mu_2(x,\mu_1) = \begin{cases} \exp\left\{-\frac{1}{2}\left(\frac{\frac{|x-m|}{\sqrt{-2\ln(\mu_1)}} - M_\sigma}{\Sigma_\sigma}\right)^2\right\} & ; \ \mu_1 \in \left[\exp\left\{-\frac{1}{2}\left(\frac{x-m}{\sigma_1}\right)^2\right\}, \right. \\ & \qquad \left. \exp\left\{-\frac{1}{2}\left(\frac{x-m}{\sigma_2}\right)^2\right\}\right] \\ 0 & ; \ \text{otherwise} \end{cases} \tag{A.2}$$

When $x = m$ ($m = 0.5$ in Fig. 1.1), every $\sigma' \in [\sigma_1, \sigma_2]$ gives $\mu_1 = \exp\{-\frac{1}{2}\left(\frac{x-m}{\sigma'}\right)^2\} = 1$. In this case, we set σ' equal to that value of σ which maximizes $\mu_2(x,\mu_1)|_{(m,1)}$, i.e., we set $\sigma' = M_\sigma$; consequently, $\mu_2(m, 1) = 1$ and $\mu_2(m, \mu_1) = 0$ for $\mu_1 \neq 1$.

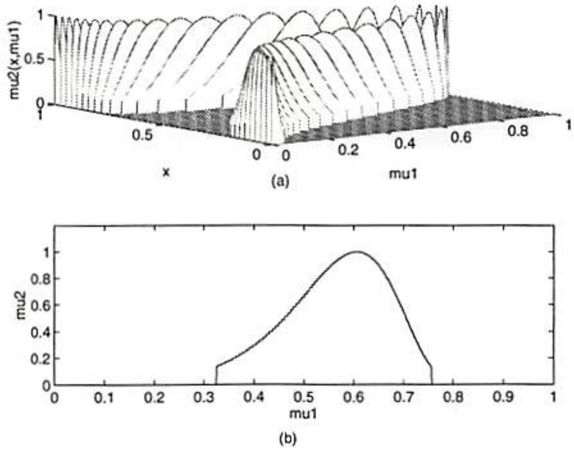


Figure A.1: (a) Three dimensional representation of the type-2 set in Example A.1, assuming that the standard deviation is a Gaussian type-1 set with mean $\frac{\sigma_1+\sigma_2}{2} = 0.15$ and standard deviation $\frac{\sigma_2-\sigma_1}{4} = 0.025$, contained in $[\sigma_1, \sigma_2] = [0.1, 0.2]$. (b) The membership grade corresponding to $x = 0.65$.

The membership grade in (A.2) is depicted in Fig. A.1 (b). Figure A.1 (a) shows a 3-D representation of this type-2 set. Observe, from (A.2), that the membership grade corresponding to any x (i.e., $\mu_1 - \mu_2$ plot for a fixed x) is generally non-Gaussian. Each of the slices in the 3-D plot was constructed by evaluating (A.2) for different values of x. □

Example A.2 Consider Example 1.2. Suppose that the mean of this Gaussian is a type-1 fuzzy set with domain $[m_1, m_2]$ that is characterized by a Gaussian membership function with mean $M_m = \frac{m_1+m_2}{2}$ and standard deviation $\Sigma_m = \frac{m_2-m_1}{4}$. Figure A.2 (a) shows 3-D diagrams for Example 1.2, when $m_1 = 0.4$ and $m_2 = 0.6$. The secondary memberships are computed as follows.

For any x (e.g., $x = 0.65$ in Fig. 1.2), if a primary membership $\mu_1 \in [0,1]$ is such that $\mu_1 = \exp\{-\frac{1}{2}\left(\frac{x-m'}{\sigma}\right)^2\}$ for some $m' \in [m_1, m_2]$ ($m_1 = 0.4$ and $m_2 = 0.6$ in Fig. 1.2), then the corresponding secondary membership $\mu_2(x, \mu_1)$ is set equal to the membership of m' in the type-1 fuzzy set m, i.e., we set

$$\mu_2(x, \mu_1) = e^{-\frac{1}{2}\left(\frac{m'-M_m}{\Sigma_m}\right)^2} \quad \text{where} \quad \mu_1 = e^{-\frac{1}{2}\left(\frac{x-m'}{\sigma}\right)^2} \tag{A.3}$$

In Fig. 1.2, for $x = 0.65$, this occurs for $\mu_1 \in [0.4578, 0.9692]$. If μ_1 is such that no $m \in [m_1, m_2]$ satisfies $\mu_1 = \exp\{-\frac{1}{2}\left(\frac{x-m'}{\sigma}\right)^2\}$, we set $\mu_2(x, \mu_1) = 0$. In Fig. 1.2, for $x = 0.65$, this occurs for $\mu_1 \notin [0.4578, 0.9692]$. Observe also, from Fig. 1.2, that in the interval $[m_1, m_2]$, there may be more than one value of m' which satisfies (A.3). In this case, we choose that value for m' which maximizes $\mu_2(x, \mu_1)$. In Fig. 1.2, this occurs for $x \in [0.4, 0.6]$.

Note that $\mu_1 = \exp\left\{-\frac{1}{2}\left(\frac{x-m'}{\sigma}\right)^2\right\}$ implies that

$$m' = \begin{cases} x + \sigma\sqrt{-2\ln(\mu_1)} & ; \quad x < m_1 \\ x \pm \sigma\sqrt{-2\ln(\mu_1)} & ; \quad m_1 \le x \le m_2 \\ x - \sigma\sqrt{-2\ln(\mu_1)} & ; \quad x > m_2 \end{cases} \tag{A.4}$$

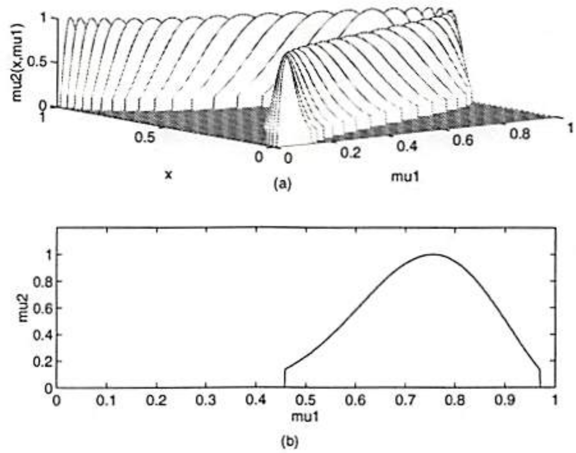


Figure A.2: (a) Three dimensional representation of the type-2 set in Example A.2, assuming that the mean is a Gaussian type-1 set with mean $\frac{m_1+m_2}{2} = 0.5$ and standard deviation $\frac{m_2-m_1}{4} = 0.05$, contained in $[m_1, m_2] = [0.4, 0.6]$. (b) The membership grade corresponding to $x = 0.65$.

Consequently, using (A.4), (A.3) can be rewritten as follows (see Fig. 1.2) : For $x < m_1$,

$$\mu_2(x,\mu_1) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x+\sigma\sqrt{-2\ln(\mu_1)}-M_m}{\Sigma_m}\right)^2} & ; \ \mu_1 \in [\mu_1^2, \mu_1^1] \\ 0 & ; \ \text{otherwise} \end{cases} \tag{A.5}$$

For $m_1 \le x \le (m_1+m_2)/2$,

$$\mu_2(x,\mu_1) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x+\sigma\sqrt{-2\ln(\mu_1)}-M_m}{\Sigma_m}\right)^2} & ; \ \mu_1 \in [\mu_1^2, \mu_1^1] \\ \max\left\{ e^{-\frac{1}{2}\left(\frac{x+\sigma\sqrt{-2\ln(\mu_1)}-M_m}{\Sigma_m}\right)^2}, e^{-\frac{1}{2}\left(\frac{x-\sigma\sqrt{-2\ln(\mu_1)}-M_m}{\Sigma_m}\right)^2} \right\} & ; \ \mu_1 \in [\mu_1^1, 1] \\ 0 & ; \ \text{otherwise} \end{cases} \tag{A.6}$$

For $(m_1+m_2)/2 \le x \le m_2$,

$$\mu_2(x,\mu_1) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x-\sigma\sqrt{-2\ln(\mu_1)}-M_m}{\Sigma_m}\right)^2} & ; \ \mu_1 \in [\mu_1^1, \mu_1^2] \\ \max\left\{ e^{-\frac{1}{2}\left(\frac{x+\sigma\sqrt{-2\ln(\mu_1)}-M_m}{\Sigma_m}\right)^2}, e^{-\frac{1}{2}\left(\frac{x-\sigma\sqrt{-2\ln(\mu_1)}-M_m}{\Sigma_m}\right)^2} \right\} & ; \ \mu_1 \in [\mu_1^2, 1] \\ 0 & ; \ \text{otherwise} \end{cases} \tag{A.7}$$

For $x > m_2$,

$$\mu_2(x,\mu_1) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x-\sigma\sqrt{-2\ln(\mu_1)}-M_m}{\Sigma_m}\right)^2} & ; \ \mu_1 \in [\mu_1^1, \mu_1^2] \\ 0 & ; \ \text{otherwise} \end{cases} \tag{A.8}$$

where

$$\mu_1^1 = e^{-\frac{1}{2}\left(\frac{x-m_1}{\sigma}\right)^2} \tag{A.9}$$

and

$$\mu_1^2 = e^{-\frac{1}{2}\left(\frac{x-m_2}{\sigma}\right)^2} \tag{A.10}$$

Figure A.2 (b) shows the membership grade corresponding to $x = 0.65$. Observe, from (A.5) - (A.8), that the membership grade corresponding to any x is generally non-Gaussian. Each of the slices in the 3-D plot was constructed by evaluating (A.5) - (A.8) for different values of x. □

Appendix B

A Note on the Extension Principle

The Extension Principle [30] allows the domain of definition of a mapping or a relation to be extended from points in U to fuzzy subsets of U. If f is a mapping from U to V and $\tilde{A}$ is a fuzzy subset of U, such that

$$\tilde{A} = \sum_{i=1}^{n} \mu_i / u_i , \tag{B.1}$$

then

$$f(\tilde{A}) = \sum_{i=1}^{n} \mu_i / f(u_i) \tag{B.2}$$

If f is a mapping from a Cartesian product $U_1 \times U_2 \times \cdots \times U_n$ to V and if $\tilde{A}$ is a fuzzy set (relation) in $U_1 \times U_2 \times \cdots \times U_n$ characterized by the membership function $\mu_{\tilde{A}}(u_1, \ldots, u_n)$, where $u_i \in U_i$, then

$$f(\tilde{A}) = \int_V \mu_{\tilde{A}}(u_1, \ldots, u_n) / f(u_1, \ldots, u_n) \tag{B.3}$$

Many times, we don't know $\tilde{A}$, but instead only know projections of $\tilde{A}$, $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n$ on $U_1, U_2, \ldots, U_n$, respectively. If $\tilde{A} = \tilde{A}_1 \times \cdots \times \tilde{A}_n$, we can use the following expression for $\mu_{\tilde{A}}(u_1, \ldots, u_n)$ [Zadeh uses only the minimum t-norm]

$$\mu_{\tilde{A}}(u_1, \ldots, u_n) = \mu_{\tilde{A}_1}(u_1) \star \mu_{\tilde{A}_2}(u_2) \star \cdots \star \mu_{\tilde{A}_n}(u_n) \tag{B.4}$$

Let us consider the case $n = 2$, for which $*$ is a binary operation defined on $U \times V$ with values in W, i.e., if $u \in U$ and $v \in V$, then $w = u * v \in W$. Now, if $\tilde{A} = \sum_{i=1}^{n} \mu_i / u_i$ and $\tilde{B} = \sum_{j=1}^{m} \nu_j / v_j$ are fuzzy sets in U and V, respectively, then

$$\begin{aligned} \tilde{A} * \tilde{B} &= \Big(\sum_{i=1}^{n} \mu_i / u_i\Big) * \Big(\sum_{j=1}^{m} \nu_j / v_j\Big) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} (\mu_i \star \nu_j) \Big/ (u_i * v_j) \end{aligned} \tag{B.5}$$

The validity of (B.5) depends on the assumption that u_i and v_j are "non-interactive", or that there is no constraint on (u_i, v_j) [we can think of u_i and v_j as being "independent" in some sense]. If there is a constraint on (u, v), which is expressed as a relation $\tilde{\text{R}}$ with a membership function $\mu_{\tilde{\text{R}}}$, then the expression for $\tilde{\text{A}} * \tilde{\text{B}}$ should be written as

$$\begin{aligned}\tilde{\text{A}} * \tilde{\text{B}} &= \Big(\sum_{i=1}^{n} \mu_i/u_i\Big) * \Big(\sum_{j=1}^{m} \nu_j/v_j\Big) \cap \tilde{\text{R}} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m}[\mu_i \star \nu_j \star \mu_{\tilde{\text{R}}}(u_i, v_j)]\Big/(u_i * v_j)\end{aligned} \tag{B.6}$$

If $\tilde{\text{R}}$ is a crisp relation [i.e., if the constraint on (u_i, v_j) is expressible as a crisp relation R], then the right-hand side of (B.6) will contain only those terms which satisfy the constraint.

Example B.1 Let $U = 1 + \cdots + 10$ and let $\tilde{\text{A}}$ be a fuzzy subset of U defined as

$$\tilde{\text{A}} = 1/1 + 0.6/4 + 0.4/5 \tag{B.7}$$

$\tilde{\text{A}}^2$ can be found in two ways. If we take f as the operation of squaring, then using (B.2),

$$\tilde{\text{A}}^2 = 1/1 + 0.6/16 + 0.4/25 \tag{B.8}$$

If we write $\tilde{\text{A}}^2$ as $\tilde{\text{A}} \times \tilde{\text{A}}$, (B.5) gives us (assuming minimum t-norm)

$$\begin{aligned}\tilde{\text{A}} \times \tilde{\text{A}} &= (1/1 + 0.6/4 + 0.4/5) \times (1/1 + 0.6/4 + 0.4/5) \\ &= (1/1) \times (1/1 + 0.6/4 + 0.4/5) + (0.6/4) \times (1/1 + 0.6/4 + 0.4/5) \\ &\quad +(0.4/5) \times (1/1 + 0.6/4 + 0.4/5) \\ &= 1/1 + 0.6/4 + 0.4/5 + 0.6/4 + 0.6/16 \\ &\quad +0.4/20 + 0.4/5 + 0.4/20 + 0.4/25 \\ &= 1/1 + 0.6/4 + 0.4/5 + 0.6/16 + 0.4/20 + 0.4/25\end{aligned} \tag{B.9}$$

From (B.8) and (B.9), we see that using (B.5), $\tilde{\text{A}}^2 \neq \tilde{\text{A}} \times \tilde{\text{A}}$! This happened, because we did not use the right form of the Extension Principle. In order to get $\tilde{\text{A}}^2 = \tilde{\text{A}} \times \tilde{\text{A}}$, we have to use the restricted form of the Extension Principle [i.e., (B.6)] to evaluate $\tilde{\text{A}} \times \tilde{\text{A}}$. The restriction is crisp in this case and can be expressed as

$$\mu_R(u_i, v_j) = \begin{cases} 1 \quad ; & u_i = v_j \\ 0 \quad ; & \text{otherwise} \end{cases} \tag{B.10}$$

Using (B.6) with (B.10), we get

$$\begin{aligned}\tilde{\text{A}} \times \tilde{\text{A}} &= (1/1 + 0.6/4 + 0.4/5) \times (1/1 + 0.6/4 + 0.4/5) \cap R \\ &= [(1/1) \times (1/1 + 0.6/4 + 0.4/5) + (0.6/4) \times (1/1 + 0.6/4 + 0.4/5) \\ &\quad +(0.4/5) \times (1/1 + 0.6/4 + 0.4/5)] \cap R \\ &= [1 \wedge 1 \wedge \mu_R(1,1)]/(1 \times 1) + [1 \wedge 0.6 \wedge \mu_R(1,4)]/(1 \times 4) \\ &\quad +[1 \wedge 0.4 \wedge \mu_R(1,5)]/(1 \times 5) + [0.6 \wedge 1 \wedge \mu_R(4,1)]/(4 \times 1) \\ &\quad +[0.6 \wedge 0.6 \wedge \mu_R(4,4)]/(4 \times 4) + [0.6 \wedge 0.4 \wedge \mu_R(4,5)]/(4 \times 5) \\ &\quad +[0.4 \wedge 1 \wedge \mu_R(5,1)]/(5 \times 1) + [0.4 \wedge 0.6 \wedge \mu_R(5,4)]/(5 \times 4) \\ &\quad +[0.4 \wedge 0.4 \wedge \mu_R(5,5)]/(5 \times 5) \\ &= 1/1 + 0/4 + 0/5 + 0/4 + 0.6/16 + 0/20 + 0/5 + 0/20 + 0.4/25\end{aligned}$$

$$
\begin{aligned}
&= 1/1 + 0.6/16 + 0.4/25 \\
&= \tilde{A}^2 \qquad \text{(B.11)}
\end{aligned}
$$

□

Observe that using the restricted form of the Extension Principle can complicate computations quite a lot. There are also a few problems/difficulties associated with using the restricted form of the Extension Principle.

1. If we have to perfom an operation between two given fuzzy sets $\tilde{A}$ and $\tilde{B}$, it may not always be easy to define a restriction between the two. For example, if we have to find $\tilde{A} \sqcup \tilde{B}$, where $\tilde{A} = 1/1 + 0.6/4 + 0.4/5$ and $\tilde{B} = 1/1 + 0.7/3$, without any other information about $\tilde{A}$ and $\tilde{B}$, there is no fixed way of defining a relation between $\tilde{A}$ and $\tilde{B}$; in fact, it may not even be possible to tell if $\tilde{A}$ and $\tilde{B}$ are related at all.

2. When performing operations like $\tilde{A} \sqcap \tilde{B} \sqcup \tilde{A}$, if we are given a restriction on $\tilde{A}$ and $\tilde{B}$, it may be easy to use the restricted from of the Extension Principle; however, if we are given $\tilde{A} \sqcap \tilde{B}$ and $\tilde{A}$, it may not be easy (or it may not even be possible) to define a restriction on the elements of $\tilde{A} \sqcap \tilde{B}$ and $\tilde{A}$ so as to use the restricted form.

3. When we use the product t-norm, even using the restricted form of the Extension Principle may not give us equalities in cases like the one considered in Example B.1, as demonstrated in the following example :

   Example B.2 Consider the same type-1 set $\tilde{A} \subset U$ considered in Example B.1. Computing $\tilde{A}^2$ by considering squaring as an operation on $\tilde{A}$ gives us the same $\tilde{A}^2$ as in (B.8); however, computing $\tilde{A} \times \tilde{A}$ using the product t-norm, gives us [where we use the same restriction R as in (B.10)]

$$
\begin{aligned}
\tilde{A} \times \tilde{A} &= (1/1 + 0.6/4 + 0.4/5) \times (1/1 + 0.6/4 + 0.4/5) \cap R \\
&= [(1/1) \times (1/1 + 0.6/4 + 0.4/5) + (0.6/4) \times (1/1 + 0.6/4 + 0.4/5) \\
&\quad +(0.4/5) \times (1/1 + 0.6/4 + 0.4/5)] \cap R \\
&= [1 \times 1 \times \mu_R(1,1)]/(1 \times 1) + [1 \times 0.6 \times \mu_R(1,4)]/(1 \times 4) \\
&\quad +[1 \times 0.4 \times \mu_R(1,5)]/(1 \times 5) + [0.6 \times 1 \times \mu_R(4,1)]/(4 \times 1) \\
&\quad +[0.6 \times 0.6 \times \mu_R(4,4)]/(4 \times 4) + [0.6 \times 0.4 \times \mu_R(4,5)]/(4 \times 5) \\
&\quad +[0.4 \times 1 \times \mu_R(5,1)]/(5 \times 1) + [0.4 \times 0.6 \times \mu_R(5,4)]/(5 \times 4) \\
&\quad +[0.4 \times 0.4 \times \mu_R(5,5)]/(5 \times 5) \\
&= 1/1 + 0/4 + 0/5 + 0/4 + 0.36/16 + 0/20 + 0/5 + 0/20 + 0.16/25 \\
&= 1/1 + 0.36/16 + 0.16/25 \\
&\neq \tilde{A}^2 \qquad \text{(B.12)}
\end{aligned}
$$

   □

Though, as shown in Example B.2, product t-norm does not give intuitive results with the restricted form of the Extension Principle, in view of the desirable properties of the product t-norm, we continue to use it in our work. In order to avoid the above mentioned difficulties with the restricted form of the Extension Principle, we adopt the following approach.

When we need to extend an operation of the form $f(\theta_1, \ldots, \theta_n)$ to an operation $f(\tilde{A}_1, \ldots, \tilde{A}_n)$, we will not extend the individual operations, like multiplication, addition, etc. involved in f; rather, we will use the following definition :

$$f(\tilde{A}_1, \ldots, \tilde{A}_n) = \int_{\theta_1} \cdots \int_{\theta_n} \mu_{\tilde{A}_1}(\theta_1) \star \cdots \star \mu_{\tilde{A}_n}(\theta_n) \Big/ f(\theta_1, \ldots, \theta_n) \tag{B.13}$$

where $\theta_i \in \tilde{A}_i$ for $i = 1, \ldots, n$. For example, if $f(\theta_1, \theta_2) = [\theta_1\theta_2] \big/ [\theta_1 + \theta_2]$, we write the extension of f to type-1 sets $\tilde{A}_1$ and $\tilde{A}_2$ as

$$f(\tilde{A}_1, \tilde{A}_2) = \int_{\theta_1} \int_{\theta_2} \mu_{\tilde{A}_1}(\theta_1) \star \mu_{\tilde{A}_2}(\theta_2) \Big/ \frac{\theta_1\theta_2}{\theta_1 + \theta_2} \tag{B.14}$$

where $\theta_i \in \tilde{A}_i$ for $i = 1, 2$; and **not** as

$$f(\tilde{A}_1, \tilde{A}_2) = \frac{\tilde{A}_1 \times \tilde{A}_2}{\tilde{A}_1 + \tilde{A}_2} \tag{B.15}$$

When discussing properties of membership grades in Chapter 3, however, we use the *unrestricted* form of the Extension Principle, just like Mizumoto and Tanaka do in [17].

Appendix C

Proofs in Chapter 2

C.1 Proof of Theorem 2.1

In the proof of Theorem 2.1, given next, we represent fuzzy sets $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$ as follows :

$$\tilde{\mathrm{F}} = \int_{v\in\Re} f(v)/v \quad (C.1)$$

$$\tilde{\mathrm{G}} = \int_{w\in\Re} g(w)/w \quad (C.2)$$

As is apparent from (C.1) and (C.2), fuzzy sets $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$ can, in general, have the real line as their domain. If a real number w_0 is not in $\tilde{\mathrm{F}}$ (or $\tilde{\mathrm{G}}$), $f(w_0)$ [or $g(w_0)$] will be zero. With this understanding, we will sometimes use the notation $v \in \Re$ and $v \in \tilde{\mathrm{F}}$ or $w \in \Re$ and $w \in \tilde{\mathrm{G}}$ interchangeably.

Proof :

(a-I) The *join* operation between $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$ can be expressed, as

$$\tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}} = \int_{v\in\Re}\int_{w\in\Re} [f(v) \wedge g(w)]/(v \vee w) \quad (C.3)$$

Let's see what operations are involved here. For every pair of points $\{v, w\}$, such that $v \in \tilde{\mathrm{F}}$ and $w \in \tilde{\mathrm{G}}$, we find the maximum of v and w and the minimum of their memberships, so that $v \vee w$ is an element of $\tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}}$ and $f(v) \wedge g(w)$ is the corresponding membership grade. If more than one $\{v, w\}$ pair gives the same maximum (i.e., the same element in $\tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}}$), we use the maximum of all the corresponding membership grades as the membership of this element. So, every element of the resulting set is obtained as a result of the *max* operation on one or more $\{v, w\}$ pairs, and it's membership is the maximum of all the results of the *min* operation on memberships of v and w.

We analyze the *join* operation by picking a point in $\tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}}$ and finding it's membership grade. Figures 2.1 (a) and (b) depict an example of the *join* operation. Let $\theta \in \tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}}$. As noted in the preceding paragraph, θ must be the result of the *max* operation on one or more $\{v, w\}$ pairs; hence, the possible admissible pairs can only be $\{v, \theta\}$ where $v \in (-\infty, \theta]$ and $\{\theta, w\}$ where $w \in (-\infty, \theta]$. To find the membership of θ, we have to perform the *min* operation between the memberships of all these possible pairs $\{v, w\}$ and then take the maximum of them. For example, to find the membership grade of the point $\theta = 3$ in the union of $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$, first we compare $g(3)$ with each $f(v)$ for $v \in (-\infty, 3]$ or $v \leq 3$, find the minimum in each of these comparisons and finally find the maximum of all these answers; then we compare $f(3)$ with all $g(w)$ for $w \in (-\infty, 3]$ or $w \leq 3$ and do a similar minimax operation, and finally find the maximum of the results of these two minimax operations.

We break this process into three steps : (1) find the minima between the memberships of all the pairs $\{v,\theta\}$ such that $v \in (-\infty,\theta]$ and then find their supremum; (2) do the same with all the pairs $\{\theta, w\}$ such that $w \in (-\infty, \theta]$; and, (3) find the maximum of the two suprema, i.e.,

$$\mu_{(\tilde{F}\sqcup\tilde{G})}(\theta) = \phi_1(\theta) \vee \phi_2(\theta) \tag{C.4}$$

where,

$$\phi_1(\theta) = \sup_{v\in(-\infty,\theta]} \{f(v) \wedge g(\theta)\} \tag{C.5}$$

and

$$\phi_2(\theta) = \sup_{w\in(-\infty,\theta]} \{f(\theta) \wedge g(w)\} \tag{C.6}$$

In (C.5), $g(\theta)$ is a constant with respect to v, and in (C.6), $f(\theta)$ is a constant with respect to w; therefore,

$$\phi_1(\theta) = g(\theta) \wedge \sup_{v\in(-\infty,\theta]} f(v) \tag{C.7}$$

$$\phi_2(\theta) = f(\theta) \wedge \sup_{w\in(-\infty,\theta]} g(w) \tag{C.8}$$

We break θ into the following three ranges : $\theta < v_0$, $v_0 \leq \theta \leq v_1$ and $\theta > v_1$ (see Fig. 2.1). Recall that $f(v_0) = 1$ and $g(v_1) = 1$ and that $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$ are both convex. Also, observe that *convexity of $\tilde{F}$ is equivalent to the condition that f is monotonic non-decreasing in $(-\infty, v_0]$ and monotonic non-increasing in* $[v_0,\infty)$ (see Appendix C.2). Similarly, *convexity of $\tilde{G}$ is equivalent to the condition that g is monotonic non-decreasing in $(-\infty, v_1]$ and monotonic non-increasing for* $[v_1,\infty)$.

$\underline{\theta = \theta_1 < v_0}$: See Fig. 2.1 (a). Since f and g both are monotonic non-decreasing in $(-\infty, v_0]$,

$$\sup_{v\in(-\infty,\theta]} f(v) = f(\theta), \tag{C.9}$$

and

$$\sup_{w\in(-\infty,\theta]} g(w) = g(\theta)\ ; \tag{C.10}$$

therefore, from (C.7) and (C.8), we have

$$\phi_1(\theta) = \phi_2(\theta) = g(\theta) \wedge f(\theta) \tag{C.11}$$

Using (C.11) in (C.4), we get

$$\mu_{(\tilde{F}\sqcup\tilde{G})}(\theta) = g(\theta) \wedge f(\theta)\ ;\ \theta < v_0 \tag{C.12}$$

$\underline{v_0 \leq \theta = \theta_2 \leq v_1}$: See Fig. 2.1 (a). Recall that $f(v_0) = 1$ and that g is monotonic non-decreasing in $(-\infty, v_1]$; therefore, $\sup_{v\in(-\infty,\theta]} f(v) = 1$ and $\sup_{w\in(-\infty,\theta]} g(w) = g(\theta)$. Using these facts in (C.7) and (C.8), we have that in this range

$$\phi_1(\theta) = g(\theta) \wedge 1 = g(\theta) \tag{C.13}$$

and

$$\phi_2(\theta) = f(\theta) \wedge g(\theta) \tag{C.14}$$

Using (C.13) and (C.14) in (C.4), we have

$$\mu_{(\tilde{F}\sqcup\tilde{G})}(\theta) = g(\theta) \vee [f(\theta) \wedge g(\theta)] \tag{C.15}$$

Observe that, if $f(\theta) \leq g(\theta)$, the RHS of (C.15) simplifies to $g(\theta) \vee [f(\theta)] = g(\theta)$ and if $f(\theta) \geq g(\theta)$, the RHS gives $g(\theta) \vee [g(\theta)] = g(\theta)$. So, in either case

$$\mu_{(F \sqcup G)}(\theta) = g(\theta) \; ; \; v_0 \leq \theta \leq v_1 \tag{C.16}$$

$\underline{\theta = \theta_3 > v_1}$: For θ in this range [see Fig. 2.1 (a)], both f and g have already attained their maximum values; therefore,

$$\sup_{v \in (-\infty, \theta]} f(v) = 1 \tag{C.17}$$

$$\sup_{w \in (-\infty, \theta]} g(w) = 1 \tag{C.18}$$

Consequently,

$$\phi_1(\theta) = g(\theta) \tag{C.19}$$

$$\phi_2(\theta) = f(\theta) \; ; \tag{C.20}$$

therefore, from (C.4),

$$\mu_{(\tilde{F} \sqcup \tilde{G})}(\theta) = f(\theta) \vee g(\theta) \; ; \; \theta > v_1 \tag{C.21}$$

From (C.12), (C.16) and (C.21), we get (2.19).
(a-II) The *meet* operation between $\tilde{F}$ and $\tilde{G}$ can be expressed, as

$$\tilde{F} \sqcap \tilde{G} = \int_{v \in \Re} \int_{w \in \Re} [f(v) \wedge g(w)] / (v \wedge w) \tag{C.22}$$

This equation looks very similar to (C.3). The operations involved here are the same as for the *join* operation, except for the fact that every element of $\tilde{F} \sqcap \tilde{G}$ is obtained as a result of the *min* operation on one or more $\{v, w\}$ pairs, where $v \in \tilde{F}$ and $w \in \tilde{G}$. Consider $\theta \in \tilde{F} \sqcap \tilde{G}$. The possible pairs $\{v, w\}$ that can give us θ as a result of the *min* operation are $\{v, \theta\}$ where $v \in [\theta, \infty)$ and $\{\theta, w\}$ where $w \in [\theta, \infty)$. To find the membership grade of θ, we find the minimum of the memberships for each of these $\{v, w\}$ pairs and then take the maximum of all these results. Again, we break this process into three steps : first we find the minima of the membership grades of all the pairs $\{v, \theta\}$ such that $v \in [\theta, \infty)$ and then find their supremum; then we do the same with all the pairs $\{\theta, w\}$ such that $w \in [\theta, \infty)$; and, finally, we find the maximum of the two suprema, i.e.,

$$\mu_{(\tilde{F} \sqcap \tilde{G})}(\theta) = \phi_3(\theta) \vee \phi_4(\theta) \tag{C.23}$$

where,

$$\phi_3(\theta) = \sup_{v \in [\theta, \infty)} \{f(v) \wedge g(\theta)\} \tag{C.24}$$

$$\phi_4(\theta) = \sup_{w \in [\theta, \infty)} \{f(\theta) \wedge g(w)\} \tag{C.25}$$

Again using similar reasoning as in part **(a-I)** of this proof, we have

$$\phi_3(\theta) = g(\theta) \wedge \sup_{v \in [\theta, \infty)} f(v) \tag{C.26}$$

$$\phi_4(\theta) = f(\theta) \wedge \sup_{w \in [\theta, \infty)} g(w) \tag{C.27}$$

We consider three ranges for θ : $\theta > v_1$, $v_0 \leq \theta \leq v_1$ and $\theta < v_0$.

$\underline{\theta = \theta_3 > v_1}$: See Fig. 2.1 (a). f and g, both, are monotonic non-increasing in (v_1, ∞); therefore,

$$\sup_{v \in [\theta, \infty)} f(v) = f(\theta) \tag{C.28}$$

$$\sup_{w \in [\theta, \infty)} g(w) = g(\theta) \tag{C.29}$$

Using (C.28) and (C.29) in (C.26) and (C.27), we get

$$\phi_3(\theta) = \phi_4(\theta) = f(\theta) \wedge g(\theta) \tag{C.30}$$

Therefore, from (C.23), we have

$$\mu_{(\tilde{F} \sqcap \tilde{G})}(\theta) = f(\theta) \wedge g(\theta) \tag{C.31}$$

$\underline{v_0 \leq \theta = \theta_2 \leq v_1}$: See Fig. 2.1 (a). Recall that f is monotonic non-increasing in $[v_0, \infty)$ and that $g(v_1) = 1$. This gives us

$$\sup_{v \in [\theta, \infty)} f(v) = f(\theta) \tag{C.32}$$

$$\sup_{w \in [\theta, \infty)} g(w) = 1 \tag{C.33}$$

Using (C.32) and (C.33) in (C.26) and (C.27), we have

$$\phi_3(\theta) = g(\theta) \wedge f(\theta) \tag{C.34}$$

$$\phi_4(\theta) = f(\theta) \tag{C.35}$$

Using (C.34) and (C.35) in (C.23), we have

$$\mu_{(\tilde{F} \sqcap \tilde{G})}(\theta) = [g(\theta) \wedge f(\theta)] \vee f(\theta) \tag{C.36}$$

Reasoning as in part **(a-I)** [see Eqs. (C.15) and (C.16)], we get

$$\mu_{(\tilde{F} \sqcap \tilde{G})}(\theta) = f(\theta)\,;\; v_0 \leq \theta \leq v_1 \tag{C.37}$$

$\underline{\theta = \theta_1 < v_0}$: See Fig. 2.1 (a). We have that $f(v_0) = 1$ and $g(v_1) = 1$; therefore,

$$\sup_{v \in [\theta, \infty)} f(v) = 1 \tag{C.38}$$

$$\sup_{w \in [\theta, \infty)} g(w) = 1 \tag{C.39}$$

Using (C.38) and (C.39) in (C.26) and (C.27), we have

$$\phi_3(\theta) = g(\theta) \tag{C.40}$$

$$\phi_4(\theta) = f(\theta)\,; \tag{C.41}$$

therefore, from (C.23), we have

$$\mu_{(\tilde{F} \sqcap \tilde{G})}(\theta) = f(\theta) \vee g(\theta)\,;\; \theta < v_0 \tag{C.42}$$

From (C.31), (C.37) and (C.42), we get (2.20).

(b-I) In [17], Mizumoto and Tanaka show that results of *join* or *meet* operations, using *max* t-conorm and *min* t-norm, on convex and normal type-1 sets are also convex and normal. Using this fact, we generalize the result in part (a) of Theorem 2.1 to more than two sets.

Consider n convex, normal, type-1 fuzzy sets $\tilde{F}_1, \ldots, \tilde{F}_n$ characterized by membership functions $f_1, \ldots, f_n$, respectively. Let $v_1, v_2, \ldots, v_n$ be real numbers such that $v_1 \leq v_2 \leq \cdots \leq v_n$ and $f_1(v_1) = f_2(v_2) = \cdots = f_n(v_n) = 1$.

Using (2.19), we have

$$\mu_{\tilde{F}_{n-1} \sqcup \tilde{F}_n}(\theta) = \begin{cases} f_{n-1}(\theta) \wedge f_n(\theta) & ; \quad \theta < v_{n-1} \\ f_n(\theta) & ; \quad v_{n-1} \leq \theta \leq v_n \\ f_{n-1}(\theta) \vee f_n(\theta) & ; \quad \theta > v_n \end{cases} \tag{C.43}$$

Using the associative property, we have (we are interested mainly in dealing with type-1 sets which are membership grades of type-2 sets; for more discussion on properties of type-1 fuzzy membership grades, see Chapter 3)

$$\tilde{F}_{n-2} \sqcup \tilde{F}_{n-1} \sqcup \tilde{F}_n = \tilde{F}_{n-2} \sqcup (\tilde{F}_{n-1} \sqcup \tilde{F}_n) \tag{C.44}$$

Let $f_{(n-1)n} = \mu_{\tilde{F}_{n-1} \sqcup \tilde{F}_n}$. Since $\tilde{F}_{n-1} \sqcup \tilde{F}_n$ is also a convex, normal, type-1 fuzzy set $[f_n(v_n) = 1$, and from (C.43) we see that $f_{(n-1)n}(v_n) = f_n(v_n)]$, another application of (2.19) gives us

$$\mu_{\tilde{F}_{n-2} \sqcup \tilde{F}_{n-1} \sqcup \tilde{F}_n}(\theta) = \begin{cases} f_{(n-1)n}(\theta) \wedge f_{n-2}(\theta) & ; \quad \theta < v_{n-2} \\ f_{(n-1)n}(\theta) & ; \quad v_{n-2} \leq \theta \leq v_n \\ f_{(n-1)n}(\theta) \vee f_{n-2}(\theta) & ; \quad \theta > v_n \end{cases} \tag{C.45}$$

Since $v_{n-2} \leq v_{n-1}$, (C.43) and (C.45) can be rewritten as follows :

$$\mu_{\tilde{F}_{n-1} \sqcup \tilde{F}_n}(\theta) = \begin{cases} f_{n-1}(\theta) \wedge f_n(\theta) & ; \quad \theta < v_{n-2} \\ f_{n-1}(\theta) \wedge f_n(\theta) & ; \quad v_{n-2} \leq \theta < v_{n-1} \\ f_n(\theta) & ; \quad v_{n-1} \leq \theta \leq v_n \\ f_{n-1}(\theta) \vee f_n(\theta) & ; \quad \theta > v_n \end{cases} \tag{C.46}$$

$$\mu_{\tilde{F}_{n-2} \sqcup \tilde{F}_{n-1} \sqcup \tilde{F}_n}(\theta) = \begin{cases} f_{(n-1)n}(\theta) \wedge f_{n-2}(\theta) & ; \quad \theta < v_{n-2} \\ f_{(n-1)n}(\theta) & ; \quad v_{n-2} \leq \theta < v_{n-1} \\ f_{(n-1)n}(\theta) & ; \quad v_{n-1} \leq \theta \leq v_n \\ f_{(n-1)n}(\theta) \vee f_{n-2}(\theta) & ; \quad \theta > v_n \end{cases} \tag{C.47}$$

Substituting for $f_{(n-1)n}$ into (C.47) from (C.46), we obtain

$$\mu_{\tilde{F}_{n-2} \sqcup \tilde{F}_{n-1} \sqcup \tilde{F}_n}(\theta) = \begin{cases} f_{n-2}(\theta) \wedge f_{n-1}(\theta) \wedge f_n(\theta) & ; \quad \theta < v_{n-2} \\ f_{n-1}(\theta) \wedge f_n(\theta) & ; \quad v_{n-2} \leq \theta \leq v_{n-1} \\ f_n(\theta) & ; \quad v_{n-1} \leq \theta \leq v_n \\ f_{n-2}(\theta) \vee f_{n-1}(\theta) \vee f_n(\theta) & ; \quad \theta > v_n \end{cases} \tag{C.48}$$

Again, $\tilde{F}_{n-2} \sqcup \tilde{F}_{n-1} \sqcup \tilde{F}_n$ is also a convex and normal type-1 set, therefore (2.19) can be applied again. Continuing in this fashion, we get (2.21).
(b-II) The proof is very much similar to that of part (b) - I. Starting with $\tilde{F}_1 \sqcap \tilde{F}_2$ and using (2.20) repeatedly, we get (2.22).

C.2 Proof of Assertion in the Proof of Theorem 2.1

The convexity of $\tilde{F} = \int f(\theta)/\theta$ is equivalent to the condition [17]

$$f(v_2) \geq \min\{f(v_1), f(v_3)\} \tag{C.49}$$

if v_2 is between v_1 and v_3, i.e., if $v_1 \leq v_2 \leq v_3$ or $v_3 \leq v_2 \leq v_1$. (See Figs. 2.1 and 2.3 for examples of arbitrarily shaped convex membership functions.) We first prove that convexity of f implies the monotonicity conditions on f, and then prove that the monotonicity conditions on f imply its convexity.

(I) Since f is a membership function for a normalized type-1 fuzzy set (see Theorem 2.1), we know that $f(v) \leq 1$ for all v. Also, we know that $f(v_0) = 1$; therefore, letting $v_1 = v_0$ in (C.49), we get

$$f(v_2) \geq \min\{1, f(v_3)\} \; ; \; v_0 \leq v_2 \leq v_3 \text{ or } v_3 \leq v_2 \leq v_0 \tag{C.50}$$

i.e.,

$$f(v_2) \geq f(v_3) \; ; \; v_0 \leq v_2 \leq v_3 \text{ or } v_3 \leq v_2 \leq v_0 \tag{C.51}$$

In other words, f is monotonic non-decreasing in $(-\infty, v_0]$ and monotonic non-increasing in $[v_0, \infty)$.

(II) Now, assume that f is monotonic non-decreasing in $(-\infty, v_0]$ and monotonic non-increasing in $[v_0, \infty)$, with $f(v_0) = 1$. Consider two points v_1 and v_3, such that $v_1 \leq v_3$. We will show that any point v_2 between v_1 and v_3, satisifes (C.49). There are three possibilities for v_1 and v_3 :

1. $\underline{v_1 \leq v_3 < v_0}$: Since f is monotonic non-decreasing in this range, for any point v_2 between v_1 and v_3, $f(v_2) \geq f(v_1)$; therefore, (C.49) is satisfied.

2. $\underline{v_1 \leq v_0 \leq v_3}$: Since f is monotonic non-decreasing in $(-\infty, v_0]$, if $v_1 \leq v_2 \leq v_0$, $f(v_2) \geq f(v_1)$. Also, since f is monotonic non-increasing in $[v_0, \infty)$, if $v_0 \leq v_2 \leq v_3$, $f(v_2) \geq f(v_3)$. In either case, (C.49) is satisfied.

3. $\underline{v_0 < v_1 \leq v_3}$: In this range, f is monotonic non-increasing; therefore, for any v_2 between v_1 and v_3, $f(v_2) \geq f(v_3)$, which implies that (C.49) is satisfied.

Since (C.49) is satisfied in all the three cases, we conclude that f is convex. □

C.3 Proof of Corollary 2.1

(a) Let $f(v_0) = 1$. Using Theorem 2.1, we have

$$\mu_{\tilde{F} \sqcup \tilde{G}}(v) = \begin{cases} f(\theta) \wedge f(\theta - k) & ; \; \theta < v_0 \\ f(\theta - k) & ; \; v_0 \leq \theta \leq v_0 + k \\ f(\theta) \vee f(\theta - k) & ; \; \theta > v_0 + k \end{cases} \tag{C.52}$$

We have made use of the fact that the membership function of $\tilde{G}$ is a shifted version of f. The point v_1 in Theorem 2.1, now becomes $(v_0 + k)$. As shown in Appendix C.1, the convexity of $\tilde{F}$ implies that f is monotonic non-decreasing in $(-\infty, v_0]$ and monotonic non-increasing in $[v_0, \infty)$, which implies that $f(\theta) > f(\theta - k)$ for $\theta < v_0$ and $f(\theta) < f(\theta - k)$ for $\theta > v_0 + k$. Using these facts in (C.52), we have

$$\begin{aligned} \mu_{\tilde{F} \sqcup \tilde{G}}(\theta) = f(\theta - k) = \mu_{\tilde{G}}(\theta) \; ; \; \forall \theta \in \Re \\ \Rightarrow \tilde{F} \sqcup \tilde{G} = \tilde{G} \end{aligned} \tag{C.53}$$

A very similar proof can be used for the *meet* operation.

(b) A repeated application of part (a) yields part(b). □

C.4 Proof of Theorem 2.2

From (2.16), we have

$$\neg \tilde{F} = \int_{\theta \in \Re} f(\theta)/(1 - \theta) \tag{C.54}$$

Let $y = (1-\theta)$, then $\theta = (1-y)$ and $\theta \in \Re \Rightarrow y \in \Re$; therefore,

$$\neg\tilde{\mathrm{F}} = \int_{y\in\Re} f(1-y)/y \tag{C.55}$$

$$= \int_{\theta\in\Re} f(1-\theta)/\theta \tag{C.56}$$

□

C.5 Join under Product t-norm

(a) Consider the two convex normal type-1 fuzzy sets, $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$ used in Theorem 2.1. The *join* operation between $\tilde{\mathrm{F}}$ and $\tilde{\mathrm{G}}$, using the product t-norm can be represented as

$$\tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}} = \int_{v\in\Re}\int_{w\in\Re} [f(v)g(w)]/(v \vee w) \tag{C.57}$$

where $\vee$ denotes the maximum. Equation (C.57) is the same as (C.3) with the *min* replaced by a product.

The following analysis is very similar to that in Theorem 2.1. If θ is an element of $\tilde{\mathrm{F}} \sqcup \tilde{\mathrm{G}}$, then the membership grade of θ can be determined by finding all the pairs $\{v, w\}$ such that $v \in \tilde{\mathrm{F}}$, $w \in \tilde{\mathrm{G}}$ and $v \vee w = \theta$; multiplying the membership grades of v and w in each pair; and then finding the maximum of these products of membership grades. The possible admissible $\{v, w\}$ pairs that can give us θ as the result of the *max* operation are $\{v, \theta\}$ where $v \in (-\infty, \theta]$ and $\{\theta, w\}$ where $w \in (-\infty, \theta]$. We find the products of membership grades of v and w from each such pair and take the maximum of all these products as the membership grade of θ. We break this process into three steps : (1) find the product of the memberships of all the pairs $\{v, \theta\}$ where $v \in (-\infty, \theta]$ and then find their supremum; (2) do the same with all the pairs $\{\theta, w\}$ where $w \in (-\infty, \theta]$; and (3) find the maximum of the two suprema, i.e.,

$$\mu_{(\tilde{\mathrm{F}}\sqcup\tilde{\mathrm{G}})}(\theta) = \psi_1(\theta) \vee \psi_2(\theta) \tag{C.58}$$

where

$$\psi_1(\theta) = \sup_{v\in(-\infty,\theta]} \{f(v)g(\theta)\} \tag{C.59}$$

Since $g(\theta)$ is a constant for a given θ,

$$\psi_1(\theta) = g(\theta) \sup_{v\in(-\infty,\theta]} f(v) \tag{C.60}$$

Similarly,

$$\psi_2(\theta) = \sup_{w\in(-\infty,\theta]} \{f(\theta)g(w)\} \tag{C.61}$$

$$= f(\theta) \sup_{w\in(-\infty,\theta]} g(w) \tag{C.62}$$

We break θ into the following three ranges : $\theta < v_0$, $v_0 \leq \theta \leq v_1$ and $\theta > v_1$.
$\underline{\theta = \theta_1 < v_0}$: See Fig. 2.1 (a). Since, $f(v)$ and $g(w)$ both are monotonic non-decreasing in $(-\infty, v_0]$,

$$\sup_{v\in(-\infty,\theta]} f(v) = f(\theta), \tag{C.63}$$

and

$$\sup_{w \in (-\infty, \theta]} g(w) = g(\theta) \tag{C.64}$$

Consequently, from (C.60) and (C.62), we get

$$\psi_1(\theta) = \psi_2(\theta) = f(\theta)g(\theta) \tag{C.65}$$

which implies that [see (C.58)]

$$\mu_{(\tilde{F} \sqcup \tilde{G})}(\theta) = f(\theta)g(\theta)\ ;\ \theta < v_0 \tag{C.66}$$

$\underline{v_0 \leq \theta = \theta_2 \leq v_1}$: See Fig. 2.1 (a). Recall that $f(v_0) = 1$ and that g is monotonic non-decreasing in $(-\infty, v_1]$; therefore,

$$\sup_{v \in (-\infty, \theta]} f(v) = 1\ , \tag{C.67}$$

and

$$\sup_{w \in (-\infty, \theta]} g(w) = g(\theta) \tag{C.68}$$

From (C.60) and (C.62), we get

$$\psi_1(\theta) = g(\theta) \tag{C.69}$$

$$\psi_2(\theta) = f(\theta)g(\theta) \tag{C.70}$$

Since $f(\theta) \leq 1$, $\psi_1(\theta) \geq \psi_2(\theta)$. Consequently, from (C.58)

$$\mu_{(\tilde{F} \sqcup \tilde{G})}(\theta) = g(\theta)\ ;\ v_0 \leq \theta \leq v_1 \tag{C.71}$$

$\underline{\theta = \theta_3 > v_1}$: See Fig. 2.1 (a). For θ in this range, both f and g have already attained their maximum values, i.e.,

$$\sup_{v \in (-\infty, \theta]} f(v) = 1\ , \tag{C.72}$$

and

$$\sup_{w \in (-\infty, \theta]} g(w) = 1\ ; \tag{C.73}$$

therefore, from (C.60) and (C.62), we get

$$\psi_1(\theta) = g(\theta) \tag{C.74}$$

$$\psi_2(\theta) = f(\theta) \tag{C.75}$$

Consequently, from (C.58),

$$\mu_{(\tilde{F} \sqcup \tilde{G})}(\theta) = f(\theta) \vee g(\theta)\ ;\ \theta > v_1 \tag{C.76}$$

Combining (C.66), (C.71) and (C.76), we get (2.23).

(b) The proof of (2.24) is very similar to the proof of (2.21) in Appendix C.1. The only thing that we have to show is that using *max* t-conorm and *product* t-norm, the *meet* of two convex, normal type-1 fuzzy sets is also a convex, normal fuzzy type-1 set. Consider the convex, normal type-1 sets $\tilde{F}$ and $\tilde{G}$ described in part (a) of this proof. We must show that $\tilde{F} \sqcup \tilde{G}$ is also convex and normal under *max* t-conorm and product t-norm. To show convexity, we use the equivalent condition proved in Appendix C.2.

Since $\tilde{F}$ and $\tilde{G}$, both, are convex and normal, f and g both are monotonic non-decreasing in $(-\infty, v_0]$ (recall that $v_0 \leq v_1$), which implies that fg is also monotonic non-decreasing in $(-\infty, v_0]$. Also, $g(v_0) \geq f(v_0)g(v_0)$ and g is monotonic non-decreasing in $[v_0, v_1]$. Consequently, $\mu_{\tilde{F} \sqcup \tilde{G}}(\theta)$ is monotonic non-decreasing in $(-\infty, v_1]$. Since, $f(\theta)$ and $g(\theta)$ are both monotonic non-increasing for $\theta > v_1$, $f(\theta) \vee g(\theta)$ is also monotonic non-increasing for $\theta > v_1$, which implies that $\mu_{\tilde{F}\sqcup\tilde{G}}(\theta)$ is monotonic non-increasing for $\theta > v_1$. Additionally, $\mu_{\tilde{F}\sqcup\tilde{G}}(v_1) = g(v_1) \vee f(v_1) = 1$; hence, $\tilde{F} \sqcup \tilde{G}$ is also convex and normal.

Suppose that we have n convex, normal, type-1 fuzzy sets $\tilde{F}_1, \ldots, \tilde{F}_n$ characterized by membership functions $f_1, \ldots, f_n$, respectively. Let $v_1, v_2, \ldots, v_n$ be real numbers such that $v_1 \leq v_2 \leq \cdots \leq v_n$ and $f_1(v_1) = f_2(v_2) = \cdots = f_n(v_n) = 1$, then proceeding exactly as in the case of the proof of part (b-I) in Theorem 2.1 (see Appendix C.1), by using the associative property of the *join* operation and by repeated application of (2.23), we get (2.24). □

C.6 Meet of Gaussians under Product t-norm

Consider the case when $f(v)$ and $g(w)$ (as in Theorem 2.1) are Gaussians with support $[0, 1]$ with means m_f, m_g and standard deviations σ_f, σ_g, respectively. Then,

$$\tilde{F} \sqcap \tilde{G} = \int_v \int_w e^{-\frac{1}{2}(\frac{v-m_f}{\sigma_f})^2} e^{-\frac{1}{2}(\frac{w-m_g}{\sigma_g})^2} / (vw) \tag{C.77}$$

Recall that the integral in the above equation denotes union in the continuum. If θ is an element of $\tilde{F} \sqcap \tilde{G}$, then the membership grade of θ can be found by : finding all the pairs $\{v, w\}$ such that $v \in \tilde{F}$, $w \in \tilde{G}$ and $vw = \theta$; multiplying the membership grades of v and w in each pair; and then finding the maximum of these products of membership grades, i.e.,

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) = \sup\{e^{-\frac{1}{2}(\frac{v-m_f}{\sigma_f})^2} e^{-\frac{1}{2}(\frac{w-m_g}{\sigma_g})^2}; vw = \theta; v \in \tilde{F}; w \in \tilde{G}\} \tag{C.78}$$

Given any v (assuming $v \neq 0$), the constraint $vw = \theta$ gives us $w = \theta/v$. Further, since $w \in [0, 1]$, it follows that $\theta/v \leq 1$ or $v \geq \theta$. So, given any $\theta \in [0, 1]$, the acceptable $\{v, w\}$ pairs that can give θ as the result of the product operation are $\{(v, \frac{\theta}{v}); 0 \leq \theta \leq v \leq 1\}$; therefore, from (C.78), we have

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) = \sup_{v\in[\theta,1]} e^{-\frac{1}{2}[(\frac{v-m_f}{\sigma_f})^2 + (\frac{\frac{\theta}{v}-m_g}{\sigma_g})^2]} \tag{C.79}$$

Observe that, when $\theta = m_f m_g$, $v = m_f$ maximizes the above quantity, making the exponent 0. This implies that

$$\mu_{\tilde{F}\sqcap\tilde{G}}(m_f m_g) = 1, \tag{C.80}$$

which shows that our result is consistent with the type-1 case result, $m_f \star m_g = m_f m_g$, obtained by reducing type-1 sets $\tilde{F}$ and $\tilde{G}$ to singletons, having unity membership at m_f and m_g respectively and zero membership at all other points. The result of the meet operation is then a singleton also, with unity membership at $m_f m_g$ and zero membership at all other points.

For $\theta \neq m_f m_g$, the only thing that is easily observable is that the exponent in (C.79) does not reduce to zero, implying any θ other than $m_f m_g$ will have a membership grade less than unity. Now, let's see if we can determine an expression for θ in terms of v.

Let us call the quantity in the square bracket on the RHS of (C.79) $J(v)$. The v that achieves the supremum in (C.79), minimizes $J(v)$. In order to find $\mu_{\tilde{F}\sqcap\tilde{G}}(\theta)$, we have to find an expression for v that minimizes

$$J(v) = (\frac{v - m_f}{\sigma_f})^2 + (\frac{\frac{\theta}{v} - m_g}{\sigma_g})^2 \tag{C.81}$$

subject to the constraint $v \in [\theta, 1]$. Differentiating $J(v)$ and equating the derivative to 0, we see that the v_* that achieves the minimum satisfies (with the constraint $v_* \in [\theta, 1]$)

$$\begin{aligned} J'(v_*) = 0 \quad &\Leftrightarrow \quad 2(\frac{v_* - m_f}{\sigma_f})(\frac{1}{\sigma_f}) + 2(\frac{\frac{\theta}{v_*} - m_g}{\sigma_g})(-\frac{\theta}{\sigma_g v_*^2}) = 0 \\ &\Leftrightarrow \quad \frac{v_*}{\sigma_f^2} - \frac{m_f}{\sigma_f^2} - \frac{\theta^2}{\sigma_g^2 v_*^3} + \frac{m_g \theta}{\sigma_g^2 v_*^2} = 0 \\ &\Leftrightarrow \quad v_*^4 - m_f v_*^3 + \theta \frac{\sigma_f^2}{\sigma_g^2} m_g v_* - \theta^2 \frac{\sigma_f^2}{\sigma_g^2} = 0 \qquad \text{(C.82)} \end{aligned}$$

Let us call the polynomial on the LHS of (C.82) $D(v)$. Observe that

$$\begin{aligned} D(m_f) &= \not{m}_f^4 - \not{m}_f^4 + \theta \frac{\sigma_f^2}{\sigma_g^2} m_g m_f - \theta^2 \frac{\sigma_f^2}{\sigma_g^2} \\ &= \theta \frac{\sigma_f^2}{\sigma_g^2} (m_f m_g - \theta) \qquad \text{(C.83)} \\ D(\frac{\theta}{m_g}) &= (\frac{\theta}{m_g})^4 - m_f (\frac{\theta}{m_g})^3 + \theta \frac{\sigma_f^2}{\sigma_g^2} \not{m}_g \frac{\theta}{\not{m}_g} - \theta^2 \frac{\sigma_f^2}{\sigma_g^2} \\ &= (\frac{\theta}{m_g})^3 (\frac{\theta}{m_g} - m_f) \\ &= \frac{\theta^3}{m_g^4} (\theta - m_f m_g) \qquad \text{(C.84)} \end{aligned}$$

From (C.83) and (C.84), we observe that $D(m_f)$ and $D(\theta/m_g)$ are of opposite signs [since θ, m_f and m_g are all in $[0, 1]$, the quantity $(\theta - m_f m_g)$ decides the sign]. This implies that $D(v)$ always has a root between m_f and θ/m_g. As long as $\theta/m_g < 1$, v_* always satisifes the constraint that $v_* \in [\theta, 1]$, because $m_f \leq 1$; however, after a critical value, say $\theta = \theta_c$, $v_* > 1$, and then $v = 1$ minimizes $J(v)$ while satisfying the constraint. The critical value θ_c can be found from (C.82) by expressing θ in terms of v. Rearranging (C.82) and solving for θ, we get

$$\begin{aligned} \theta^2 (\frac{\sigma_f^2}{\sigma_g^2}) - \theta (\frac{\sigma_f^2}{\sigma_g^2} m_g v_*) + (m_f v_*^3 - v_*^4) = 0 \\ \Rightarrow \theta = \frac{v_*}{2} [m_g \pm \sqrt{m_g^2 + 4 \frac{\sigma_g^2}{\sigma_f^2} v_* (v_* - m_f)}] \qquad \text{(C.85)} \end{aligned}$$

We are interested in finding those values of θ for which $v_* \geq 1$. Obviously, this implies that $v_* \geq m_f$, because $m_f \leq 1$ and therefore the second term in the bracket on the RHS of (C.85) is greater than or equal to m_g; hence, keeping the positive root of the above equation (recall that $\theta \geq 0$), we get

$$\theta = \frac{v_*}{2} [m_g + \sqrt{m_g^2 + 4 \frac{\sigma_g^2}{\sigma_f^2} v_* (v_* - m_f)}] \qquad \text{(C.86)}$$

The critical value of θ can be found by substituting $v_* = 1$ in (C.86) and is

$$\theta_c = \frac{1}{2} [m_g + \sqrt{m_g^2 + 4 \frac{\sigma_g^2}{\sigma_f^2} (1 - m_f)}] \qquad \text{(C.87)}$$

So, for $\theta < \theta_c$, v_* can be obtained by solving Eq. (C.82) without using the constraint and then picking the root that satisfies the constraint and minimizes $J(v)$. If there is more than one root that satisfies the constraints, we check the value of $J(v)$ at each of the roots and pick the root at which $J(v)$ is minimum. For $\theta \geq \theta_c$, $v = 1$ minimizes $J(v)$.

In what follows, we attempt to solve Eq. (C.82). We rewrite (C.82) as (for notational simplicity, we drop the subscript "*")

$$v^4 - av^3 + bv - c = 0 \tag{C.88}$$

where

$$\begin{aligned} a &= m_f \\ b &= \frac{\theta m_g \sigma_f^2}{\sigma_g^2} \\ c &= \frac{\theta^2 \sigma_f^2}{\sigma_g^2} \end{aligned} \tag{C.89}$$

In the following, we use a standard procedure for solving quartic (4th order) equations [6]. Substituting $y = v - a/4$ (i.e., $v = y + a/4$) into (C.88), we get

$$(y + \frac{a}{4})^4 - a(y + \frac{a}{4})^3 + b(y + \frac{a}{4}) - c = 0 \tag{C.90}$$

which, upon simplification, gives

$$y^4 - (\frac{3}{8}a^2)y^2 + (b - \frac{a^3}{8})y + (\frac{ab}{4} - \frac{3}{256}a^4 - c) = 0 \tag{C.91}$$

This equation has the following resolvent equation [i.e., if the roots of the following equation are found, the roots of (C.91) can be calculated from them]

$$z^3 - (\frac{3}{4}a^2)z^2 + (\frac{3}{16}a^4 + 4c - ab)z - (b - \frac{a^3}{8})^2 = 0 \tag{C.92}$$

In order to find roots of (C.92), we simplify it further by substituting $z = t + a^2/4$ (i.e., $t = z - a^2/4$). Upon simplification, we get

$$t^3 + (4c - ab)t + (a^2 c - b^2) = 0 \tag{C.93}$$

Let $\alpha = (4c - ab)$, $\beta = (a^2 c - b^2)$ and $D = (\beta/2)^2 + (\alpha/3)^3$; additionally, let

$$\begin{aligned} A &= (-\frac{\beta}{2} + \sqrt{D})^{1/3} \\ B &= -(\frac{\beta}{2} + \sqrt{D})^{1/3} \end{aligned} \tag{C.94}$$

Then, the three roots of (C.93) are

$$\begin{aligned} t_1 &= A + B \\ t_2 &= -(\frac{A+B}{2}) + i(\frac{A-B}{2})\sqrt{3} \\ t_3 &= -(\frac{A+B}{2}) - i(\frac{A-B}{2})\sqrt{3} \end{aligned} \tag{C.95}$$

If $D > 0$, one of these three roots is real and the other two are complex conjugates (which are discarded); if $D = 0$, all the three roots are real and at least two of them are equal; and if $D < 0$, all three roots are real and unequal.

The three roots of the resolvent equation (C.92) can be obtained by adding $a^2/4$ to each of t_1, t_2 and t_3, as

$$\begin{aligned} z_1 &= A + B + \frac{a^2}{4} \\ z_2 &= -(\frac{A+B}{2}) + i(\frac{A-B}{2})\sqrt{3} + \frac{a^2}{4} \\ z_3 &= -(\frac{A+B}{2}) - i(\frac{A-B}{2})\sqrt{3} + \frac{a^2}{4} \end{aligned} \tag{C.96}$$

From these, we can obtain the roots of (C.91) as

$$\begin{aligned} y_1 &= \frac{\sqrt{z_1} + \sqrt{z_2} + \sqrt{z_3}}{2} \\ y_2 &= \frac{\sqrt{z_1} - \sqrt{z_2} - \sqrt{z_3}}{2} \\ y_3 &= \frac{-\sqrt{z_1} + \sqrt{z_2} - \sqrt{z_3}}{2} \\ y_4 &= \frac{-\sqrt{z_1} - \sqrt{z_2} + \sqrt{z_3}}{2} \end{aligned} \tag{C.97}$$

Finally, these four roots of Eq. (C.91) give us the roots of (C.88), as

$$\begin{aligned} v_1 &= \frac{\sqrt{z_1} + \sqrt{z_2} + \sqrt{z_3}}{2} + \frac{a}{4} \\ v_2 &= \frac{\sqrt{z_1} - \sqrt{z_2} - \sqrt{z_3}}{2} + \frac{a}{4} \\ v_3 &= \frac{-\sqrt{z_1} + \sqrt{z_2} - \sqrt{z_3}}{2} + \frac{a}{4} \\ v_4 &= \frac{-\sqrt{z_1} - \sqrt{z_2} + \sqrt{z_3}}{2} + \frac{a}{4} \end{aligned} \tag{C.98}$$

Summarizing, we have four choices for v_* :

$$\begin{aligned} v_1 &= \frac{\sqrt{z_1} + \sqrt{z_2} + \sqrt{z_3}}{2} + \frac{a}{4} \\ v_2 &= \frac{\sqrt{z_1} - \sqrt{z_2} - \sqrt{z_3}}{2} + \frac{a}{4} \\ v_3 &= \frac{-\sqrt{z_1} + \sqrt{z_2} - \sqrt{z_3}}{2} + \frac{a}{4} \\ v_4 &= \frac{-\sqrt{z_1} - \sqrt{z_2} + \sqrt{z_3}}{2} + \frac{a}{4} \end{aligned} \tag{C.99}$$

where

$$\begin{aligned} z_1 &= A + B + \frac{a^2}{4} \\ z_2 &= -(\frac{A+B}{2}) + i(\frac{A-B}{2})\sqrt{3} + \frac{a^2}{4} \end{aligned}$$

$$z_3 = -(\frac{A+B}{2}) - i(\frac{A-B}{2})\sqrt{3} + \frac{a^2}{4} \tag{C.100}$$

and

$$A = [\frac{b^2 - a^2c}{2} + \sqrt{(\frac{a^2-b^2}{2})^2 + (\frac{4c-ab}{3})^3}]^{1/3} \tag{C.101}$$

$$B = [\frac{b^2 - a^2c}{2} - \sqrt{(\frac{a^2-b^2}{2})^2 + (\frac{4c-ab}{3})^3}]^{1/3} \tag{C.102}$$

with a, b, c as in (C.89).

Of these four choices for v_*, we choose the one that is real, satisfies the constraint $v_* \in [\theta, 1]$, and minimizes $J(v)$. [This can be checked by examining $J(\theta)$, $J(1)$ and the value of the second drivative of J at the root.] As mentioned earlier, if there is more than one root satisfying the constraints, we pick the one at which $J(v)$ attains the minimum value.

Summarizing, we have

$$\mu_{\bar{\mathrm{F}}\sqcap\bar{\mathrm{G}}}(\theta) = e^{-\frac{1}{2}[(\frac{v_* - m_f}{\sigma_f})^2 + (\frac{\frac{\theta}{v_*} - m_g}{\sigma_g})^2]} \tag{C.103}$$

where v_* is obtained by solving (C.82) if $\theta < \theta_c$ and $v_* = 1$ if $\theta \geq \theta_c$, where θ_c is given in (C.87). Observe that computations begin by choosing a value for $\theta \in [0, 1]$ and must be repeated for every $\theta \in [0, 1]$, so that θ must, in practice, be discretized.

Figures 2.11 - 2.15 show some examples of Gaussian curves and the result of the *meet* operation between them. The *meet* curve in all the figures was obtained numerically. The order of the curves (i.e., the order $\{m_f, m_g\}$ or $\{\sigma_f, \sigma_g\}$) is not important, which means that the *meet* operation is commutative, as we would expect it to be.

C.7 Solving for the Gaussian Meet Approximation

The problem of maximizing the RHS of (2.55) reduces to the problem of minimizing the objective function

$$H(v) = (\frac{v - m_f}{\sigma_f})^2 + (\frac{\theta - m_g v}{k\sigma_g})^2 \tag{C.104}$$

This minimization also needs to be performed with the constraint $v \in [\theta, 1]$; however, for simplicity, first we minimize H unconstrained and then handle the constraint. Observe that $m_f, m_g, \sigma_f, \sigma_g$ are all positive (in addition, we will always have $m_f, m_g \in [0, 1]$). It can be easily seen that H is convex ($H'' = 2/\sigma_f^2 + 2(m_g/k\sigma_g)^2 > 0$); therefore, equating the first derivative of H to zero (and assuming that the infimum is obtained at $v = v^*$), we get

$$2(\frac{v^* - m_f}{\sigma_f})(\frac{1}{\sigma_f}) + 2(\frac{\theta - m_g v^*}{k\sigma_g})(\frac{-m_g}{k\sigma_g}) = 0$$

$$v^*(\frac{1}{\sigma_f^2} + \frac{m_g^2}{k^2\sigma_g^2}) = \theta\frac{m_g}{k^2\sigma_g^2} + \frac{m_f}{\sigma_f^2} \tag{C.105}$$

$$v^* = \frac{\theta m_g \sigma_f^2 + m_f k^2 \sigma_g^2}{m_g^2\sigma_f^2 + k^2\sigma_g^2} \tag{C.106}$$

Substituting (C.106) into (C.104), we get

$$\inf_v H(v) = (\frac{\theta - m_f m_g}{\sqrt{m_g^2\sigma_f^2 + k^2\sigma_g^2}})^2 \tag{C.107}$$

Now, let us handle the constraint. Since H is convex in v, H is monotonic increasing for $v > v^*$ and monotonic decreasing for $v < v^*$; therefore, if $v^* < \theta$, the minimum in the constraint set is fixed at $v = \theta$, and, if $v^* > 1$, the constrained minimum is fixed at $v = 1$. Otherwise, v^* is as in (C.106).

Observe that the condition $v^* \in [\theta, 1]$ translates into conditions on θ and also depends on the parameters m_f, σ_f, m_g, σ_g and k, since all of them appear on the RHS of (C.106). Next, we analyse these conditions.

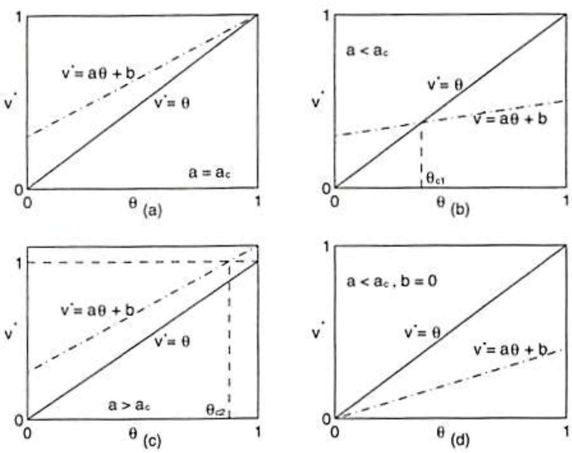


Figure C.1: Plots of $v^* = a\theta + b$ versus θ and different possibilities that can arise depending on the value of a and b. The critical value of a is $a_c = 1 - b$. (a) $a = a_c$. In this case, the constrained minimum is always equal to the true minimum. (b) $a < a_c$. (c) $a > a_c$. (d) The special case, when $b = 0$ and $a > a_c$. In this case, the constrained minimum is equal to the true minimum only at $\theta = 0$.

From (C.106), we can see that v^* is affine in θ. Let

$$a = \frac{m_g \sigma_f^2}{m_g^2 \sigma_f^2 + k^2 \sigma_g^2} \qquad \text{(C.108)}$$

$$b = \frac{m_f k^2 \sigma_g^2}{m_g^2 \sigma_f^2 + k^2 \sigma_g^2} \qquad \text{(C.109)}$$

Then,

$$v^* = a\theta + b, \qquad \text{(C.110)}$$

i.e., if we plot v^* versus θ, we get a straight line. Figures C.1 (a) - (d) show some examples.

Observe that $a \geq 0$ and $0 \leq b \leq 1$ always. For any b, if a is such that the portion of the line in $[0, 1]$ is contained completely in the area above the line $v^* = \theta$ and below the line $v^* = 1$ [Fig. C.1 (a)], then the constrained minimum is always equal to the unconstrained minimum. Let's call the critical value of a that achieves this a_c. To find a_c, we use the condition that $a_c\theta + b = 1$ when $\theta = 1$ (this condition is required for the portion of the line in $[0, 1]$ to be contained above $v^* = \theta$ and below $v^* = 1$). This gives us

$$a_c = 1 - b \qquad \text{(C.111)}$$

If $a < a_c$, after some critical value of θ, θ_{c1}, in $[0, 1]$ [Fig. C.1 (b)], $v^* < \theta$. θ_{c1} is the point of intersection of the lines $v^* = a\theta + b$ and $v^* = \theta$; therefore, θ_{c1} can be found as

$$a\theta_{c1} + b = \theta_{c1}$$

$$\theta_{c1} = \frac{b}{1-a} \tag{C.112}$$

where a and b are as in (C.108) and (C.109).

Similarly, we can see from Fig. C.1 (c) that when $a > a_c$, after some other critical value of θ, θ_{c2}, $v^* > 1$. θ_{c2} is the point of intersection of the lines $v^* = a\theta + b$ and $v^* = 1$; therefore,

$$a\theta_{c2} + b = 1$$
$$\theta_{c2} = \frac{1-b}{a} \tag{C.113}$$

where a and b are as in (C.108) and (C.109).

The above discussion can be summarized in terms of three cases, as follows : Let θ_{c1} and θ_{c2} be as in (C.112) and (C.113), respectively; then,

1. $\underline{a < a_c}$

$$v^* = \begin{cases} \frac{\theta m_g \sigma_f^2 + m_f k^2 \sigma_g^2}{m_g^2 \sigma_f^2 + k^2 \sigma_g^2} & ; \quad \theta \le \theta_{c1} \\ \theta & ; \quad \theta > \theta_{c1} \end{cases} \tag{C.114}$$

Substituting (C.114) into (C.104), we get

$$\inf_{v \in [\theta,1]} H(v) = \begin{cases} (\frac{\theta - m_f m_g}{\sqrt{m_g^2 \sigma_f^2 + k^2 \sigma_g^2}})^2 & ; \quad \theta \le \theta_{c1} \\ (\frac{\theta - m_f}{\sigma_f})^2 + \theta^2 (\frac{1 - m_g}{k \sigma_g})^2 & ; \quad \theta > \theta_{c1} \end{cases} \tag{C.115}$$

Substituting (C.115) into (2.55), we obtain

$$\mu_{\tilde{E}}(\theta) = \begin{cases} e^{-\frac{1}{2}(\frac{\theta - m_f m_g}{\sqrt{m_g^2 \sigma_f^2 + k^2 \sigma_g^2}})^2} & ; \quad \theta \le \theta_{c1} \\ e^{-\frac{1}{2}(\frac{\theta - m_f}{\sigma_f})^2} e^{-\frac{\theta^2}{2}(\frac{1 - m_g}{k \sigma_g})^2} & ; \quad \theta > \theta_{c1} \end{cases} \tag{C.116}$$

2. $\underline{a = a_c}$

$$v^* = \frac{\theta m_g \sigma_f^2 + m_f k^2 \sigma_g^2}{m_g^2 \sigma_f^2 + k^2 \sigma_g^2} \quad ; \quad \theta \in [0,1] \tag{C.117}$$

Substituting (C.117) into (C.104), we get (C.107), i.e.,

$$\inf_{v \in [\theta,1]} H(v) = (\frac{\theta - m_f m_g}{\sqrt{m_g^2 \sigma_f^2 + k^2 \sigma_g^2}})^2 \quad ; \quad \theta \in [0,1] \tag{C.118}$$

Substituting (C.118) into (2.55), we obtain

$$\mu_{\tilde{E}}(\theta) = e^{-\frac{1}{2}(\frac{\theta - m_f m_g}{\sqrt{m_g^2 \sigma_f^2 + k^2 \sigma_g^2}})^2} \quad ; \quad \theta \in [0,1] \tag{C.119}$$

3. $\underline{a > a_c}$

$$v^* = \begin{cases} \frac{\theta m_g \sigma_f^2 + m_f k^2 \sigma_g^2}{m_g^2 \sigma_f^2 + k^2 \sigma_g^2} & ; \quad \theta \le \theta_{c2} \\ 1 & ; \quad \theta > \theta_{c2} \end{cases} \tag{C.120}$$

Substituting (C.120) into (C.104), we get

$$\inf_{v\in[\theta,1]} H(v) = \begin{cases} (\frac{\theta - m_f m_g}{\sqrt{m_g^2\sigma_f^2 + k^2\sigma_g^2}})^2 & ; \ \theta \leq \theta_{c2} \\ (\frac{1-m_f}{\sigma_f})^2 + (\frac{\theta - m_g}{k\sigma_g})^2 & ; \ \theta > \theta_{c2} \end{cases} \tag{C.121}$$

Substituting (C.121) into (2.55), we obtain

$$\mu_{\tilde{E}}(\theta) = \begin{cases} e^{-\frac{1}{2}(\frac{\theta - m_f m_g}{\sqrt{m_g^2\sigma_f^2 + k^2\sigma_g^2}})^2} & ; \ \theta \leq \theta_{c2} \\ e^{-\frac{1}{2}(\frac{1-m_f}{\sigma_f})^2} e^{-\frac{1}{2}(\frac{\theta - m_g}{k\sigma_g})^2} & ; \ \theta > \theta_{c2} \end{cases} \tag{C.122}$$

Now, we come back to the question of choosing an expression for $\mu_{\tilde{F}\sqcap\tilde{G}}$. Recall, that we solved this modified optimization problem because we wanted to find a simple expression for $\tilde{\mu}_{\tilde{F}\sqcap\tilde{G}}$, and it was for this reason that we simplified the actual optimization problem. Although (C.116), (C.119) and (C.122) give the exact solution to the simplified problem, the expressions are still too complicated. Even if we were to accept (C.116), (C.119) and (C.122) as they are, we are still going to have an approximate solution to the actual problem; therefore, it seems very reasonable to choose the simplest of the three possible expressions for $\mu_{\tilde{E}}(\theta)$ and just use that as our approximation of $\mu_{\tilde{F}\sqcap\tilde{G}}(\theta)$. So, we choose the expression for the case $a = a_c$ as the required approximation (in effect, this is equivalent to simplifying the problem even further by disregarding the constraints $v^* \geq \theta$ and $v^* \leq 1$); therefore,

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) \approx e^{-\frac{1}{2}(\frac{\theta - m_f m_g}{\sqrt{m_g^2\sigma_f^2 + k^2\sigma_g^2}})^2} \tag{C.123}$$

As explained in Chapter 2, this expression is also consistent with the type-1 case.

Now, we have to choose some value for k. As explained in Chapter 2, we need some value in $[0, 1]$. Since the *meet* operation is commutative, we want our approximation to also be commutative. This will make generalization to the case of more than two Gaussians easy. By observing (C.123), it is apparent that if we choose $k = m_f$, the approximation becomes commutative (if we interchange $\{m_f, \sigma_f\}$ and $\{m_g, \sigma_g\}$, we still get the same result); so,

$$\mu_{\tilde{F}\sqcap\tilde{G}}(\theta) \approx e^{-\frac{1}{2}(\frac{\theta - m_f m_g}{\sqrt{m_g^2\sigma_f^2 + m_f^2\sigma_g^2}})^2} \tag{C.124}$$

□

C.8 Error Bounds for the Gaussian Meet Approximation

To obtain bounds on the Gaussian approximation error, we first find bounds on the result of the actual *meet* operation between two Gaussians. As explained in Chapter 2, just after (2.57), using $k = 1$ ($k = 0$) in (2.55) is equivalent to finding an upper (lower) bound on the result of the actual *meet* operation. Let's find an upper bound first.

C.8.1 Upper Bound on the Meet between Gaussians

If we just substitute $k = 1$ in the expressions for $\mu_{\tilde{E}}(\theta)$, the resulting curves, generally, will not be Gaussian [see (C.116), (C.119) and (C.122)]. Here, we try to find a Gaussian upper bound for the *meet* operation, to facilitate generalization of this operation to more than two Gaussians. For this purpose, we consider $H(v)$ in (C.104), the objective function for the Gaussian approximation.

For any k, the unconstrained infimum of $H(v)$ should always be less than or equal to its constrained infimum; hence,

$$\begin{aligned} \inf_v H(v) &\leq \inf_{v\in[\theta,1]} H(v) \\ \Rightarrow e^{-\frac{1}{2}[\inf_v H(v)]} &\geq e^{-\frac{1}{2}[\inf_{v\in[\theta,1]} H(v)]} \\ \Rightarrow e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{\sqrt{m_g^2\sigma_f^2 + k^2\sigma_g^2}}\right)^2} &\geq \mu_{\tilde{E}}(\theta) \end{aligned} \tag{C.125}$$

where, we have made use of (C.107) and (2.55). A suitable upper bound on the result of the *meet* between two Gaussians can therefore be obtained by substituting $k = 1$ in (C.125), i.e.,

$$\mu^U_{\tilde{F}\sqcap\tilde{G}}(\theta) = e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{\sqrt{m_g^2\sigma_f^2 + \sigma_g^2}}\right)^2} \tag{C.126}$$

Observe that (C.126) is not symmetrical in $\{m_f, \sigma_f\}$ and $\{m_g, \sigma_g\}$, i.e., if we interchange the fuzzy sets $\tilde{F}$ and $\tilde{G}$, we get a different expression for the upper bound,

$$\mu^U_{\tilde{G}\sqcap\tilde{F}}(\theta) = e^{-\frac{1}{2}\left(\frac{\theta - m_g m_f}{\sqrt{m_f^2\sigma_g^2 + \sigma_f^2}}\right)^2} \tag{C.127}$$

Both (C.126) and (C.127) give an upper bound for the *meet* between $\tilde{F}$ and $\tilde{G}$; therefore, to ensure that the upper bound is independent of the order of the two Gaussians, we choose the minimum of these two functions as the upper bound [because both (C.126) and (C.127) are upper bounds, the minimum of the two is also an upper bound], i.e.,

$$\mu^U_{\tilde{F}\sqcap\tilde{G}}(\theta) = e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{\sigma_{u2}}\right)^2} \tag{C.128}$$

where

$$\sigma_{u2} = \min\left\{\sqrt{m_g^2\sigma_f^2 + \sigma_g^2}, \sqrt{m_f^2\sigma_g^2 + \sigma_f^2}\right\} \tag{C.129}$$

Let's see how this result generalizes to the *meet* of more than two Gaussians. Suppose that we have to find the *meet* between three type-1 Gaussian fuzzy sets $\tilde{F}_1$, $\tilde{F}_2$ and $\tilde{F}_3$, having means m_1, m_2 and m_3, respectively, and standard deviations σ_1, σ_2 and σ_3, respectively. If we perform the *meet* between $\tilde{F}_1$ and $\tilde{F}_2$ first, (C.128) gives us the following upper bound

$$\mu^U_{\tilde{F}_1\sqcap\tilde{F}_2}(\theta) = e^{-\frac{1}{2}\left(\frac{\theta - m_1 m_2}{\sigma_{u12}}\right)^2} \tag{C.130}$$

where

$$\sigma_{u12} = \min\left\{\sqrt{m_1^2\sigma_2^2 + \sigma_1^2}, \sqrt{m_2^2\sigma_1^2 + \sigma_2^2}\right\} \tag{C.131}$$

which can also be rewritten as

$$\sigma_{u12} = \sqrt{\min\{m_1^2\sigma_2^2 + \sigma_1^2, m_2^2\sigma_1^2 + \sigma_2^2\}} \tag{C.132}$$

Now, an upper bound on the *meet* of $\tilde{F}_1$, $\tilde{F}_2$ and $\tilde{F}_3$ can be found by finding the upper bound on the *meet* of $\tilde{F}_3$ and the Gaussian in (C.130), i.e.,

$$\mu^U_{\tilde{F}_1\sqcap\tilde{F}_2\sqcap\tilde{F}_3}(\theta) = e^{-\frac{1}{2}\left(\frac{\theta - m_1 m_2 m_3}{\sigma_{u123}}\right)^2} \tag{C.133}$$

where

$$
\begin{aligned}
\sigma_{u123} &= \sqrt{\min\{m_1^2 m_2^2 \sigma_3^2 + \sigma_{u12}^2,\ m_3^2 \sigma_{u12}^2 + \sigma_3^2\}} \\
&= \Big[\min\Big\{m_1^2 m_2^2 \sigma_3^2 + \min\{m_1^2\sigma_2^2 + \sigma_1^2,\ m_2^2\sigma_1^2 + \sigma_2^2\}, \\
&\qquad m_3^2 \min\{m_1^2\sigma_2^2 + \sigma_1^2,\ m_2^2\sigma_1^2 + \sigma_2^2\} + \sigma_3^2\Big\}\Big]^{\frac{1}{2}} \\
&= \Big[\min\Big\{\min\{\sigma_1^2 + m_1^2\sigma_2^2 + m_1^2 m_2^2 \sigma_3^2,\ \sigma_2^2 + m_2^2\sigma_1^2 + m_1^2 m_2^2 \sigma_3^2\}, \\
&\qquad \min\{\sigma_3^2 + m_3^2\sigma_2^2 + m_2^2 m_3^2 \sigma_1^2,\ \sigma_3^2 + m_3^2\sigma_1^2 + m_1^2 m_3^2 \sigma_2^2\}\Big\}\Big]^{\frac{1}{2}} \\
&= \Big[\min\Big\{\sigma_1^2 + m_1^2\sigma_2^2 + m_1^2 m_2^2 \sigma_3^2,\ \sigma_2^2 + m_2^2\sigma_1^2 + m_1^2 m_2^2 \sigma_3^2, \\
&\qquad \sigma_3^2 + m_3^2\sigma_2^2 + m_2^2 m_3^2 \sigma_1^2,\ \sigma_3^2 + m_3^2\sigma_1^2 + m_1^2 m_3^2 \sigma_2^2\Big\}\Big]^{\frac{1}{2}} \qquad \text{(C.134)}
\end{aligned}
$$

This expression is also not symmetric in $\tilde{F}_1$, $\tilde{F}_2$ and $\tilde{F}_3$, i.e.,

$$\mu^U_{\tilde{F}_2 \sqcap \tilde{F}_3 \sqcap \tilde{F}_1}(\theta) = e^{-\frac{1}{2}(\frac{\theta - m_1 m_2 m_3}{\sigma_{u231}})^2} \qquad \text{(C.135)}$$

where

$$
\begin{aligned}
\sigma_{u231} &= \Big[\min\Big\{\sigma_1^2 + m_1^2\sigma_2^2 + m_1^2 m_2^2 \sigma_3^2,\ \sigma_1^2 + m_1^2\sigma_3^2 + m_1^2 m_3^2 \sigma_2^2, \\
&\qquad \sigma_2^2 + m_2^2\sigma_3^2 + m_2^2 m_3^2 \sigma_1^2,\ \sigma_3^2 + m_3^2\sigma_2^2 + m_2^2 m_3^2 \sigma_1^2\Big\}\Big]^{\frac{1}{2}}, \qquad \text{(C.136)}
\end{aligned}
$$

and,

$$\mu^U_{\tilde{F}_1 \sqcap \tilde{F}_3 \sqcap \tilde{F}_2}(\theta) = e^{-\frac{1}{2}(\frac{\theta - m_1 m_2 m_3}{\sigma_{u132}})^2} \qquad \text{(C.137)}$$

where

$$
\begin{aligned}
\sigma_{u132} &= \Big[\min\Big\{\sigma_2^2 + m_2^2\sigma_1^2 + m_1^2 m_2^2 \sigma_3^2,\ \sigma_2^2 + m_2^2\sigma_3^2 + m_2^2 m_3^2 \sigma_1^2, \\
&\qquad \sigma_1^2 + m_1^2\sigma_3^2 + m_1^2 m_3^2 \sigma_2^2,\ \sigma_3^2 + m_3^2\sigma_1^2 + m_1^2 m_3^2 \sigma_2^2\Big\}\Big]^{\frac{1}{2}} \qquad \text{(C.138)}
\end{aligned}
$$

Observe that we have considered all the possible orderings of $\tilde{F}_1$, $\tilde{F}_2$ and $\tilde{F}_3$ that would give us distinct results for $\mu^U(\theta)$, e.g., since the expression for the upper bound for the *meet* of two Gaussians is commutative, $\mu^U_{(\tilde{F}_1 \sqcap \tilde{F}_2) \sqcap \tilde{F}_3} = \mu^U_{(\tilde{F}_2 \sqcap \tilde{F}_1) \sqcap \tilde{F}_3}$.

We choose the minimum of the three Gaussians in (C.133), (C.135) and (C.137) as the final upper bound, i.e.,

$$\mu^U_{\tilde{F}_1 \sqcap \tilde{F}_2 \sqcap \tilde{F}_3} = e^{-\frac{1}{2}(\frac{\theta - m_1 m_2 m_3}{\sigma_{u3}})^2} \qquad \text{(C.139)}$$

where (some terms are common to σ_{u123}, σ_{u231} and σ_{u132})

$$\sigma_{u3} = \min\{\sigma_{u123},\ \sigma_{u231},\ \sigma_{132}\}$$

$$= \Bigg[\min \Big\{ \sigma_1^2 + m_1^2\sigma_2^2 + m_1^2m_2^2\sigma_3^2,\ \sigma_1^2 + m_1^2\sigma_3^2 + m_1^2m_3^2\sigma_2^2,$$
$$\sigma_2^2 + m_2^2\sigma_1^2 + m_1^2m_2^2\sigma_3^2,\ \sigma_2^2 + m_2^2\sigma_3^2 + m_2^2m_3^2\sigma_1^2,$$
$$\sigma_3^2 + m_3^2\sigma_1^2 + m_1^2m_3^2\sigma_2^2,\ \sigma_3^2 + m_3^2\sigma_2^2 + m_2^2m_3^2\sigma_1^2 \Big\} \Bigg]^{\frac{1}{2}} \tag{C.140}$$

Continuing in this fashion, the upper bound on the *meet* between n Gaussians $\tilde{F}_1, \ldots, \tilde{F}_n$ having means $m_1, \ldots, m_n$ and standard deviations $\sigma_1, \ldots, \sigma_n$, respectively, is given as

$$\mu^U_{\sqcap_{i=1}^n \tilde{F}_i} = \exp\{-\frac{1}{2}(\frac{\theta - \prod_{i=1}^n m_i}{\bar{\sigma}_u})^2\} \tag{C.141}$$

where

$$\bar{\sigma}_u = \min\{\Sigma_1, \Sigma_2, \ldots, \Sigma_{n!}\} \tag{C.142}$$

where

$$\Sigma_j = \sqrt{c_{i_1}^2\sigma_{i_1}^2 + c_{i_2}^2\sigma_{i_2}^2 + \quad \cdots \quad + c_{i_n}^2\sigma_{i_n}^2} \quad ; \quad j = 1, \ldots, n! \tag{C.143}$$

where $\{i_1, \ldots, i_n\}$ indicates a permutation of $\{1, 2, \ldots, n\}$, and the c_{i_k}s ($k = 1, \ldots, n$) are calculated as follows :

$$\begin{aligned} c_{i_1} &= 1, \\ c_{i_k} &= m_{i_{(k-1)}} c_{i_{(k-1)}} \quad \text{for} \quad k = 2, \ldots, n. \end{aligned} \tag{C.144}$$

In order to illustrate the use and validity of (C.143) and (C.144), we consider the following example.

Example C.1 For $n = 3$, $\bar{\sigma}_u$ can be calculated, as follows. There are $3! = 6$ possible permutations of $\{1, 2, 3\}$. For each one of these permutations, we use (C.144) to calcuate the c_{i_j}'s, and (C.143) to calculate the Σ_j's :

$$\begin{array}{lll}
\{i_1, i_2, i_3\} = \{1, 2, 3\} & : \ c_1 = 1; & c_2 = m_1c_1 = m_1; \quad c_3 = m_2c_2 = m_1m_2 \\
& & \Sigma_1 = \sqrt{\sigma_1^2 + m_1^2\sigma_2^2 + m_1^2m_2^2\sigma_3^2} \\
\{i_1, i_2, i_3\} = \{1, 3, 2\} & : \ c_1 = 1; & c_3 = m_1c_1 = m_1; \quad c_2 = m_3c_3 = m_1m_3 \\
& & \Sigma_2 = \sqrt{\sigma_1^2 + m_1^2\sigma_3^2 + m_1^2m_3^2\sigma_2^2} \\
\{i_1, i_2, i_3\} = \{2, 1, 3\} & : \ c_2 = 1; & c_1 = m_2c_2 = m_2; \quad c_3 = m_1c_1 = m_1m_2 \\
& & \Sigma_3 = \sqrt{\sigma_2^2 + m_2^2\sigma_1^2 + m_1^2m_2^2\sigma_3^2} \\
\{i_1, i_2, i_3\} = \{2, 3, 1\} & : \ c_2 = 1; & c_3 = m_2c_2 = m_2; \quad c_1 = m_3c_3 = m_2m_3 \\
& & \Sigma_4 = \sqrt{\sigma_2^2 + m_2^2\sigma_3^2 + m_2^2m_3^2\sigma_1^2} \\
\{i_1, i_2, i_3\} = \{3, 1, 2\} & : \ c_3 = 1; & c_1 = m_3c_3 = m_3; \quad c_2 = m_1c_1 = m_1m_3 \\
& & \Sigma_5 = \sqrt{\sigma_3^2 + m_3^2\sigma_1^2 + m_1^2m_3^2\sigma_2^2} \\
\{i_1, i_2, i_3\} = \{3, 2, 1\} & : \ c_3 = 1; & c_2 = m_3c_3 = m_3; \quad c_1 = m_2c_2 = m_2m_3 \\
& & \Sigma_6 = \sqrt{\sigma_3^2 + m_3^2\sigma_2^2 + m_2^2m_3^2\sigma_1^2}
\end{array} \tag{C.145}$$

Using $\Sigma_1, \ldots, \Sigma_6$ from (C.145) in (C.142), we can verify that the $\bar{\sigma}_u$ we obtain, is the same as in (C.140). □

As is apparent from (C.142), the calculation of σ_u is computationally intensive. To find an upper bound on the *meet* between 4 Gaussians, we need to find the minimum of $4! = 24$ terms; when 5 Gaussians are involved, the number of terms rises to 120, and so on. Observe, though, that the minimum in (C.142) gives the tightest of all bounds (i.e., tightest of the bounds that we have derived). If one just wants to find any upper bound, the minimization in (C.142) is not necessary. Each of the Σ_js ($j = 1, 2, \ldots, n!$) in (C.142) is the standard deviation of one of the upper bounds of the *meet*; so, we can just choose one of the $n!$ terms and use it in (C.141) to get an upper bound on the *meet*.

C.8.2 Effect of Clipping

So far, in our derivations, we have assumed that we have perfect Gaussians, i.e., we have neglected the fact that the actual curves are Gaussians contained in $[0,1]$ and may, therefore, be clipped (i.e., any portion of the Gaussians lying outside the interval $[0,1]$ is cut-off); however, this clipping does not change the upper bounds in (C.128) and (C.141). The reason can be explained as follows. Let $\tilde{F}_c$ and $\tilde{G}_c$ be clipped type-1 fuzzy sets having Gaussian membership functions contained in $[0,1]$. Though the membership functions of $\tilde{F}_c$ and $\tilde{G}_c$ are defined only on $[0,1]$, in numerical calculations the membership functions are treated as if they are 0 before 0 and after 1. Therefore, $\mu_{\tilde{F}_c} \leq \mu_{\tilde{F}}$ and $\mu_{\tilde{G}_c} \leq \mu_{\tilde{G}}$, where $\tilde{F}$ and $\tilde{G}$ are type-1 sets whose membership functions are perfect (unclipped) Gaussians. (Figure C.2 shows an example of clipped and unclipped Gaussians.) Consequently, $\mu_{\tilde{F}_c \sqcap \tilde{G}_c} \leq \mu_{\tilde{F} \sqcap \tilde{G}}$ and therefore the upper bound derived above also holds in the case of clipped Gaussians. We will have to consider the explicit effects of clipping when deriving the lower bound for the *meet* in Section C.8.3.

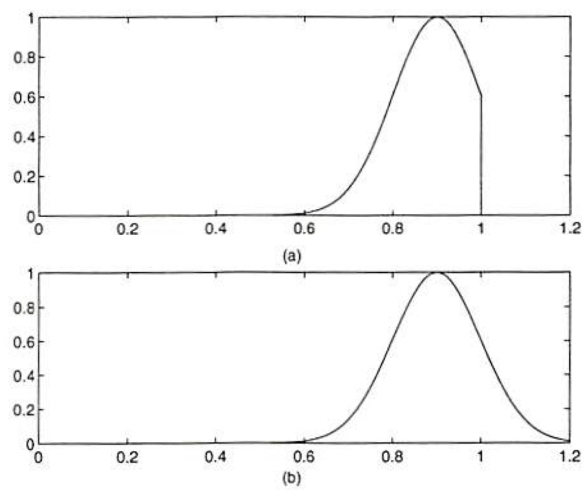


Figure C.2: A Gaussian contained in $[0,1]$ may be clipped as shown in (a). Figure (b) shows the unclipped version of the same Gaussian.

C.8.3 Lower Bound on the Meet between Gaussians

As we have seen earlier, substituting $k=0$ into (2.55) is analogous to finding a lower bound on the *meet* between two Gaussians. Now, however, we also have to include the clipping effects mentioned in Section C.8.2, because a lower bound that assumes perfect (unclipped) Gaussians may not work for clipped Gaussians, since, as mentioned in Section C.8.2, $\mu_{\tilde{F}_c \sqcap \tilde{G}_c} \leq \mu_{\tilde{F} \sqcap \tilde{G}}$, where $\tilde{F}_c$ and $\tilde{G}_c$ are clipped versions of Gaussian type-1 sets $\tilde{F}$ and $\tilde{G}$. So, instead of just substituting $k=0$ into (2.55), we go back to the beginning of the derivation for the Gaussian meet approximation.

Recall that using $k=0$ is equivalent to assuming that one of the Gaussians has zero standard deviation (Section 2.3.2), i.e., it is equivalent to assuming that type-1 fuzzy set $\tilde{F}$ (or $\tilde{G}$) has a membership equal to 1 at m_f (or m_g) and equal to 0 at all other points. Let's see what happens if this is really the case.

Assume that $\mu_{\tilde{F}}(m_f) = 1$ and $\mu_{\tilde{F}}(\theta) = 0$ for $\theta \neq m_f$, and $\mu_{\tilde{G}}(\theta)$ is a Gaussian contained in $[0,1]$, i.e.,

$$\mu_{\tilde{G}}(\theta) = \begin{cases} e^{-\frac{1}{2}(\frac{\theta - m_g}{\sigma_g})^2} & ; \quad \theta \in [0,1] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.146}$$

The result of the *meet* of $\tilde{F}$ and $\tilde{G}$ is just a scaled version of $\tilde{G}$, i.e.,

$$\begin{aligned} \mu_{\tilde{F} \sqcap \tilde{G}}(\theta) &= \mu_{\tilde{G}}(\theta / m_f) \\ &= \begin{cases} e^{-\frac{1}{2}(\frac{\frac{\theta}{m_f} - m_g}{\sigma_g})^2} & ; \quad \frac{\theta}{m_f} \in [0,1] \\ 0 & ; \quad \text{otherwise} \end{cases} \\ &= \begin{cases} e^{-\frac{1}{2}(\frac{\theta - m_f m_g}{m_f \sigma_g})^2} & ; \quad \theta \in [0, m_f] \\ 0 & ; \quad \text{otherwise} \end{cases} \end{aligned} \tag{C.147}$$

Equation (C.147) follows from the definition of the *meet* operation under product t-norm [see Eq. (2.36) and also Section 2.3.1]. In Section 2.3.1, we ignored the clipping effects mentioned in Section C.8.2; but, here, we have to take them into account to make sure that the lower bound holds in all possible cases. Since the Gaussian $\mu_{\tilde{G}}(\theta)$ is contained in $[0,1]$, the resulting function in (C.147) is nonzero only in $[0, m_f]$.

Now, if we assume that G is the singleton, i.e., if $\mu_{\tilde{G}}(m_g) = 0$ and $\mu_{\tilde{G}}(\theta) = 0$ for $\theta \neq m_g$, and $\mu_{\tilde{F}}(\theta)$ is a Gaussian contained in $[0,1]$, then we get

$$\mu_{\tilde{F} \sqcap \tilde{G}}(\theta) = \begin{cases} e^{-\frac{1}{2}(\frac{\theta - m_f m_g}{m_g \sigma_f})^2} & ; \quad \theta \in [0, m_g] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.148}$$

The actual lower bound can be taken as the maximum (since each of them is a lower bound) of the results in (C.147) and (C.148). If we assume that $m_f \leq m_g$, then the lower bound is

$$\mu^L_{\tilde{F} \sqcap \tilde{G}}(\theta) = \begin{cases} \exp\left\{ -\frac{1}{2}(\frac{\theta - m_f m_g}{\max\{m_f \sigma_g, m_g \sigma_f\}})^2 \right\} & ; \quad \theta \in [0, m_f] \\ \exp\left\{ -\frac{1}{2}(\frac{\theta - m_f m_g}{m_g \sigma_f})^2 \right\} & ; \quad \theta \in [m_f, m_g] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.149}$$

If we assume that $m_g \leq m_f$, the lower bound is

$$\mu^L_{\tilde{F} \sqcap \tilde{G}}(\theta) = \begin{cases} \exp\left\{ -\frac{1}{2}(\frac{\theta - m_f m_g}{\max\{m_f \sigma_g, m_g \sigma_f\}})^2 \right\} & ; \quad \theta \in [0, m_g] \\ \exp\left\{ -\frac{1}{2}(\frac{\theta - m_f m_g}{m_f \sigma_g})^2 \right\} & ; \quad \theta \in [m_g, m_f] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.150}$$

To simplify (C.149) and (C.150) a little bit, we ignore the part of the curve lying outside $[0, \min\{m_f, m_g\}]$. Doing this will make the lower bound a little bit loose, but will let us generalize to the *meet* of more than two sets easily. Hence, without any assumptions on m_f and m_g,

$$\mu^L_{\tilde{F} \sqcap \tilde{G}}(\theta) = \begin{cases} \exp\left\{ -\frac{1}{2}(\frac{\theta - m_f m_g}{\max\{m_f \sigma_g, m_g \sigma_f\}})^2 \right\} & ; \quad \theta \in [0, \min\{m_f, m_g\}] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.151}$$

A generalization of (C.151) to the case of more than two Gaussians is a bit tedious; therefore, we take a different approach as illustrated by the following :

Example C.2 Consider the meet between three Gaussians $\tilde{F}_1$, $\tilde{F}_2$, $\tilde{F}_3$, in $[0,1]$, with means m_1, m_2, m_3 and standard deviations σ_1, σ_2, σ_3. We consider three cases to find a lower bound for the *meet*. First, we assume that $\tilde{F}_2$ and $\tilde{F}_3$ are singletons at m_2 and m_3, respectively (zero standard deviations), and $\tilde{F}_1$ is a Gaussian contained in $[0,1]$ with mean m_1 and standard deviation σ_1. This is equivalent to finding the *meet* between $\tilde{F}_1$ and a singleton at $m_2 m_3$, because *meet* under product t-norm is multiplication under product t-norm. Let the result of the *meet* be $\tilde{F}^1_{23}$. Proceeding as in the derivation of (C.147) and (C.148), the membership function of $\tilde{F}^1_{23}$ can be obtained as

$$\mu_{\tilde{F}^1_{23}} = \begin{cases} \exp\left\{-\frac{1}{2}\left(\frac{\theta - m_1 m_2 m_3}{m_2 m_3 \sigma_1}\right)^2\right\} & ; \quad \theta \in [0, m_2 m_3] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.152}$$

Then, we find the *meet* between $\tilde{F}_1$, $\tilde{F}_2$ and $\tilde{F}_3$ by assuming that $\tilde{F}_1$ and $\tilde{F}_3$ are singletons at m_1 and m_3, respectively, and $\tilde{F}_2$ is not a singleton. Let us call the result of this *meet* operation $\tilde{F}^2_{13}$. Its membership function is

$$\mu_{\tilde{F}^2_{13}} = \begin{cases} \exp\left\{-\frac{1}{2}\left(\frac{\theta - m_1 m_2 m_3}{m_1 m_3 \sigma_2}\right)^2\right\} & ; \quad \theta \in [0, m_1 m_3] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.153}$$

Finally, we find the *meet* between $\tilde{F}_1$, $\tilde{F}_2$ and $\tilde{F}_3$ by assuming that $\tilde{F}_1$ and $\tilde{F}_2$ are singletons at m_1 and m_2, respectively, and $\tilde{F}_3$ is not a singleton. Let us call the result of this *meet* operation $\tilde{F}^3_{12}$. Its membership function is

$$\mu_{\tilde{F}^3_{12}} = \begin{cases} \exp\left\{-\frac{1}{2}\left(\frac{\theta - m_1 m_2 m_3}{m_1 m_2 \sigma_3}\right)^2\right\} & ; \quad \theta \in [0, m_1 m_2] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.154}$$

Since $\mu_{\tilde{F}^1_{23}}$, $\mu_{\tilde{F}^2_{13}}$ and $\mu_{\tilde{F}^3_{12}}$, each give a lower bound on the *meet* between $\tilde{F}_1$, $\tilde{F}_2$ and $\tilde{F}_3$, we choose the maximum of them [again, as in (C.151), we keep only that part of the maximum function which lies in $[0, \min\{m_1 m_2, m_2 m_3, m_1 m_3\}]$, for simplicity]. This gives us

$$\mu^L_{\tilde{F}_1 \sqcap \tilde{F}_2 \sqcap \tilde{F}_3}(\theta) = \begin{cases} \exp\left\{-\frac{1}{2}\left(\frac{\theta - m_1 m_2 m_3}{\bar{\sigma}_3}\right)^2\right\} & ; \quad \theta \in [0, l_3] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.155}$$

where

$$\bar{\sigma}_3 = \max\{m_1 m_2 \sigma_3, m_1 m_3 \sigma_2, m_2 m_3 \sigma_1\}, \tag{C.156}$$

and

$$l_3 = \min\{m_1 m_2, m_2 m_3, m_1 m_3\} \tag{C.157}$$

□

Now, suppose that we have n Gaussian fuzzy sets, $\tilde{F}_1, \tilde{F}_2, \ldots, \tilde{F}_n$, in $[0,1]$ having means $m_1, m_2, \ldots, m_n$ and standard deviations $\sigma_1, \sigma_2, \ldots, \sigma_n$. A lower bound on their *meet* is obtained by generalizing (C.155) to the case of n Gaussians, i.e.,

$$\mu^L_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) = \begin{cases} \exp\left\{-\frac{1}{2}\left(\frac{\theta - \prod_{i=1}^n m_i}{\bar{\sigma}_l}\right)^2\right\} & ; \quad \theta \in [0, l_n] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{C.158}$$

where $(i = 1, 2, \cdots, n)$

$$\bar{\sigma}_l = \max\left\{\sigma_1 \prod_{i;i\neq 1} m_i, \sigma_2 \prod_{i;i\neq 2} m_i, \cdots, \sigma_j \prod_{i;i\neq j} m_i, \cdots, \sigma_n \prod_{i;i\neq n} m_i\right\} \tag{C.159}$$

and

$$l_n = \min\Big\{ \prod_{i;i\neq 1} m_i, \prod_{i;i\neq 2} m_i, \cdots, \prod_{i;i\neq j} m_i, \cdots, \prod_{i;i\neq n} m_i \Big\} \qquad \text{(C.160)}$$

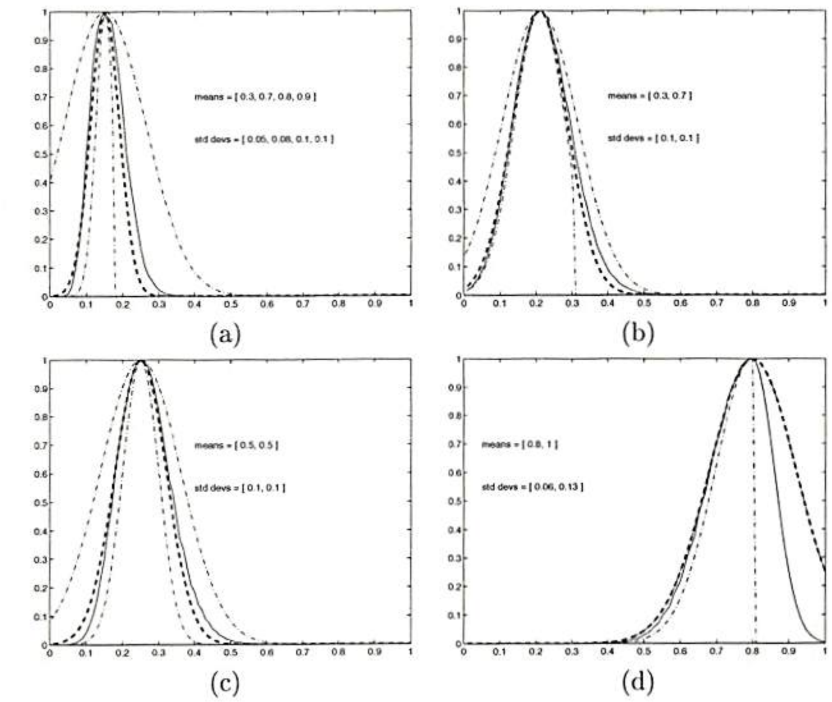


Figure C.3: Some examples of upper and lower bounds given in (C.141) and (C.158). In each figure, the solid line shows the actual result of the *meet* (computed numerically), the thick dashed line shows the Gaussian approximation and the dash-dotted lines show upper and lower bounds. The Gaussians in (c) are coincident (same means and standard deviations). The Gaussians in (d) are the same as in Fig. 2.15 (f). Since one of the Gaussians is centered at 1 (half of it is clipped), the approximation does not work as well as in the other cases; however, the upper and lower bounds still hold. In this case, the upper bound coincides with the approximation.

Figure C.3 shows some examples of the just-derived upper and lower bounds. Both, the upper and the lower bounds, that we have derived are quite conservative. It may be possible to derive tighter bounds; however, we will not pursue this issue any further.

C.8.4 Bounds on the Gaussian Approximation Error

Before proceeding to find bounds on the error for the *meet* between Gaussians and the Gaussian approximation of the *meet*, we shall show that, just as the upper and lower bounds derived in this section enclose the actual *meet* curve between them, they also enclose the Gaussian approximation for the *meet*, i.e., we shall show that

$$\mu^L_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) \leq \hat{\mu}_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) \leq \mu^U_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) \quad ; \quad \theta \in [0,1] \qquad \text{(C.161)}$$

Recall that to find the Gaussian approximation for the *meet* between two Gaussian fuzzy sets, $\tilde{F}$ and $\tilde{G}$, we solved an optimization problem and arrived at the following expression for the Gaussian's standard deviation [see Eq. (C.123)] ,

$$\sigma(k) = \sqrt{m_g^2\sigma_f^2 + k^2\sigma_g^2} \quad , \quad k \in [0,1] \tag{C.162}$$

Obviously, $\sigma(0) \leq \sigma(k) \leq \sigma(1)$ for any $k \in [0,1]$. In particular, $\sigma(0) \leq \sigma(m_f) \leq \sigma(1)$, i.e.,

$$m_g\sigma_f \leq \sqrt{m_g^2\sigma_f^2 + m_f^2\sigma_g^2} \leq \sqrt{m_g^2\sigma_f^2 + \sigma_g^2} \tag{C.163}$$

These results are also true when we change the order of $\tilde{F}$ and $\tilde{G}$; hence,

$$m_f\sigma_g \leq \sqrt{m_g^2\sigma_f^2 + m_f^2\sigma_g^2} \leq \sqrt{m_f^2\sigma_g^2 + \sigma_f^2} \tag{C.164}$$

Combining (C.163) and (C.164), we get

$$\max\{m_f\sigma_g\,,\, m_g\sigma_f\} \leq \sqrt{m_f^2\sigma_g^2 + m_g^2\sigma_f^2} \leq \min\left\{\sqrt{m_f^2\sigma_g^2 + \sigma_f^2}\,,\, \sqrt{m_g^2\sigma_f^2 + \sigma_g^2}\right\} \tag{C.165}$$

Consequently,

$$e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{\max\{m_f\sigma_g, m_g\sigma_f\}}\right)^2} \leq e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{\sqrt{m_f^2\sigma_g^2 + m_g^2\sigma_f^2}}\right)^2} \leq e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{\min\left\{\sqrt{m_f^2\sigma_g^2+\sigma_f^2},\, \sqrt{m_g^2\sigma_f^2+\sigma_g^2}\right\}}\right)^2} \tag{C.166}$$

Also, from (C.151), it follows that

$$\mu^L_{\tilde{F}\sqcap\tilde{G}} \leq e^{-\frac{1}{2}\left(\frac{\theta - m_f m_g}{\max\{m_f\sigma_g, m_g\sigma_f\}}\right)^2} \tag{C.167}$$

If we denote the Gaussian approximation in (C.124) by $\hat{\mu}_{\tilde{F}\sqcap\tilde{G}}$, it follows from (C.128), (C.129), (C.166), and (C.167) that

$$\mu^L_{\tilde{F}\sqcap\tilde{G}}(\theta) \leq \hat{\mu}_{\tilde{F}\sqcap\tilde{G}}(\theta) \leq \mu^U_{\tilde{F}\sqcap\tilde{G}}(\theta) \quad ; \quad \theta \in [0,1] \tag{C.168}$$

This is in general true for the *meet* of any number of Gaussians, as can be verified from (2.61), (C.141), and (C.158) in a similar manner; hence,

$$\mu^L_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) \leq \hat{\mu}_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) \leq \mu^U_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) \quad ; \quad \theta \in [0,1]$$

Suppose that we have n Gaussian fuzzy sets, $\tilde{F}_1, \tilde{F}_2, \ldots, \tilde{F}_n$, in $[0,1]$ having means $m_1, m_2, \ldots, m_n$ and standard deviations $\sigma_1, \sigma_2, \ldots, \sigma_n$; then, we have shown earlier (in Sections C.8.1 and C.8.3) that

$$\mu^L_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) \leq \mu_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) \leq \mu^U_{\sqcap_{i=1}^n \tilde{F}_i}(\theta) \quad ; \quad \theta \in [0,1] \tag{C.169}$$

From (C.161) and (C.169), we can see that (dropping the subscript "$\sqcap_{i=1}^n \tilde{F}_i$" for notational convenience) if $\hat{\mu}(\theta) \leq \mu(\theta)$,

$$\mu(\theta) - \hat{\mu}(\theta) \leq \mu^U(\theta) - \hat{\mu}(\theta) \tag{C.170}$$

Similarly, if $\hat{\mu}(\theta) \geq \mu(\theta)$,

$$\hat{\mu}(\theta) - \mu(\theta) \leq \hat{\mu}(\theta) - \mu^L(\theta) \tag{C.171}$$

From (C.170) and (C.171), we can bound the approximation error as

$$\begin{aligned}\left|\hat{\mu}_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta) - \mu_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)\right| &\leq \max\left\{\left[\mu^U_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta) - \hat{\mu}_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)\right],\right.\\ &\left.\left[\hat{\mu}_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta) - \mu^L_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)\right]\right\};\ \theta \in [0,1] \qquad \text{(C.172)}\end{aligned}$$

where $\hat{\mu}_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)$ is given in (2.61), $\mu^U_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)$ is given in (C.141) and $\mu^L_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)$ is given in (C.158).

Equation (C.172) gives an upper bound on the approximation error at any point $\theta \in [0,1]$. To use (C.172), one needs to compute the upper and lower bounds $\left[\mu^U_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta) \text{ and } \mu^L_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)\right]$ along with the Gaussian approximation $\left[\hat{\mu}_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)\right]$. The approximation error at θ is less than or equal to the larger of $\left[\mu^U_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta) - \hat{\mu}_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)\right]$ and $\left[\hat{\mu}_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta) - \mu^L_{\sqcap_{i=1}^{n}\tilde{F}_i}(\theta)\right]$.

Observe, from Fig. C.3, that the approximation error is less in the high membership regions than in the low membership regions, and is equal to zero at the point having unity membership (this point is equal to the product of the centers of all the participating Gaussians).

C.9 Proof of Theorem 2.4

We prove the theorem in two parts : (a) we prove that $\alpha_i\tilde{F}_i + \beta$ is a Gaussian fuzzy number with mean $\alpha_i m_i + \beta$ and standard deviation $|\alpha_i\sigma_i|$; and (b) we prove that $\sum_{i=1}^{n}\tilde{F}_i$ is a Gaussian fuzzy number with mean $\sum_{i=1}^{n} m_i$ and standard deviation Σ'', where

$$\Sigma'' = \begin{cases} \sqrt{\sum_{i=1}^{n}\sigma_i^2} & , \text{ if product } t\text{-norm is used} \\ \sum_{i=1}^{n}\sigma_i & , \text{ if minimum } t\text{-norm is used} \end{cases} \qquad \text{(C.173)}$$

(a) Consider

$$\tilde{F}_i = \int_v e^{-\frac{1}{2}\left(\frac{v-m_i}{\sigma_i}\right)^2} \Big/ v \qquad \text{(C.174)}$$

Multiplying $\tilde{F}_i$ by a constant $\alpha_i(= 1/\alpha_i)$ yields [see Section 2.4.1]

$$\begin{aligned}\alpha_i\tilde{F}_i &= \int_v \left[e^{-\frac{1}{2}\left(\frac{v-m_i}{\sigma_i}\right)^2} \star 1\right]\Big/(\alpha_i v)\\ &= \int_v e^{-\frac{1}{2}\left(\frac{v-m_i}{\sigma_i}\right)^2}\Big/(\alpha_i v) \qquad \text{(C.175)}\end{aligned}$$

Now, adding a crisp constant $\beta(= 1/\beta)$ to $\alpha_i\tilde{F}_i$, we get [see Section 2.4.2]

$$\begin{aligned}\alpha_i\tilde{F}_i + \beta &= \int_v \left[e^{-\frac{1}{2}\left(\frac{v-m_i}{\sigma_i}\right)^2} \star 1\right]\Big/(\alpha_i v + \beta)\\ &= \int_v e^{-\frac{1}{2}\left(\frac{v-m_i}{\sigma_i}\right)^2}\Big/(\alpha_i v + \beta) \qquad \text{(C.176)}\end{aligned}$$

Let $\alpha_i v + \beta = v'$; this gives $v = (v' - \beta)/\alpha_i$, which when substituted into (C.176), leads to

$$\alpha_i\tilde{F}_i + \beta = \int_{v'} \exp\left\{-\frac{1}{2}\left[\frac{(\frac{v'-\beta}{\alpha_i}) - m_i}{\sigma_i}\right]^2\right\}\Big/ v'$$

$$= \int_{v'} \exp\Big\{ -\frac{1}{2}\Big[\frac{v' - (\alpha_i m_i + \beta)}{\alpha_i \sigma_i}\Big]^2 \Big\} \Big/ v' \qquad (C.177)$$

which shows that $\alpha_i \tilde{F}_i + \beta$ is a Gaussian fuzzy number with mean $\alpha_i m_i + \beta$ and standard deviation $|\alpha_i \sigma_i|$. Note that this result does not depend on the kind of t-norm used, since α_i and β are crisp numbers.

(b) Consider $\tilde{F}_1$ and $\tilde{F}_2$, with means m_1 and m_2 and standard deviations σ_1 and σ_2 respectively. The sum of these two fuzzy numbers can be expressed as [see Section 2.4.2]

$$\tilde{F}_1 + \tilde{F}_2 = \int_{v \in \tilde{F}_1} \int_{w \in \tilde{F}_2} e^{-\frac{1}{2}\left(\frac{v - m_1}{\sigma_1}\right)^2} \star e^{-\frac{1}{2}\left(\frac{w - m_2}{\sigma_2}\right)^2} /[v + w] \qquad (C.178)$$

where $\star$ indicates the chosen t-norm.

(i) Product t-norm : In this case, (C.178) reduces to

$$\tilde{F}_1 + \tilde{F}_2 = \int_{v \in \tilde{F}_1} \int_{w \in \tilde{F}_2} e^{-\frac{1}{2}\left(\frac{v - m_1}{\sigma_1}\right)^2} e^{-\frac{1}{2}\left(\frac{w - m_2}{\sigma_2}\right)^2} /[v + w] \qquad (C.179)$$

If θ is an element of $\tilde{F}_1 + \tilde{F}_2$, the membership grade of θ in $\tilde{F}_1 + \tilde{F}_2$ can be obtained by considering all the $\{v, w\}$ pairs such that $v \in \tilde{F}_1$ and $w \in \tilde{F}_2$ and $v + w = \theta$, multiplying the memberships of v and w in every pair, and, choosing the maximum of all these membership products. In other words,

$$\begin{aligned} \mu_{(\tilde{F}_1 + \tilde{F}_2)}(\theta) &= \sup_v e^{-\frac{1}{2}\left(\frac{v - m_1}{\sigma_1}\right)^2} e^{-\frac{1}{2}\left[\frac{(\theta - v) - m_2}{\sigma_2}\right]^2} \\ &= \sup_v e^{-\frac{1}{2}\left[\left(\frac{v - m_1}{\sigma_1}\right)^2 + \left(\frac{(\theta - v) - m_2}{\sigma_2}\right)^2\right]} \qquad (C.180) \end{aligned}$$

Let us call the expression in the square bracket in the exponent of (C.180) $J(v)$, i.e.,

$$J(v) = \Big(\frac{v - m_1}{\sigma_1}\Big)^2 + \Big[\frac{(\theta - v) - m_2}{\sigma_2}\Big]^2 \qquad (C.181)$$

The value of v that maximizes the exponent on the RHS of (C.180) can be obtained by minimizing $J(v)$. Note that J is convex ($J'' = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} > 0$), so equating the first derivative of J to zero (assuming the minimum is reached at v^*), we get

$$\begin{aligned} 2\Big(\frac{v^* - m_1}{\sigma_1}\Big)\Big(\frac{1}{\sigma_1}\Big) + 2\Big[\frac{(\theta - v^*) - m_2}{\sigma_2}\Big]\Big(-\frac{1}{\sigma_2}\Big) &= 0 \\ \frac{v^* - m_1}{\sigma_1^2} + \frac{v^* - (\theta - m_2)}{\sigma_2^2} &= 0 \\ v^*\Big(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\Big) &= \frac{m_1}{\sigma_1^2} + \frac{\theta - m_2}{\sigma_2^2} \\ v^* &= \frac{m_1 \sigma_2^2 + (\theta - m_2)\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \qquad (C.182) \end{aligned}$$

Substituting (C.182) into (C.181), we get

$$\inf_v J(v) = \left[\frac{\theta - (m_1 + m_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right]^2 \tag{C.183}$$

Substituting (C.183) into (C.180), we get

$$\mu_{(\tilde{F}_1 + \tilde{F}_2)}(\theta) = e^{-\frac{1}{2}\left[\frac{\theta - (m_1 + m_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right]^2} \tag{C.184}$$

This result generalizes easily to the case of more than two Gaussians. Let $m_{12} = m_1 + m_2$ and $\sigma_{12} = \sqrt{\sigma_1^2 + \sigma_2^2}$, so that $\tilde{F}_1 + \tilde{F}_2$ is a Gaussian with mean m_{12} and standard deviation σ_{12}. Now, if a third Gaussian fuzzy set $\tilde{F}_3$, with mean m_3 and standard deviation σ_3, adds to this sum, the mean and standard deviation of the resulting Gaussian are

$$m_{123} = m_{12} + m_3 = m_1 + m_2 + m_3 \tag{C.185}$$

$$\sigma_{123} = \sqrt{\sigma_{12}^2 + \sigma_3^2} = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} \tag{C.186}$$

Generalizing the result to the case of n Gaussians, we see that $\sum_{i=1}^n \tilde{F}_i$ is a Gaussian fuzzy number with mean $\sum_{i=1}^n m_i$ and standard deviation $\sqrt{\sum_{i=1}^n \sigma_i^2}$.
(ii) Minimum *t*-norm : In this case, (C.178) reduces to

$$\tilde{F}_1 + \tilde{F}_2 = \int_{v \in \tilde{F}_1} \int_{w \in \tilde{F}_2} e^{-\frac{1}{2}\left(\frac{v - m_1}{\sigma_1}\right)^2} \wedge e^{-\frac{1}{2}\left(\frac{w - m_2}{\sigma_2}\right)^2} / [v + w] \tag{C.187}$$

If θ is an element of $\tilde{F}_1 + \tilde{F}_2$, the membership grade of θ in $\tilde{F}_1 + \tilde{F}_2$ can be obtained by considering all the $\{v, w\}$ pairs such that $v \in \tilde{F}_1$ and $w \in \tilde{F}_2$ and $v + w = \theta$, finding the minimum of the memberships of v and w in every pair, and, choosing the maximum of all these minimums. In other words,

$$\mu_{(\tilde{F}_1 + \tilde{F}_2)}(\theta) = \sup_v \left[e^{-\frac{1}{2}\left(\frac{v - m_1}{\sigma_1}\right)^2} \wedge e^{-\frac{1}{2}\left[\frac{(\theta - v) - m_2}{\sigma_2}\right]^2}\right] \tag{C.188}$$

We make use of the fact that the supremum of the minimum of two Gaussians is reached at their point of intersection lying between their means. To solve for the point of intersection, we equate the equations of the two Gaussians.

$$\begin{aligned} e^{-\frac{1}{2}\left(\frac{v_* - m_1}{\sigma_1}\right)^2} &= e^{-\frac{1}{2}\left[\frac{(\theta - v_*) - m_2}{\sigma_2}\right]^2} \\ \Rightarrow \quad \left(\frac{v_* - m_1}{\sigma_1}\right)^2 &= \left[\frac{(\theta - v_*) - m_2}{\sigma_2}\right]^2 \\ \Rightarrow \quad \left(\frac{v_* - m_1}{\sigma_1}\right)^2 &= \left[\frac{(\theta - m_2) - v_*}{\sigma_2}\right]^2 \\ \Rightarrow \quad \frac{v_* - m_1}{\sigma_1} &= \pm\frac{(\theta - m_2) - v_*}{\sigma_2} \end{aligned} \tag{C.189}$$

The positive square-root on the RHS of (C.189) gives us the point of intersection lying between the means $[m_1$ and $(\theta - m_2)]$. Solving further, we find

$$v_*\left(\frac{1}{\sigma_1} + \frac{1}{\sigma_2}\right) = \frac{m_1}{\sigma_1} + \frac{(\theta - m_2)}{\sigma_2}$$

$$\Rightarrow \quad v_* \quad = \quad \frac{m_1\sigma_2 + (\theta - m_2)\sigma_1}{\sigma_1 + \sigma_2} \tag{C.190}$$

Since v_* is the point of intersection of the two Gaussians, it has the same membership in $\tilde{F}_1$ and $\tilde{F}_2$; therefore, membership grade of θ in $\tilde{F}_1 + \tilde{F}_2$ is

$$\begin{aligned} \mu_{(\tilde{F}_1+\tilde{F}_2)}(\theta) &= \exp\left\{-\frac{1}{2}\left(\frac{v_* - m_1}{\sigma_1}\right)^2\right\} \\ &= \exp\left\{-\frac{1}{2}\left[\frac{\frac{m_1\sigma_2+(\theta-m_2)\sigma_1}{\sigma_1+\sigma_2} - m_1}{\sigma_1}\right]^2\right\} \\ &= \exp\left\{-\frac{1}{2}\left(\frac{\theta - (m_1 + m_2)}{\sigma_1 + \sigma_2}\right)^2\right\} \end{aligned} \tag{C.191}$$

This result generalizes easily to the case of more than two Gaussians. Let $m_{12} = m_1 + m_2$ and $\sigma_{12} = \sigma_1 + \sigma_2$, so that $\tilde{F}_1 + \tilde{F}_2$ is a Gaussian with mean m_{12} and standard deviation σ_{12}. Now, if a third Gaussian fuzzy set $\tilde{F}_3$, with mean m_3 and standard deviation σ_3, adds to this sum, the mean and standard deviation of the resulting Gaussian are

$$m_{123} = m_{12} + m_3 = m_1 + m_2 + m_3 \tag{C.192}$$

$$\sigma_{123} = \sigma_{12} + \sigma_3 = \sigma_1 + \sigma_2 + \sigma_3 \tag{C.193}$$

Generalizing to the case of n Gaussians, we see that $\sum_{i=1}^{n} \tilde{F}_i$ is a Gaussian fuzzy number with mean $\sum_{i=1}^{n} m_i$ and standard deviation $\sum_{i=1}^{n} \sigma_i$.

Combining parts (a) and (b), we get the desired result. □

C.10 Proof of the Claim in Example 2.5

Consider an interval type-2 set resulting from a Gaussian type-1 set, whose mean is uncertain in the interval $[m_1, m_2]$, and whose standard deviation is σ. We show that, if $(m_2 - m_1)$ is small compared to σ, the centroid of this type-2 set is approximately an interval type-1 set with domain $[m_1, m_2]$.

Figure C.4 (a) depicts an example of such a set, $\tilde{\tilde{A}}$. The type-1 Gaussians with centers m_1 and m_2 and standard deviation σ are also shown. From the discussion in Example 2.5, we know that : (1) the centroid of $\tilde{\tilde{A}}$ is some interval, $[c_l, c_r]$, which contains $[m_1, m_2]$, and (2) c_l is the centroid of an embedded type-1 set, $\tilde{A}_l$ [see Fig. C.4 (b)] whose membership function assigns the highest possible memberships to all the points to the left of c_l, and the lowest possible memberships to all the points to the right of c_l. We now focus on the left end-point, c_l, of this interval. The discussion about c_r is similar.

We show that $c_l > m_1 - \Delta$, where

$$\Delta = \frac{m_2 - m_1}{2} \tag{C.194}$$

Consider the embedded type-1 set, $\tilde{A}_1$, shown in Fig. C.4 (c). The membership function of this embedded set assigns the highest possible memberships to the points less than $(m_1 - \Delta)$, and the lowest possible membership to the points greater than $(m_1 - \Delta)$. It is easy to verify that the area under this curve, to the left of $(m_1 - \Delta)$, is the same as the area under it to the right of $(m_1 + \Delta)$ [see Fig. C.4 (c)]; therefore, the centroid of this curve, c', lies between $(m_1 - \Delta)$ and $(m_1 + \Delta)$. Now consider any other embedded set, $\tilde{A}_2$, which assigns the highest possible memberships to all the points to the left of some point m' and the lowest possible memberships to all the points to the right of m', where $m' < m_1 - \Delta$. It is easy to see that the centroid of $\tilde{A}_2$ will lie to the right of

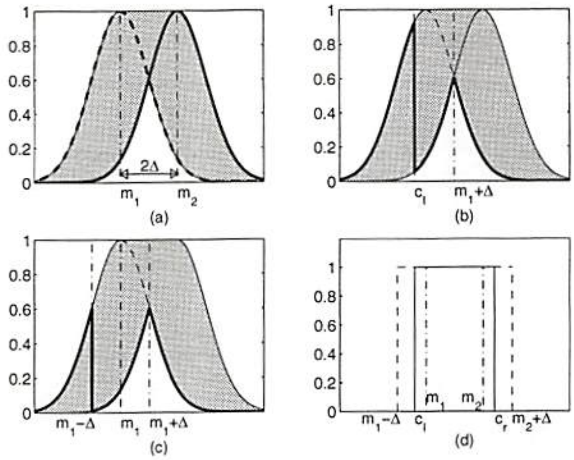


Figure C.4: Figures for Appendix C.10. (a) An interval type-2 set $\tilde{\tilde{A}}$ resulting from a Gaussian type-1 set with standard deviation equal to σ and mean uniformly uncertain in the interval $[m_1, m_2]$. The Gaussian type-1 sets having standard deviations σ and means m_1 (thick dashed line) and m_2 (thick solid line) are also shown. (b) The embedded type-1 set, $\tilde{A}_l$, whose centroid equals c_l is shown with a thick solid line. (c) The embedded type-1 set, $\tilde{A}_1$, whose membership function assigns the highest possible memberships to the points to the left of $m_1 - \Delta$ and the lowest possible memberships to the points to the right of $m_1 - \Delta$, where $\Delta = (m_2 - m_1)/2$, is shown with a thick solid line. The Gaussian with center m_1 is shown with a thin dashed line. (d) The centroid of $\tilde{\tilde{A}}$ is a crisp set with domain $[c_l, c_r]$, where $m_1 - \Delta < c_l \leq m_1$ and $m_2 \leq c_r < m_2 + \Delta$.

c'. We, therefore, conclude that $c_l > m_1 - \Delta$. Using this result with the fact that $[c_l, c_r]$ includes $[m_1, m_2]$, we get

$$m_1 - \Delta < c_l \leq m_1 \tag{C.195}$$

Obviously, as $\Delta \to 0$ [i.e., as $(m_2 - m_1) \to 0$], $c_l \to m_1$. We will make use of (C.195) in the sequel, to show the dependence of c_l on σ.

To show the effect of σ on c_l, we write the expression for the centroid of $\tilde{A}_l$ [see Fig. C.4 (b)]. Let $G(x; m, \sigma) = \exp\{-\frac{1}{2}(\frac{x-m}{\sigma})^2\}$; then

$$\begin{aligned} c_l = \tilde{C}_{\tilde{\tilde{A}}_l} &= \frac{\int_{-\infty}^{c_l} xG(x; m_1, \sigma)dx + \int_{c_l}^{m_1+\Delta} xG(x; m_2, \sigma)dx + \int_{m_1+\Delta}^{\infty} xG(x; m_1, \sigma)dx}{\int_{-\infty}^{c_l} G(x; m_1, \sigma)dx + \int_{c_l}^{m_1+\Delta} G(x; m_2, \sigma)dx + \int_{m_1+\Delta}^{\infty} G(x; m_1, \sigma)dx} \\ &= \frac{\int_{-\infty}^{\infty} xG(x; m_1, \sigma)dx - \int_{c_l}^{m_1+\Delta} x[G(x; m_1, \sigma) - G(x; m_2, \sigma)]dx}{\int_{-\infty}^{\infty} G(x; m_1, \sigma)dx - \int_{c_l}^{m_1+\Delta} [G(x; m_1, \sigma) - G(x; m_2, \sigma)]dx} \\ &= \frac{\sqrt{2\pi}\sigma m_1 - I_1}{\sqrt{2\pi}\sigma - I_2} \end{aligned} \tag{C.196}$$

where

$$I_1 = \int_{c_l}^{m_1+\Delta} x[G(x; m_1, \sigma) - G(x; m_2, \sigma)]dx, \tag{C.197}$$

$$I_2 = \int_{c_l}^{m_1+\Delta} [G(x; m_1, \sigma) - G(x; m_2, \sigma)]dx \tag{C.198}$$

and, we have made use of the facts that : (1) the area under a Gaussian having standard deviation σ is $\sqrt{2\pi}\sigma$, and (2) since the centroid of a Gaussian is equal to its mean,

$$\int_{-\infty}^{\infty} xG(x; m_1, \sigma)dx = m_1 \int_{-\infty}^{\infty} G(x; m_1, \sigma)dx = \sqrt{2\pi}\sigma m_1 \tag{C.199}$$

From (C.195) and (C.196), we have

$$0 \leq m_1 - c_l = \frac{I_1 - m_1 I_2}{\sqrt{2\pi}\sigma - I_2} \tag{C.200}$$

Observe, from (C.197) and (C.198), that

$$I_1 \leq \int_{c_l}^{m_1+\Delta} (m_1 + \Delta)[G(x; m_1, \sigma) - G(x; m_2, \sigma)]dx = (m_1 + \Delta)I_2 \tag{C.201}$$

Using (C.201) with (C.200), we get

$$0 \leq m_1 - c_l \leq \frac{\Delta I_2}{\sqrt{2\pi}\sigma - I_2} \tag{C.202}$$

Now, observe, from (C.198), (C.195) and the fact that $G(x; m_1, \sigma) > G(x; m_2, \sigma)$ for $x < m_1 + \Delta$, that

$$I_2 < \int_{m_1-\Delta}^{m_1+\Delta} [G(x; m_1, \sigma) - G(x; m_2, \sigma)]dx \tag{C.203}$$

Observe also that, for $x \in [m_1 - \Delta, m_1 + \Delta]$,

$$\begin{aligned} 0 &\leq G(x; m_1, \sigma) - G(x; m_2, \sigma) \\ &= G(x; m_1, \sigma)\Big[1 - \exp\Big\{-\frac{1}{2}\Big[\Big(\frac{x - m_2}{\sigma}\Big)^2 - \Big(\frac{x - m_1}{\sigma}\Big)^2\Big]\Big\}\Big] \\ &= G(x; m_1, \sigma)\Big[1 - \exp\Big\{-\frac{2\Delta}{\sigma^2}(m_1 + \Delta - x)\Big\}\Big] \\ &\leq G(x; m_1, \sigma)\Big[1 - \exp\Big\{-\Big(\frac{2\Delta}{\sigma}\Big)^2\Big\}\Big] \end{aligned} \tag{C.204}$$

In the last step of (C.204), we have made use of the fact that $x \geq m_1 - \Delta$.

Using (C.204) with (C.203), we have

$$\begin{aligned} I_2 &< \Big[1 - \exp\Big\{-\Big(\frac{2\Delta}{\sigma}\Big)^2\Big\}\Big] \int_{m_1-\Delta}^{m_1+\Delta} G(x; m_1, \sigma)dx \\ &\leq \Big[1 - \exp\Big\{-\Big(\frac{2\Delta}{\sigma}\Big)^2\Big\}\Big] \int_{-\infty}^{\infty} G(x; m_1, \sigma)dx \\ &= \Big[1 - \exp\Big\{-\Big(\frac{2\Delta}{\sigma}\Big)^2\Big\}\Big]\sqrt{2\pi}\sigma \end{aligned} \tag{C.205}$$

Observe that

$$\begin{aligned} \sqrt{2\pi}\sigma\Big[1 - \exp\Big\{-\Big(\frac{2\Delta}{\sigma}\Big)^2\Big\}\Big] &= \sqrt{2\pi}\sigma\Big[1 - \Big(1 + \sum_{i=1}^{\infty}(-1)^i \frac{1}{i!}\Big(\frac{2\Delta}{\sigma}\Big)^{2i}\Big)\Big] \\ &= \sqrt{2\pi}\sigma\Big[\sum_{i=1}^{\infty}(-1)^{i+1} \frac{1}{i!}\Big(\frac{2\Delta}{\sigma}\Big)^{2i}\Big)\Big] \end{aligned}$$

$$= 2\sqrt{2\pi}\Delta\Big[\sum_{i=1}^{\infty}(-1)^{i+1}\frac{1}{i!}\Big(\frac{2\Delta}{\sigma}\Big)^{2i-1}\Big)\Big] \quad \text{(C.206)}$$

It is clear, from (C.205) and (C.206), that, as $2\Delta/\sigma = (m_2 - m_1)/\sigma \to 0$, $I_2 \to 0$; hence, from (C.202), $c_l \to m_1$.

In a similar manner, it can be shown that $m_2 \leq c_r < m_2 + \Delta$, and as $(m_2 - m_1)/\sigma \to 0$, $c_r \to m_2$ [see Fig. C.4 (d)]. □

C.11 Proof of Theorem 2.5

Let $G(x; m, \sigma) \triangleq \exp\Big\{-\frac{1}{2}\Big(\frac{x-m}{\sigma}\Big)^2\Big\}$. Equation (2.75) can now be rewritten as

$$\tilde{Y} = \int_{z_1}\cdots\int_{z_M}\int_{w_1}\cdots\int_{w_M} \mathcal{T}_{i=1}^{M} G(z_l; m_l, \sigma_l) \star \mathcal{T}_{l=1}^{M} G(w_l; h_l, \Delta_l) \Big/ \frac{\sum_{l=1}^{M} w_l z_l}{\sum_{l=1}^{M} w_l} \quad \text{(C.207)}$$

where $\tilde{Y} \triangleq \tilde{Y}(\tilde{Z}_1, \ldots, \tilde{Z}_M, \tilde{W}_1, \ldots, \tilde{W}_M)$.

If we let $\gamma_l = z_l - m_l$ and $\delta_l = w_l - h_l$ for $l = 1, \ldots, M$, (C.207) becomes

$$\tilde{Y} = \int_{\gamma_1}\cdots\int_{\gamma_M}\int_{\delta_1}\cdots\int_{\delta_M} \mathcal{T}_{l=1}^{M} G(\gamma_l; 0, \sigma_l) \star \mathcal{T}_{l=1}^{M} G(\delta_l; 0, \Delta_l) \Big/ \frac{\sum_{l=1}^{M}(h_l + \delta_l)(m_l + \gamma_l)}{\sum_{l=1}^{M}(h_l + \delta_l)} \quad \text{(C.208)}$$

Theoretically, each δ_l can take any value in the interval $[0, 1]$; however, only those values which lie within 2 or 3 standard deviations of h_l contribute significantly to the union in (C.208); therefore, we assume that each δ_l takes values between $\pm k\Delta_l$, where $k = 2$ or 3. Similarly, we assume that each γ_l takes values between $\pm k\sigma_l$.

The term to the right of the slash in (C.208) can be rewritten as

$$\begin{aligned} \frac{\sum_l w_l z_l}{\sum_l w_l} &= \frac{\sum_l (h_l + \delta_l)(m_l + \gamma_l)}{\sum_l (h_l + \delta_l)} \\ &= \frac{\sum_l h_l m_l + \sum_l h_l \gamma_l + \sum_l \delta_l m_l + \sum_l \delta_l \gamma_l}{\sum_l h_l + \sum_l \delta_l} \end{aligned} \quad \text{(C.209)}$$

where the limits on each sum are from 1 to M.

In what follows, we express the term on the RHS of (C.209) as an affine combination of Gaussian γ_l's and δ_l's so that we can make use of Theorem 2.4 to find an (approximate) expression for $\tilde{Y}$ in (C.207). We expand the denominator of (C.209) by first rewriting it as

$$\frac{1}{\sum_l h_l + \sum_l \delta_l} = \frac{1}{\sum_l h_l}\left(\frac{1}{1 + \frac{\sum_l \delta_l}{\sum_l h_l}}\right) \quad \text{(C.210)}$$

If $|\sum_l \delta_l| < \sum_l h_l$ (since δ_l varies between $-k\Delta_l$ and $k\Delta_l$, this is equivalent to assuming that $k\sum_l \Delta_l < \sum_l h_l$), we can express the parenthetical term on the RHS of (C.210), as

$$\frac{1}{1 + \frac{\sum_l \delta_l}{\sum_l h_l}} = 1 - \Big(\frac{\sum_l \delta_l}{\sum_l h_l}\Big) + \Big(\frac{\sum_l \delta_l}{\sum_l h_l}\Big)^2 - \Big(\frac{\sum_l \delta_l}{\sum_l h_l}\Big)^3 + \cdots \quad \text{(C.211)}$$

where we have made use of the identity

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots \quad \text{if} \quad |x| < 1 \tag{C.212}$$

If

$$k\frac{\sum_l \Delta_l}{\sum_l h_l} \ll 1\,, \tag{C.213}$$

which means that

$$\frac{|\sum_l \delta_l|}{\sum_l h_l} \ll 1\,, \tag{C.214}$$

we can ignore powers of $\sum_l \delta_l \Big/ \sum_l h_l$ greater than 1 in (C.211). This gives us

$$\frac{1}{1 + \frac{\sum_l \delta_l}{\sum_l h_l}} \approx 1 - \Big(\frac{\sum_l \delta_l}{\sum_l h_l}\Big) \tag{C.215}$$

Substituting (C.215) into (C.210), we get

$$\frac{1}{\sum_l h_l + \sum_l \delta_l} \approx \frac{1}{\sum_l h_l}\Big(1 - \frac{\sum_l \delta_l}{\sum_l h_l}\Big) \tag{C.216}$$

Using (C.216) in (C.209), we get

$$\frac{\sum_l w_l z_l}{\sum_l w_l} \approx \frac{\sum_l h_l m_l + \sum_l h_l \gamma_l + \sum_l \delta_l m_l + \sum_l \delta_l \gamma_l}{\sum_l h_l}\Big(1 - \frac{\sum_l \delta_l}{\sum_l h_l}\Big) \tag{C.217}$$

Ignoring all the terms containing powers of $\sum_l \delta_l / \sum_l h_l$ higher than 1, we get

$$\begin{aligned}
\frac{\sum_l w_l z_l}{\sum_l w_l} &\approx \frac{\sum_l h_l m_l}{\sum_l h_l}\Big(1 - \frac{\sum_l \delta_l}{\sum_l h_l}\Big) + \frac{\sum_l h_l \gamma_l}{\sum_l h_l}\Big(1 - \frac{\sum_l \delta_l}{\sum_l h_l}\Big) + \frac{\sum_l \delta_l m_l}{\sum_l h_l} + \frac{\sum_l \delta_l \gamma_l}{\sum_l h_l} \\
&= \frac{\sum_l h_l m_l}{\sum_l h_l} - \frac{\sum_l \delta_l}{\sum_l h_l}\Big(\frac{\sum_l h_l m_l}{\sum_l h_l}\Big) + \frac{\sum_l h_l \gamma_l}{\sum_l h_l} \\
&\quad - \frac{\sum_l \delta_l}{\sum_l h_l}\Big(\frac{\sum_l h_l \gamma_l}{\sum_l h_l}\Big) + \frac{\sum_l \delta_l m_l}{\sum_l h_l} + \frac{\sum_l \delta_l \gamma_l}{\sum_l h_l}
\end{aligned} \tag{C.218}$$

Let

$$\mathcal{M} = \frac{\sum_l h_l m_l}{\sum_l h_l}\,; \tag{C.219}$$

then (C.218) can be rewritten as

$$\frac{\sum_l w_l z_l}{\sum_l w_l} \approx \mathcal{M} - \mathcal{M}\frac{\sum_l \delta_l}{\sum_l h_l} + \frac{\sum_l \delta_l m_l}{\sum_l h_l} + \frac{\sum_l h_l \gamma_l}{\sum_l h_l} - \frac{\sum_l \delta_l}{\sum_l h_l}\Big(\frac{\sum_l h_l \gamma_l}{\sum_l h_l}\Big) + \frac{\sum_l \delta_l \gamma_l}{\sum_l h_l} \tag{C.220}$$

Next we focus on the last two terms in (C.220). Observe that

$$\Big|\frac{\sum_l \delta_l \gamma_l}{\sum_l h_l}\Big| \le \Big|\max_l \gamma_l \Big(\frac{\sum_l \delta_l}{\sum_l h_l}\Big)\Big| \tag{C.221}$$

Since γ_l takes values between $\pm k\sigma_l$, (C.221) is equivalent to

$$\left|\frac{\sum_l \delta_l \gamma_l}{\sum_l h_l}\right| \le k \max_l \sigma_l \left|\frac{\sum_l \delta_l}{\sum_l h_l}\right| \tag{C.222}$$

Similarly,

$$\left|\frac{\sum_l h_l \gamma_l}{\sum_l h_l}\right| \le k \max_l \sigma_l \tag{C.223}$$

Consequently,

$$\left| -\frac{\sum_l \delta_l}{\sum_l h_l}\Big(\frac{\sum_l h_l\gamma_l}{\sum_l h_l}\Big) + \frac{\sum_l \delta_l\gamma_l}{\sum_l h_l}\right| \le \left|\frac{\sum_l \delta_l}{\sum_l h_l}\Big(\frac{\sum_l h_l\gamma_l}{\sum_l h_l}\Big)\right| + \left|\frac{\sum_l \delta_l\gamma_l}{\sum_l h_l}\right| \le 2k \max_l \sigma_l \left|\frac{\sum_l \delta_l}{\sum_l h_l}\right| \tag{C.224}$$

Observe, from (C.220) and (C.224), that if condition (C.214) is satisfied, we can ignore the last two terms on the RHS of (C.220) in comparison with $\sum_l h_l\gamma_l \big/ \sum_l h_l$, which, according to (C.223), takes values in $\pm k \max_l \sigma_l$. Doing this gives us

$$\begin{aligned} \frac{\sum_l w_l z_l}{\sum_l w_l} &\approx \mathcal{M} - \mathcal{M}\frac{\sum_l \delta_l}{\sum_l h_l} + \frac{\sum_l \delta_l m_l}{\sum_l h_l} + \frac{\sum_l h_l\gamma_l}{\sum_l h_l} \\ &= \sum_{l=1}^{M}\Big[\gamma_l\Big(\frac{h_l}{\sum_l h_l}\Big) + \delta_l\Big(\frac{m_l - \mathcal{M}}{\sum_l h_l}\Big)\Big] + \mathcal{M} \end{aligned} \tag{C.225}$$

Using (C.225), (C.208) can be rewritten as

$$\begin{aligned} \tilde{Y} \approx &\int_{\gamma_1}\cdots\int_{\gamma_M}\int_{\delta_1}\cdots\int_{\delta_M} \mathcal{T}_{l=1}^{M} G(\gamma_l;0,\sigma_l) \star \mathcal{T}_{l=1}^{M} G(\delta_l;0,\Delta_l) \Big/ \\ &\sum_{l=1}^{M}\Big[\gamma_l\Big(\frac{h_l}{\sum_l h_l}\Big) + \delta_l\Big(\frac{m_l - \mathcal{M}}{\sum_l h_l}\Big)\Big] + \mathcal{M} \end{aligned} \tag{C.226}$$

Recall that $\gamma_l = z_l - m_l$ and $\delta_l = w_l - h_l$. Let

$$\underline{\tilde{Z}}_l = \tilde{Z}_l - m_l \quad \text{for} \quad l = 1, \ldots, M \tag{C.227}$$

and

$$\underline{\tilde{W}}_l = \tilde{W}_l - h_l \quad \text{for} \quad l = 1, \ldots, M\,; \tag{C.228}$$

so that each $\underline{\tilde{Z}}_l$ is a type-1 Gaussian fuzzy number with zero mean and standard deviation equal to σ_l, and each $\underline{\tilde{W}}_l$ is also a type-1 Gaussian fuzzy number with zero mean and standard deviation Δ_l. Observe that the RHS of (C.226) is equal to $\sum_{l=1}^{M}\Big[\underline{\tilde{Z}}_l\Big(\frac{h_l}{\sum_l h_l}\Big) + \underline{\tilde{W}}_l\Big(\frac{m_l-\mathcal{M}}{\sum_l h_l}\Big)\Big] + \mathcal{M}$ (see Section 2.4.2), i.e.,

$$\tilde{Y} \approx \sum_{l=1}^{M}\Big[\underline{\tilde{Z}}_l\Big(\frac{h_l}{\sum_l h_l}\Big) + \underline{\tilde{W}}_l\Big(\frac{m_l - \mathcal{M}}{\sum_l h_l}\Big)\Big] + \mathcal{M} \tag{C.229}$$

The result in Theorem 2.5 follows by applying Theorem 2.4 to (C.229), using the fact that all $\underline{\tilde{Z}}_l$'s and $\underline{\tilde{W}}_l$'s have zero means. □

Comment 1 : When all $\Delta_l = 0$, there is only one source of fuzziness in $\sum_l w_l z_l / \sum_l w_l$, namely, the $\tilde{Z}_l$'s. In this case, (2.75) reduces to

$$\tilde{Y}(\tilde{Z}_1, \ldots, \tilde{Z}_M, h_1, \ldots, h_M) = \int_{z_1} \cdots \int_{z_M} \mathcal{T}_{l=1}^M G(z_l; m_l, \sigma_l) \Bigg/ \frac{\sum_{l=1}^M h_l z_l}{\sum_{l=1}^M h_l} \tag{C.230}$$

Again, letting $\gamma_l = z_l - m_l$ for $l = 1, \ldots, M$, we have

$$\begin{aligned}
\tilde{Y}(\tilde{Z}_1, \ldots, \tilde{Z}_M, h_1, \ldots, h_M) &= \int_{\gamma_1} \cdots \int_{\gamma_M} \mathcal{T}_{l=1}^M G(\gamma_l; 0, \sigma_l) \Bigg/ \frac{\sum_{l=1}^M h_l (m_l + \gamma_l)}{\sum_{l=1}^M h_l} \\
&= \int_{\gamma_1} \cdots \int_{\gamma_M} \mathcal{T}_{l=1}^M G(\gamma_l; 0, \sigma_l) \Bigg/ \\
&\qquad \left[\frac{\sum_{l=1}^M h_l m_l}{\sum_{l=1}^M h_l} + \sum_{l=1}^M \gamma_l \Big(\frac{h_l}{\sum_{l=1}^M h_l} \Big) \right] \\
&= \int_{\gamma_1} \cdots \int_{\gamma_M} \mathcal{T}_{l=1}^M G(\gamma_l; 0, \sigma_l) \Bigg/ \left[\mathcal{M} + \sum_{l=1}^M \gamma_l \Big(\frac{h_l}{\sum_{l=1}^M h_l} \Big) \right] \\
&= \sum_{l=1}^M \underline{\tilde{W}}_l \Big(\frac{h_l}{\sum_{l=1}^M h_l} \Big) + \mathcal{M}
\end{aligned} \tag{C.231}$$

where $\underline{\tilde{W}}_l$ is as defined in (C.228) and $\mathcal{M}$ is as in (C.219). Applying Theorem 2.4 to (C.231), it follows that $\tilde{Y}(\tilde{Z}_1, \ldots, \tilde{Z}_M, h_1, \ldots, h_M)$ is a Gaussian type-1 set with mean $\mathcal{M}$ and standard deviation Σ, where

$$\Sigma = \begin{cases} \frac{\sqrt{\sum_{l=1}^M (h_l \sigma_l)^2}}{\sum_{l=1}^M h_l} &, \text{ if product } t\text{-norm is used} \\ \frac{\sum_{l=1}^M (h_l \sigma_l)}{\sum_{l=1}^M h_l} &, \text{ if minimum } t\text{-norm is used} \end{cases} \tag{C.232}$$

Observe that this result is exact and it can also be obtained by substituting $\Delta_l = 0$ $(l = 1, \ldots, M)$ in (2.77).

Comment 2 : It is not very easy to find an expression for the error between the approximation in Theorem 2.5 and the true set $\tilde{Y}$; however, we can find bounds on the domain of $\tilde{Y}$ easily.

Recall Eq. (C.208). If each δ_l varies between $\pm k\Delta_l$ and each γ_l varies between $\pm k\sigma_l$, the term to the right of the slash can be bounded as

$$\frac{\sum_l (h_l - k\Delta_l)(m_l - k\sigma_l)}{\sum_l (h_l + k\Delta_l)} \leq \frac{\sum_l (h_l + \delta_l)(m_l + \gamma_l)}{\sum_l (h_l + \delta_l)} \leq \frac{\sum_l (h_l + k\Delta_l)(m_l + k\sigma_l)}{\sum_l (h_l - k\Delta_l)} \tag{C.233}$$

Let $k_1 = \max_l[\Delta_l / h_l]$ for $h_l \neq 0$ and let $k_2 = \max_l[\sigma_l / m_l]$, assuming $m_l > 0$ $(l = 1, \ldots, M)$. From (C.233), we have

$$\begin{aligned}
\frac{\sum_l (h_l - kk_1 h_l)(m_l - kk_2 m_l)}{\sum_l (h_l + kk_1 h_l)} &\leq \frac{\sum_l (h_l + \delta_l)(m_l + \gamma_l)}{\sum_l (h_l + \delta_l)} \leq \\
&\frac{\sum_l (h_l + kk_1 h_l)(m_l + kk_2 m_l)}{\sum_l (h_l - kk_1 h_l)} \\
\Rightarrow \quad \mathcal{M}\Big[\frac{(1 - kk_1)(1 - kk_2)}{(1 + kk_1)}\Big] &\leq \frac{\sum_l (h_l + \delta_l)(m_l + \gamma_l)}{\sum_l (h_l + \delta_l)} \leq
\end{aligned}$$

$$\mathcal{M}\Big[\frac{(1+kk_1)(1+kk_2)}{(1-kk_1)}\Big] \tag{C.234}$$

where $\mathcal{M} = \sum_l h_l m_l \Big/ \sum_l h_l$. Since the term to the right of the slash in (C.208) indicates a general point in $\tilde{Y}$, it is clear that the entire domain of $\tilde{Y}$ lies between the bounds in (C.234).

Though the bounds in (C.234) are generally very conservative, they illustrate how $\tilde{Y}(\tilde{Z}_1, \ldots, \tilde{Z}_M, \tilde{W}_1, \ldots, \tilde{W}_M)$ collapses to $y(z_1, \ldots, z_M, w_1, \ldots, l_M)$ [see (2.74) and (2.75)], when all the type-2 uncertainties collapse to type-1 uncertainties. When all the type-2 uncertainties disappear, $k_1 = k_2 = 0$ and the upper and lower bounds on $\tilde{Y}$ both equal $\mathcal{M}$, implying that the type-1 set $\tilde{Y}$ collapses to a crisp point equal to $\mathcal{M}$, i.e., $1/\mathcal{M}$ (since $\mu_{\tilde{Y}}(\mathcal{M}) = 1$).

Similar bounds can be obtained when all or some of the $m_l < 0$.

Comment 3 : Since the bounds in (C.234) enclose the entire domain of the type-reduced set between them, $\mathcal{M}$, the unity membership point in $\tilde{Y}$, and $C_{\tilde{Y}}$, the centroid of $\tilde{Y}$, also lie between these two bounds. The difference between $\mathcal{M}$ and $C_{\tilde{Y}}$ can, therefore, be loosely bounded as (assuming $m_l > 0$ for $l = 1, \ldots, M$).

$$\begin{aligned}
|\mathcal{M} - C_{\tilde{Y}}| &\leq \mathcal{M}\Big[\frac{(1+kk_1)(1+kk_2)}{(1-kk_1)} - \frac{(1-kk_1)(1-kk_2)}{(1+kk_1)}\Big] \\
\Rightarrow \quad |\mathcal{M} - C_{\tilde{Y}}| &\leq \mathcal{M}\Big[\frac{(1+kk_1)^2(1+kk_2) - (1-kk_1)^2(1-kk_2)}{(1-k^2k_1^2)}\Big] \\
\Rightarrow \quad |\mathcal{M} - C_{\tilde{Y}}| &\leq \mathcal{M}\Big[\frac{4kk_1 + 2kk_2(1+k^2k_1^2)}{(1-k^2k_1^2)}\Big]
\end{aligned} \tag{C.235}$$

Appendix D

Weighted Average of Interval Type-1 Sets

Here, we develop a computational procedure for computing the exact result of a weighted average of interval type-1 sets. This procedure can be used to compute the centroid of an interval type-2 set (Section 2.5) or to compute the type-reduced set for an interval type-2 FLS (Section 5.9).

Since every point in the domain of an interval type-1 set has a unity membership, we can describe an interval type-1 set just by its domain, e.g., an interval type-1 set with domain $[l, r]$ can be indicated as just $[l, r]$. If we let $m = (l + r)/2$ (mean) and $s = (r - l)/2$ (spread), we can also indicate an interval type-1 set as $[m - s, m + s]$.

D.1 Exact Result : Computational Procedure

Consider the weighted average of type-1 fuzzy sets [see (2.75)], which we reproduce here for convenience

$$\tilde{\mathrm{Y}}(\tilde{\mathrm{Z}}_1, \ldots, \tilde{\mathrm{Z}}_M, \tilde{\mathrm{W}}_1, \ldots, \tilde{\mathrm{W}}_M) = \int_{z_1} \cdots \int_{z_M} \int_{w_1} \cdots \int_{w_M} \mathcal{T}_{l=1}^{M} \mu_{\tilde{\mathrm{Z}}_l}(z_l) \star \mathcal{T}_{l=1}^{M} \mu_{\tilde{\mathrm{W}}_l}(w_l) \Bigg/ \frac{\sum_{l=1}^{M} w_l z_l}{\sum_{l=1}^{M} w_l} \tag{D.1}$$

If each $\tilde{\mathrm{Z}}_l$ and $\tilde{\mathrm{W}}_l$ $(l = 1, \ldots, M)$ is an interval type-1 set, then, using the fact that $\mu_{\tilde{\mathrm{Z}}_l}(z_l) = \mu_{\tilde{\mathrm{W}}_l}(w_l) = 1$, (D.1) can be rewritten as

$$Y(Z_1, \ldots, Z_M, W_1, \ldots, W_M) = \int_{z_1} \cdots \int_{z_M} \int_{w_1} \cdots \int_{w_M} 1 \Bigg/ \frac{\sum_{l=1}^{M} w_l z_l}{\sum_{l=1}^{M} w_l} \tag{D.2}$$

where we have omitted the tilde, since all the sets involved are crisp. We present an iterative procedure to compute the actual weighted average Y, when each Z_l in (D.2) is an interval type-1 set, having center c_l and spread s_l $(s_l \geq 0)$, and when each W_l is also an interval type-1 set with center h_l and spread Δ_l $(\Delta_l \geq 0)$ [we assume that $h_l \geq \Delta_l$, so that $w_l \geq 0$ for $l = 1, \ldots, M$].

We make the following observations :

1. Since each set in the weighted average on the RHS of (D.2) is an interval type-1 set, $Y(Z_1, \ldots, Z_M, W_1, \ldots, W_M)$ will also be an interval type-1 set, i.e., it will be a crisp set having an interval on the real line as its domain. So, to find $Y(Z_1, \ldots, Z_M, W_1, \ldots, W_M)$, we need to compute just the two end-points of this interval.

2. Since $w_l \geq 0$ for all l, the partial derivative $\partial Y/\partial z_k = w_k/\sum_l w_l \geq 0$; therefore, Y always increases with increasing z_k; and, for any combination of $\{w_1, \ldots, w_M\}$ chosen so that $w_l \in W_l$, $Y(Z_1, \ldots, Z_M, W_1, \ldots, W_M)$ is maximized when $z_l = c_l + s_l$ for $l = 1, \ldots, M$; and $Y(Z_1, \ldots, Z_M, W_1, \ldots, W_M)$ is minimized when $z_l = c_l - s_l$. The right end-point of the domain of $Y(Z_1, \ldots, Z_M, W_1, \ldots, W_M)$ is, therefore, obtained by maximizing $\left[\sum_l w_l(c_l + s_l)\right] / \left[\sum_l w_l\right]$ subject to the constraints $w_l \in W_l$ for $l = 1, \ldots, M$; and, the left end-point of the domain of $Y(Z_1, \ldots, Z_M, W_1, \ldots, W_M)$ is obtained by minimizing $\left[\sum_l w_l(c_l - s_l)\right] / \left[\sum_l w_l\right]$ subject to the constraints $w_l \in W_l$ for $l = 1, \ldots, M$.

From these two observations, it is clear that in order to compute $Y(Z_1, \ldots, Z_M, W_1, \ldots, W_M)$ we only need to consider the problem of optimizing (maximizing / minimizing) the weighted average

$$S(w_1, \ldots, w_M) = \frac{\sum_{l=1}^{M} z_l w_l}{\sum_{l=1}^{M} w_l} \tag{D.3}$$

subject to the constraints $w_l \in [h_l - \Delta_l, h_l + \Delta_l]$ for $l = 1, \ldots, M$, where, $h_l \geq \Delta_l$, so that $w_l \geq 0$, for $l = 1, \ldots, M$. As explained in observation (2) above, we set $z_l = c_l + s_l$ $(l = 1, \ldots, M)$, when maximizing S, and $z_l = c_l - s_l$ $(l = 1, \ldots, M)$, when minimizing S.

Differentiating $S(w_1, \ldots, w_M)$ w.r.t. w_k gives us

$$\begin{aligned}
\frac{\partial}{\partial w_k} S(w_1, \ldots, w_M) &= \frac{\partial}{\partial w_k}\left[\frac{\sum_{l=1}^{M} z_l w_l}{\sum_{l=1}^{M} w_l}\right] \\
&= \frac{\partial}{\partial w_k}\left[\frac{z_k w_k + \sum_{l \neq k} z_l w_l}{w_k + \sum_{l \neq k} w_l}\right] \\
&= \left[\frac{1}{w_k + \sum_{l \neq k} w_l}\right](z_k) \\
&\qquad + \left(z_k w_k + \sum_{l \neq k} z_l w_l\right)\left[\frac{-1}{\left(w_k + \sum_{l \neq k} w_l\right)^2}\right] \\
&= \frac{z_k}{\sum_{l=1}^{M} w_l} - \frac{\sum_{l=1}^{M} z_l w_l}{\left(\sum_{l=1}^{M} w_l\right)^2} \\
&= \frac{z_k}{\sum_{l=1}^{M} w_l} - \left[\frac{\sum_{l=1}^{M} z_l w_l}{\sum_{l=1}^{M} w_l}\right]\frac{1}{\sum_{l=1}^{M} w_l} \\
&= \frac{z_k - S(w_1, \ldots, w_M)}{\sum_{l=1}^{M} w_l}
\end{aligned} \tag{D.4}$$

Since $\sum_{l=1}^{M} w_l > 0$, it is easy to see from (D.4) that

$$\frac{\partial}{\partial w_k} S(w_1, \ldots, w_M) \gtreqless 0 \qquad \text{if} \qquad z_k \gtreqless S(w_1, \ldots, w_M) \tag{D.5}$$

Equating $\partial S/\partial w_k$ to zero does not give us any information about the value of w_k when S is maximized or minimized, because

$$\frac{\sum_{l=1}^{M} z_l w_l}{\sum_{l=1}^{M} w_l} = z_k$$

$$\begin{aligned}
\Rightarrow \quad \sum_{l=1}^{M} z_l w_l &= z_k \sum_{l=1}^{M} w_l \\
\Rightarrow \quad z_k w_k + \sum_{\substack{l=1 \\ l\neq k}}^{M} z_l w_l &= z_k w_k + z_k \sum_{\substack{l=1 \\ l\neq k}}^{M} w_l \\
\Rightarrow \quad \frac{\sum_{l\neq k} z_l w_l}{\sum_{l\neq k} w_l} &= z_k \qquad \text{(D.6)}
\end{aligned}$$

where the sum goes from 1 to M. Observe that w_k no longer appears in (D.6). Equation (D.5), however, gives us the direction in which w_k should be changed to increase or decrease S. Observe, from (D.5), that if $z_k > S$, S increases as w_k increases; and, if $z_k < S$, S increases as w_k decreases.

Recall that the maximum value that w_k can attain is $h_k + \Delta_k$ and the minimum value that it can attain is $h_k - \Delta_k$. The discussion in the previous paragraph, therefore, implies that $S(w_1, \ldots, w_M)$ attains its maximum value if : (1) $w_k = h_k + \Delta_k$ for those values of k for which $z_k > S$, and, (2) $w_k = h_k - \Delta_k$ for those values of k for which $z_k < S$. Similarly, $S(w_1, \ldots, w_M)$ attains its minimum value, if : (1) $w_k = h_k - \Delta_k$ for those values of k for which $z_k > S$, and, (2) $w_k = h_k + \Delta_k$ for those values of k for which $z_k < S$.

The maximum of S can be obtained by following the iterative procedure given next. We set $z_l = c_l + s_l$ $(l = 1, \ldots, M)$; and, without loss of generality, assume that the z_l's are arranged in ascending order, i.e., $z_1 \leq z_2 \leq \ldots \leq z_M$.

1. Set $w_l = h_l$ for $l = 1, \ldots, M$, and compute $S' = S(h_1, \ldots, h_M)$ using (D.3).

2. Find k $(1 \leq k \leq M-1)$ such that $z_k \leq S' \leq z_{k+1}$.

3. Set $w_l = h_l - \Delta_l$ for $l \leq k$ and $w_l = h_l + \Delta_l$ for $l \geq k+1$, and compute $S'' = S(h_1 - \Delta_1, \ldots, h_k - \Delta_k, h_{k+1} + \Delta_{k+1}, \ldots, h_M + \Delta_M)$ using (D.3).

   [Since the z_l's are arranged in ascending order, observe, from (D.5) - see also the sentences after Eq. (D.6) - and the fact that $z_k \leq S' \leq z_{k+1}$, that, because we are decreasing the w_l's for $l \leq k$ and increasing the w_l's for $l \geq k+1$, $S'' \geq S'$.]

4. Check if $S'' = S'$. If yes, stop. S'' is the maximum value of $S(w_1, \ldots, w_M)$. If no, go to step 5.

5. Set S' equal to S''. Go to step 2.

It can easily be shown that *this iterative procedure converges in at most M iterations*, where one iteration consists of one pass through steps 1 to 5. At any iteration, let k' be such that $z_{k'} \leq S'' \leq z_{k'+1}$. Since $S'' \geq S'$, $k' \geq k$. If k' is the same as k, the algorithm converges at the end of the next iteration. This can be explained as follow : $k' = k$ implies that both S' and S'' are in $[z_k, z_{k+1}]$. Note that, it is still possible that $S'' \neq S'$. If this happens, however, observe, from step 3, that $S'' = S(h_1 - \Delta_1, \ldots, h_k - \Delta_k, h_{k+1} + \Delta_{k+1}, \ldots, h_M + \Delta_M)$; and, because of step 5, for the next iteration, $S' = S(h_1 - \Delta_1, \ldots, h_k - \Delta_k, h_{k+1} + \Delta_{k+1}, \ldots, h_M + \Delta_M)$. The index k chosen for the next iteration will, therefore, be the same as the index k chosen for the current iteration ($k' = k$); consequently, at the end of the next iteration, we will have $S'' = S(h_1 - \Delta_1, \ldots, h_k - \Delta_k, h_{k+1} + \Delta_{k+1}, \ldots, h_M + \Delta_M) = S'$, and the algorithm will converge. Since k can take at most $M-1$ values, the algorithm converges in at most M iterations.

The minimum of $S(w_1, \ldots, w_M)$ can be obtained by using a procedure similar to the one described above. Only two changes need to be made : (1) we must set $z_l = c_l - s_l$ for $l = 1, \ldots, M$; and, (2) in Step 3, we must set $w_l = h_l + \Delta_l$ for $l \leq k$ and $w_l = h_l - \Delta_l$ for $l \geq k+1$, to compute the weighted average $S'' = S(h_1 + \Delta_1, \ldots, h_k + \Delta_k, h_{k+1} - \Delta_{k+1}, \ldots, h_M - \Delta_M)$.

D.1.1 Centroid of an Interval Type-2 Set

Observe, from (2.69), that the centroid of an interval type-2 set $\tilde{\tilde{A}}$, whose domain is discretized into N points, is given as

$$\tilde{C}_{\tilde{\tilde{A}}} = \int_{\theta_1} \cdots \int_{\theta_N} 1 \Bigg/ \frac{\sum_{l=1}^{N} x_l \theta_l}{\sum_{l=1}^{N} \theta_l} \tag{D.7}$$

where θ_l belongs to some interval in $[0, 1]$. Equation (D.7) has the same form as (D.2), except for the fact that x_l's in (D.7) are crisp numbers unlike Z_l's in (D.2); therefore, the same computational procedure described above can be used to compute $\tilde{C}_{\tilde{\tilde{A}}}$, with the x_l's and θ_l's in (D.7) corresponding to z_l's and w_l's in (D.3), respectively. Note that in this case, $s_l = 0$ for all l, because the x_l's are crisp. If N is very large, in Step (4), we can check if $|S'' - S'| < \epsilon$ instead of $S'' = S'$, for some predecided ϵ.

D.1.2 Type-Reduction for Interval Type-2 FLS's

This computational procedure can be used to compute the type-reduced set for each of the type-reduction methods described in Section 5.3.

1. Centroid type-reduction : The centroid type-reducer [Section 5.3.1 - (5.7)] combines output sets for different rules by finding their union, so that the membership function of the combined output set, $\tilde{\tilde{B}}$, is given by (5.6).

Let the combined output set be discretized into N points, $y_1, \ldots, y_N$, and let $[L_i, R_i]$ be the domain of $\tilde{\mu}_{\tilde{\tilde{B}}}(y_i)$. To use the computational procedure described in this section, note that the sum in (D.3) now goes from 1 to N instead of from 1 to M; y_i's in (5.7) play the role of c_l's; $s_l = 0$ for all l, since the y_i's are crisp; $(L_i + R_i)/2 = h_l$, and $(R_i - L_i)/2 = \Delta_l$.

2. Center-of-sums type-reduction : The center-of-sums type-reducer (Section 5.3.2) combines output sets for different rules by summing them, so that the membership function of the combined output set, $\tilde{\tilde{B}}$, is given by (5.13). The type-reduced set can be computed using the computational procedure described in this section, in exactly the same manner as the centroid type-reducer, where now $\tilde{\tilde{B}}$ is the sum of individual output sets.

3. Height type-reduction : For the height type-reducer [Section 5.3.3 - (5.17)], $\bar{y}^l$'s play the role of c_l's; $s_l = 0$ for all l, since the $\bar{y}^l$'s are crisp; and, the output set memberships of the $\bar{y}^l$'s, $\tilde{\mu}_{\tilde{\tilde{B}}^l}(\bar{y}^l)$'s, play the role of $\tilde{W}^l$'s. If the domain of each $\tilde{\mu}_{\tilde{\tilde{B}}^l}(\bar{y}^l)$ is represented as $[L_l, R_l]$, then $h_l = (L_l + R_l)/2$, and $\Delta_l = (R_l - L_l)/2$.

4. Modified height type-reduction : The computations for a modified height type-reducer [Section 5.3.4 - (5.20)] are very similar to those for a height type-reducer, the only difference being that each output membership grade is now multiplied by a factor of $1/\delta^{l^2}$, so that $h_l = (L_l + R_l)/2\delta^{l^2}$ and $\Delta_l = (R_l - L_l)/2\delta^{l^2}$.

5. Center-of-sets type-reduction : The center-of-sets type-reduced set is given in (5.22), where the t-norm used is product, $\tilde{C}_l = \tilde{C}_{\tilde{\tilde{G}}^l}$ is the centroid of the lth consequent set, and $\tilde{E}_l = \sqcap_{i=1}^{p} \tilde{\mu}_{\tilde{\tilde{F}}_i^l}(x_i)$ is the degree of firing for the lth consequent set, for $l = 1, \ldots, M$. In this case, the computational procedure needs to be applied in two stages. In the first stage, we compute the centroids ($\tilde{C}_l$'s) of the interval type-2 consequent sets (see Section D.1.1); and, in the next stage, we compute the type-reduced set using (D.2). When computing the type-reduced set, $\tilde{C}_l$ plays the role of Z_l in (D.2). If the domain of $\tilde{C}_l$ is the interval $[L_l^c, R_l^c]$, then $c_l = (L_l^c + R_l^c)/2$ and $s_l = (R_l^c - L_l^c)/2$. The degree of firing $\tilde{E}_l$ plays the role of W_l. If the domain of $\tilde{E}_l$ is the interval $[L_l, R_l]$, then $h_l = (L_l + R_l)/2$ and $\Delta_l = (R_l - L_l)/2$.

In each of the above cases, a crisp output for the FLS can be found by computing the centroid of the type-reduced set. Since the type-reduced set is an interval, the centroid is the mid-point of its domain.

Example D.1 In this example, we illustrate the use of the just-described type-reduction methods for Interval type-2 FLS's. We consider a single input - single output type-2 FLS using product t-norm and product inference, which has rules of the form :

$$R^l : \text{IF } x \text{ is } \tilde{\tilde{\mathrm{F}}}^l, \text{ THEN } y \text{ is } \tilde{\tilde{\mathrm{G}}}^l.$$

where $x, y \in [0, 10]$.

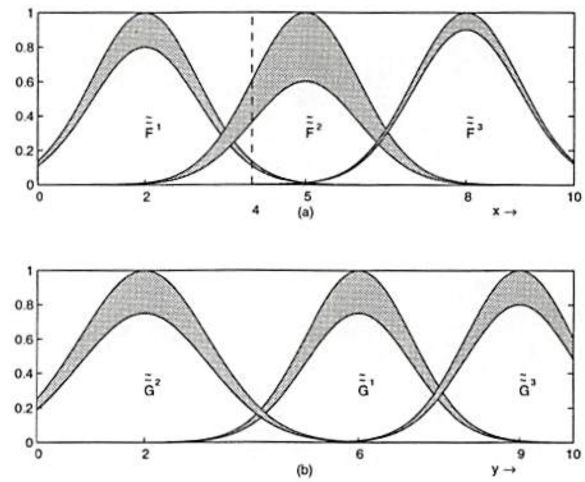


Figure D.1: (a) Antecedent sets (the vertical axis shows the primary memberships of x in the antecedent sets) and (b) consequent sets (the vertical axis shows the primary memberships of y in the consequent sets) for Example D.1. The applied input ($x = 4$) is shown in Fig. (a).

Figures D.1 (a) and (b) show the antecedent and consequent sets. Each of these sets is an interval type-2 set which can be described by two Gaussians which have the same mean and standard deviation. The two Gaussians are scaled to different heights. The maximum height reached by the taller Gaussian is unity, whereas that reached by the shorter Gaussian is s. If the mean and standard deviation of the Gaussians is m and σ, respectively, the membership grade of a domain point x' is an interval $[s\exp\{-0.5(\frac{x'-m}{\sigma})^2\}, \exp\{-0.5(\frac{x'-m}{\sigma})^2\}]$. The m values for each of the antecedent sets, $\tilde{\tilde{\mathrm{F}}}^1$, $\tilde{\tilde{\mathrm{F}}}^2$ and $\tilde{\tilde{\mathrm{F}}}^3$, are 2, 5 and 8, respectively; the σ values are 1, 1 and 1, respectively; and, the s values are 0.8, 0.6, and 0.9, respectively. For the three consequent sets, $\tilde{\tilde{\mathrm{G}}}^1$, $\tilde{\tilde{\mathrm{G}}}^2$ and $\tilde{\tilde{\mathrm{G}}}^3$, the m values are 6, 2, and 9, respectively; the σ values are 1, 1.2, and 1, respectively; and, the s values are 0.75, 0.75, and 0.8, repectively.

The applied input is $x = 4$ [shown in Fig. D.1 (a)]. It has non-zero memberships in two antecedents $\tilde{\tilde{\mathrm{F}}}_1$ and $\tilde{\tilde{\mathrm{F}}}_2$. Figures D.2 and D.3 depict the output sets and type-reduced sets for the centroid and center-of-sums type-reducers, respectively. Figure D.4 depicts the $\bar{y}^l$'s, their output set memberships $[\tilde{\mu}_{\tilde{\tilde{\mathrm{B}}}^l}(\bar{y}^l)$'s - see Eq. (5.16)] and the type-reduced sets for the height and modified height type-reducers. For the modified height type-reducer, the δ^l's were set equal to the σ values of the consequent sets : $\delta^1 = 1$, $\delta^2 = 1.2$ and $\delta^3 = 1$. Figure D.5 depicts the centroids of the consequent sets, their degrees of firing and the center-of-sets type-reduced set. The centroids of

$\tilde{\tilde{G}}^1$, $\tilde{\tilde{G}}^2$ and $\tilde{\tilde{G}}^3$ are $[5.8851, 6.1146]$, $[1.9986, 2.2516]$ and $[8.6415, 8.7857]$, respectively. The type-reduced set in each case, and the centroids of the consequent sets in the center-of-sets type-reducer are computed using the computational procedure described earlier in this section.

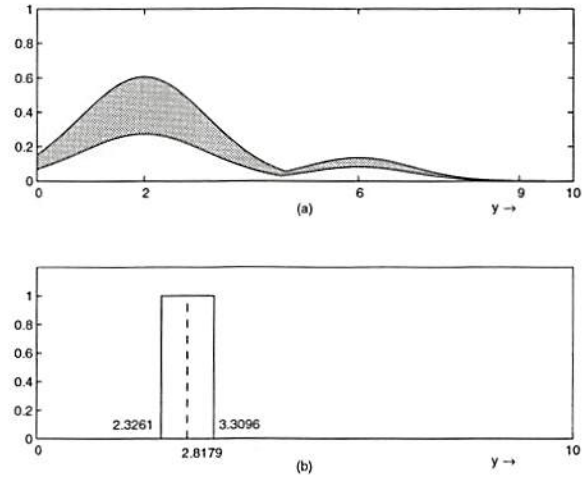


Figure D.2: (a) The combined output set for the centroid type-reducer (the vertical axis shows the primary memberships of y in the combined output set) and (b) the centroid type-reduced set (obtained using the computational procedure in Section D.1) for Example D.1.

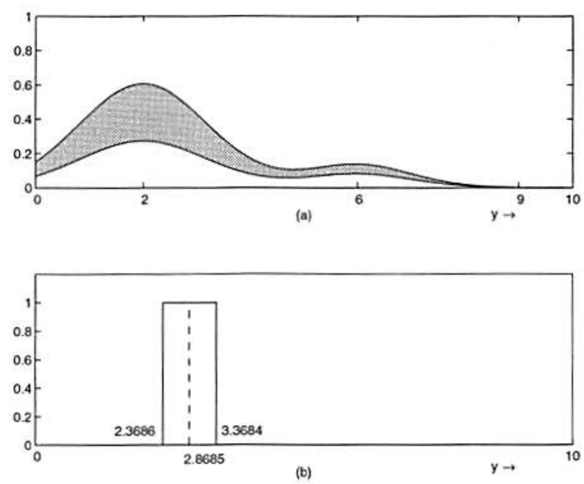


Figure D.3: (a) The combined output set for the center-of-sums type-reducer (the vertical axis shows the primary memberships of y in the combined output set) and (b) the center-of-sums type-reduced set (obtained using the computational procedure in Section D.1) for Example D.1.

The crisp output for the FLS, for each type-reduction method, can be obtained by computing the centroid of the type-reduced set. Since the type-reduced set in each case is crisp, the centroid is equal to the midpoint of its domain interval. The crisp outputs are also indicated on Figs. D.2 - D.5.

The results for this example are collected in Table D.1. In the table, we represent each interval type-1 set in terms of its center and spread. Recall that an interval type-1 set with center c' and spread s' has $[c' - s', c' + s']$ as its domain interval. □

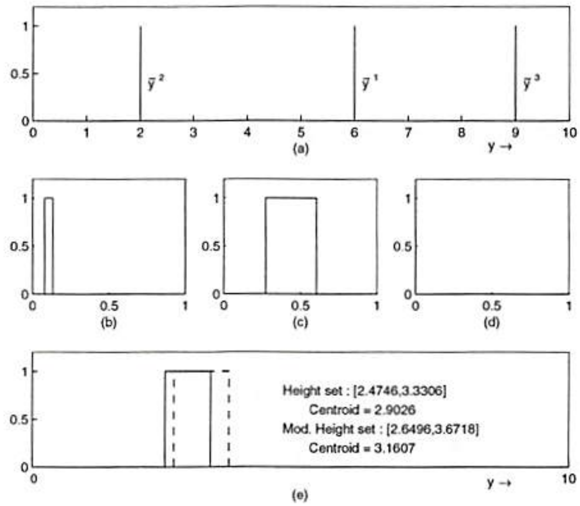


Figure D.4: (a) The $\bar{y}^l$'s, and their output set memberships : (b) $\tilde{\mu}_{\tilde{\bar{B}}^1}(\bar{y}^1)$, (c) $\tilde{\mu}_{\tilde{\bar{B}}^2}(\bar{y}^2)$, and (d) $\tilde{\mu}_{\tilde{\bar{B}}^3}(\bar{y}^3)$; and, (e) the height (solid line) and modified height (dashed line) type-reduced sets (obtained using the computational procedure in Section D.1) for Example D.1. Figures (b), (c) and (d) show plots of primary memberships (horizontal axis) versus secondary memberships (vertical axis). Figure (d) is empty, because $\tilde{\mu}_{\tilde{\bar{B}}^3}(\bar{y}^3)$ is zero (1/0). For the modified height type-reducer, $\delta^1 = 1$, $\delta^2 = 1.2$ and $\delta^3 = 1$.

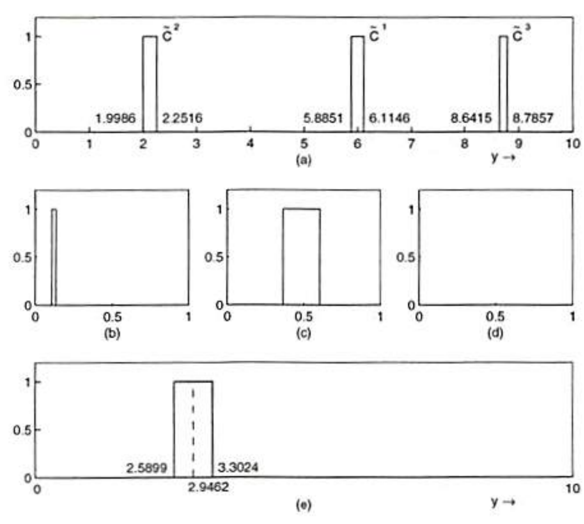


Figure D.5: The centroids of consequent sets, $\tilde{C}^l$'s, are depicted in Fig. (a) and their respective degrees of firing, $\tilde{\mu}_{\tilde{\bar{F}}^l}(x)$'s, for $l = 1$, 2 and 3 are depicted in Figs. (b), (c), and (d), respectively. Figure (e) shows the center-of-sets type-reduced set (obtained using the computational procedure in Section D.1) for Example D.1. Figures (b), (c) and (d) show plots of primary memberships (horizontal axis) versus secondary memberships (vertical axis). Figure (d) is empty, because $\tilde{\mu}_{\tilde{\bar{F}}^3}(x)$ is zero (1/0).

Table D.1: Results of Example D.1.

Type-reduced set	Center	Spread
Centroid	2.8179	0.4918
Center-of-sums	2.8685	0.4999
Height	2.9026	0.4280
Modified height	3.1607	0.5111
Center-of-sets	2.9462	0.3563

D.2 Approximate Result

In this section, we give a result similar to Theorem 2.5 for interval type-2 sets. Before giving the approximate result, we obtain a result similar to Theorem 2.4 for interval type-1 sets.

Theorem D.1 *Given n interval type-1 numbers $F_1, \ldots, F_n$, with means $m_1, m_2, \ldots, m_n$ and spreads $s_1, s_2, \ldots, s_n$, their affine combination $\sum_{i=1}^{n} \alpha_i F_i + \beta$, where α_i $(i = 1, \ldots, n)$ and β are crisp constants, is also an interval type-1 number with mean $\sum_{i=1}^{n} \alpha_i m_i + \beta$, and spread $\sum_{i=1}^{n} |\alpha_i| s_i$.* □

Proof : Consider $F_i = [m_i - s_i, m_i + s_i]$. Multiplying F_i by a crisp constant α_i $(= 1/\alpha_i)$ yields (see 2.63)

$$\alpha_i F_i = \int_v 1/(\alpha_i v) \quad ; \quad v \in [m_i - s_i, m_i + s_i] \tag{D.8}$$

Adding a crisp constant β $(= 1/\beta)$ to $\alpha_i F_i$ yields [see (2.65)]

$$\alpha_i F_i + \beta = \int_v 1/(\alpha_i v + \beta) \quad ; \quad v \in [m_i - s_i, m_i + s_i] \tag{D.9}$$

Substituting $w = \alpha_i v + \beta$, (D.9) gives us

$$\alpha_i F_i + \beta = \int_w 1/w \quad ; \quad w \in [\alpha_i m_i + \beta - |\alpha_i| s_i \,,\, \alpha_i m_i + \beta + |\alpha_i| s_i] \tag{D.10}$$

Recall that F_i can be represented as $[l_i, r_i]$, where $l_i = m_i - s_i$ and $r_i = m_i + s_i$. Observe, therefore, from (2.66) [see, also, the discussion after (2.66)], that

$$\sum_{i=1}^{n} F_i = \left[\sum_{i=1}^{n} m_i - \sum_{i=1}^{n} s_i \,,\, \sum_{i=1}^{n} m_i + \sum_{i=1}^{n} s_i \right] \tag{D.11}$$

Using (D.10) and (D.11), we get the result in Theorem D.1. □

We now give an approximation to the weighted average of interval type-1 sets.

Theorem D.2 *If each Z_l in (D.2) is an interval type-1 set, having center c_l and spread s_l, and if each W_l is also an interval type-1 set with center h_l and spread Δ_l, then Y is approximately an interval type-1 set, with center $\mathcal{C}$ and spread $\mathcal{S}$, where*

$$\mathcal{C} = \frac{\sum_{l=1}^{M} h_l c_l}{\sum_{l=1}^{M} h_l} \tag{D.12}$$

and

$$\mathcal{S} = \frac{\sum_{l=1}^{M}[(h_l s_l) + |c_l - \mathcal{C}|\Delta_l]}{\sum_{l=1}^{M} h_l} \tag{D.13}$$

provided that

$$\frac{\sum_{l=1}^{M} \Delta_l}{\sum_{l=1}^{M} h_l} \ll 1\,, \tag{D.14}$$

The approximation improves as $\left(\sum_{l=1}^{M} \Delta_l \Big/ \sum_{l=1}^{M} h_l\right)$ *grows smaller. The result is exact when* $\sum_{l=1}^{M} \Delta_l = 0$*, i.e., when* $\Delta_l = 0$ *for* $l = 1, \ldots, M$. □

Proof : The proof proceeds exactly like the proof of Theorem 2.5 in Appendix C.11, the only difference being that now the condition required for a good approximation is $\left[\sum_{l=1}^{M} \Delta_l\right] / \left[\sum_{l=1}^{M} h_l\right] \ll 1$ instead of $k\left[\sum_{l=1}^{M} \Delta_l\right] / \left[\sum_{l=1}^{M} h_l\right] \ll 1$ [see (2.78)]. The factor of k appeared in the Gaussian case, because the membership of a point in a Gaussian type-1 set is never exactly equal to 0; we just neglected the memberships outside $\pm k\Delta_l$, since they were too small. In the interval case, however, the memberships of points outside $\pm\Delta_l$ are equal to 0; and therefore, the factor of k disappears.

We get a result similar to (C.229) :

$$Y \approx \sum_{l=1}^{M} \left[\underline{Z}_l\left(\frac{h_l}{\sum_l h_l}\right) + \underline{W}_l\left(\frac{c_l - \mathcal{C}}{\sum_l h_l}\right)\right] + \mathcal{C} \tag{D.15}$$

where $\underline{Z}_l$'s are zero-mean interval type-1 sets with spreads s_l's, $\underline{W}_l$'s are zero-mean interval type-1 sets with spreads Δ_l's, and all the summations and "+" signs denote algebraic sum. The result in Theorem D.2 follows by applying Theorem D.1 to (D.15). When applying Theorem D.1, we set $n = 2M$; $F_i = \underline{Z}_i$, and $\alpha_i = h_i / \left[\sum_{l=1}^{M} h_l\right]$ for $i = 1, \ldots, M$; and, $F_i = \underline{W}_i$, and $\alpha_i = (c_i - \mathcal{C}) / \left[\sum_{l=1}^{M} h_l\right]$ for $i = M+1, \ldots, 2M$. □

Comment 1 : In this case, the true weighted average [i.e., the LHS of (D.2)], Y, will also be an interval type-1 set (since all the sets involved are interval type-1 sets); however, the approximation is useful because the actual end-points of the domain of Y can only be obtained computationally.

Comment 2 : Comments 2 and 3 at the end of Appendix C.11 apply in this case as well.

Comment 3 : Though Theorem D.2 is very much similar to Theorems 2.5, there is one difference. In case of Gaussians sets, the secondary membership functions may be clipped because they have to be contained in $[0, 1]$; and therefore, may not remain true Gaussians. We ignored these clipping effects for simplicity; therefore, the results in Theorems 2.5 contained a "clipping effect" approximation, in addition to the approximations introduced subject to conditions (2.78). In the case of interval sets, however, no clipping effects need to be considered, because any clipped version of an interval is again an interval; so, the only approximation that is introduced in the result in Theorem D.2 is the one subject to condition (D.14).

Example D.2 Consider the interval type-2 FLS in Example D.1. In this example, we obtain approximate type-reduced sets for each of the type-reducers considered in Example D.1, by using Theorem D.2. The approximate results are collected in Table D.2. The value of $\sum_l \Delta_l / \sum_l h_l$, in each case, is also shown. For the center-of-sets type-reducer, the three consequent centroids, computed approximately using Theorem D.2, are $[5.8859, 6.1138]$, $[1.9987, 2.2501]$ and $[8.6424, 8.7861]$, respectively; and, the values of $\sum_l \Delta_l / \sum_l h_l$ for the consequent sets $\tilde{\tilde{G}}^1$, $\tilde{\tilde{G}}^2$, and $\tilde{\tilde{G}}^3$, are 0.1429, 0.1429, and 0.1111, respectively.

Observe that the ratios $\sum_l \Delta_l / \sum_l h_l$ for the consequent centroid computation are much smaller than those for the type-reduction computations; consequently, the approximation is much closer to the true value (see Example D.1) for the consequent centroids than for the type-reduced sets. Observe also, from Table D.2, that, for this example, the center-of-sets type-reducer has the lowest

Table D.2: Results of Example D.2.

Type-reduced set	Approximate Center	Approximate Spread	$\sum_l \Delta_l / \sum_l h_l$
Centroid	2.7622	0.4642	0.3576
Center-of-sums	2.8135	0.4730	0.3563
Height	2.7935	0.4002	0.3537
Modified height	3.0511	0.4876	0.3453
Center-of-sets	2.9051	0.3484	0.2220

value of $\sum_l \Delta_l / \sum_l h_l$; and correspondingly, the difference between the true and approximate type-reduced set centers and spreads is the lowest for the center-of-sets type-reducer. □

Reference List

[1] J. L. Chameau, M. Gunaratne and A. G. Altschaeffl, "An Application of Type-2 Sets to Decision Making in Engineering", *Analysis of Fuzzy Information - vol. II : Artificial Intelligence and Decision Systems* - James C. Bezdek (editor), CRC Press, Inc., Boca Raton, FL. 1987.

[2] D. Driankov, H. Hellendoorn and M. Reinfrank, *An Introduction to Fuzzy Control* (2nd ed.), Springer-Verlag, 1996.

[3] D. Dubois and H. Prade, "Operations on Fuzzy Numbers", *Int. J. Systems Sci.*, vol. 9, no. 6, pp. 613-626, 1978.

[4] D. Dubois and H. Prade, "Operations in a Fuzzy-Valued Logic", *Information and Control*, vol. 43, pp. 224-240, 1979.

[5] D. Dubois and H. Prade, *Fuzzy Sets and Systems : Theory and Applications*, Academic Press, Inc., NY. 1980.

[6] W. Gellert, S. Gottwald, M. Hellwich, H. Kästner, H. Küstner (eds.), *The VNR Concise Encyclopedia of Mathematics (2nd ed.)*, Van Nostrand Reinhold, NY. 1989.

[7] M. B. Gorzalczany, "A Method of Inference in Approximate Reasoning Based on Interval-Valued Fuzzy Sets", *Fuzzy Sets and Systems*, vol. 21, pp. 1-17, 1987.

[8] N. N. Karnik, *Type-2 Fuzzy Logic Systems*, PhD Dissertation, University of Southern California, Los Angeles, CA. 1998.

[9] N. N. Karnik and J. M. Mendel, "Introduction to Type-2 Fuzzy Logic Systems", presented at the 1998 IEEE FUZZ Conference, Anchorage, AK, May.

[10] A. Kaufmann and M. M. Gupta, *Introduction to Fuzzy Arithmetic : Theory and Applications*, Van Nostrand Reinhold, NY. 1991.

[11] D. Kim and C. Kim, "Forecasing Time Series with Genetic Fuzzy Predictor Ensemble", *IEEE Transactions on Fuzzy Systems*, vol. 5, no. 4, pp. 523-535, November 1997.

[12] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic : Theory and Applications*, Prentice Hall, Upper Saddle River, NJ. 1995.

[13] B. Kosko, *Fuzzy Engineering*, Prentice Hall, Upper Saddle River, NJ. 1997.

[14] M. C. Mackey and L. Glass, "Oscillation and Chaos in Physiological Control Systems", *Science*, vol. 197, pp. 287-289, 1977.

[15] J. M. Mendel, "Fuzzy Logic Systems for Engineering : A Tutorial", *Proceedings of the IEEE*, vol. 83, no. 3, March 1995.

[16] J. M. Mendel, S. Murphy, L. C. Miller, M. Martin and N. Karnik, "The Fuzzy Logic Advisor for Social Judgements", presented at MENDEL'97, Brno, Czech Republic, June 25-27, 1997.

[17] M. Mizumoto and K. Tanaka, "Some Properties of Fuzzy Sets of Type-2", *Information and Control*, 31, 1976, pp. 312-340.

[18] M. Mizumoto and K. Tanaka, "Fuzzy Sets of Type-2 Under Algebraic Product and Algebraic Sum", *Fuzzy Sets and Systems*, vol. 5, 1981, pp. 277-290.

[19] G. Mouzouris, *Non-Parametric Modeling of Nonlinear Processes Using a Fuzzy Set Theoretic Approach*, USC-SIPI Report #300, Signal and Image Processing Institute, University of Southern California, May 1996.

[20] J. Nieminen, "On The Algebraic Structure of Fuzzy Sets of Type-2", *Kybernetica*, vol. 13, no. 4, 1977.

[21] M. Sugeno and G. T. Kang, "Structure Identification of Fuzzy Model", *Fuzzy Sets and Systems*, vol. 28, pp. 15-33, 1988.

[22] T. Takagi and M. Sugeno, "Fuzzy Identification of Systems and its Application to Modeling and Control", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 15, no. 1, 1985, pp. 116-132.

[23] I. Turksen, "Interval Valued Fuzzy Sets Based on Normal Forms", *Fuzzy Sets and Systems*, vol. 20, pp. 191-210, 1986.

[24] L.-X. Wang, *Analysis and Design of Fuzzy Systems*, USC-SIPI Report #206, Signal and Image Processing Institute, University of Southern California, June 1992.

[25] L.-X. Wang, *A Course in Fuzzy Systems and Control*, Prentice Hall, Upper Saddle River, NJ. 1997.

[26] L.-X. Wang and J. M. Mendel, "Generating Fuzzy Rules by Learning from Examples", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22, no. 6, pp. 1414-1427, November/December 1992.

[27] M. Wagenknecht and K. Hartmann, "Application of Fuzzy Sets of Type 2 to the Solution of Fuzzy Equation Systems", *Fuzzy Sets and Systems*, vol. 25, pp. 183-190, 1988.

[28] W. Wei,*Classification of Digital Modulations Using Constellation Analyses*, USC-SIPI Report #317, Signal and Image Processing Institute, University of Southern California, May 1998.

[29] R. R. Yager, "Fuzzy Subsets of Type II in Decisions", *Journal of Cybernetics*, vol. 10, pp. 137-159, 1980.

[30] L. A. Zadeh, "The Concept of A Linguistic Variable and Its Application to Approximate Reasoning - 1", *Information Sciences* 8, pp. 199 - 249. 1975.

Index

E

F

G

H

I

J

L

M

U

W