

Capítulo 7

Inferencia Y Probabilidad. Lógica Difusa.

A menudo se aplica la denominación de Sistemas Expertos a sistemas de inferencia lógica, es decir a estructuras basadas en reglas que permiten evaluar la verdad o el grado de certeza de una afirmación en función de la evaluación de los antecedentes.

La teoría básica de los sistemas expertos esta referida a la evaluación de sistemas binarios, es decir que solo pueden ser evaluados como verdaderos o falsos.

En un libro de sugestivo nombre “Teoría de Probabilidades: La Lógica de la Ciencia” el autor (Edwin Thompson Jaynes, 1994) da un ejemplo interesante para mostrar la esencia probabilística del proceso de inferencia cuando esta es realizada por una persona. El ejemplo nos remite a plantearnos un escenario en el cual un vigilante realiza una caminata nocturna y repentinamente escucha el sonido de una alarma, seguido del estrépito de vidrios rotos, al doblar una esquina observa a un sujeto enmascarado salir a la carrera de un negocio pasando a través de los restos de vidrio de una vidriera, este hombre es a su vez es seguido enérgicamente por otra persona. Sin dudar el vigilante se lanza tras el supuesto delincuente que acaba de asaltar el negocio y que es perseguido además por el dueño del mismo

Debemos en este punto observar que todo el análisis que el vigilante hace, es un supuesto que se basa en probabilidades. La probabilidad de que el enmascarado que huye a la carrera sea realmente un ladrón es alta, pero podría quizás no haber sido un ladrón... quizás una banda de muchachos pasaron momentos antes por allí y arrojaron una piedra que rompió el vidrio, y quizás en ese momento se estaba llevando a cabo una fiesta de disfraces en el interior del edificio, y algunos de los participantes salieron tras los vándalos... y uno de ellos traía aun puesta su máscara...

Vemos así que los razonamientos concernientes a problemas reales muestran que es posible considerar cierto grado de certeza o falsedad para una afirmación y no valores absolutos de verdad o falsedad.

Para incorporar este concepto a los sistemas de inferencia y a los sistemas expertos se recurre a modelos probabilísticos, es decir se asignan valores de probabilidad a las sentencias y se propagan a través de las reglas. Que un coeficiente de probabilidad se propague a través de una regla significa calcular que probabilidad se atribuye al consecuente en función de las probabilidades de los antecedentes.

Existen distintas formas de calcular como influyen las probabilidades de los antecedentes sobre los consecuentes. Esto da lugar a distintos esquemas de inferencia en IA, tales como Inferencia Probabilística, Inferencia Difusa, Inferencia Bayesiana y Propagación de Restricciones.

Campos de Aplicación y denominaciones en Lógica Difusa

Los sistemas basados en lógica Difusa se han empleado con éxito en campos tales como control automático, clasificación, soporte de decisión, sistemas expertos, y visión computarizada. A causa de su naturaleza multi-disciplinaria, el sistema de inferencia difusa es conocido bajo distintas denominaciones, tales como sistema difuso basado en reglas (Rule based fuzzy system), sistema

experto difuso (fuzzy expert system), modelo difuso (fuzzy model), memoria asociativa difusa (fuzzy associative memory), controlador lógico difuso (fuzzy logic controller), y también sencillamente (y ambigüamente) sistema difuso. Como se ve en materia de terminología se está aún lejos de un estándar.

En lo que sigue usaremos los términos fuzzy y difusa como sinónimos de manera indistinta de modo de estar familiarizados con la denominación en inglés por ser de uso frecuente. Al proceso inverso a la denominada Difusión o Fuzificación lo llamaremos con la palabra defuzzificación.

7.1. Un ejemplo comparativo

Para ilustrar como es un sistema de inferencia basado en lógica difusa, se discuten continuación dos modos diferentes de resolver mismo problema. Una es una solución convencional, dada por un sistema lineal y el otro por un sistema de lógica difusa.

El problema consiste en determinar el valor “correcto” para una propina que se debe abonar en función de algunas características del servicio recibido. Básicamente, se trata de evaluar cuanto debería ser la propina, dado un número entre cero y diez que representa la calidad del servicio recibido. Un valor promedio de propina por comida se estima en 15 % del total de la cuenta. La propina a entregar variará en función de la calidad del servicio recibido.

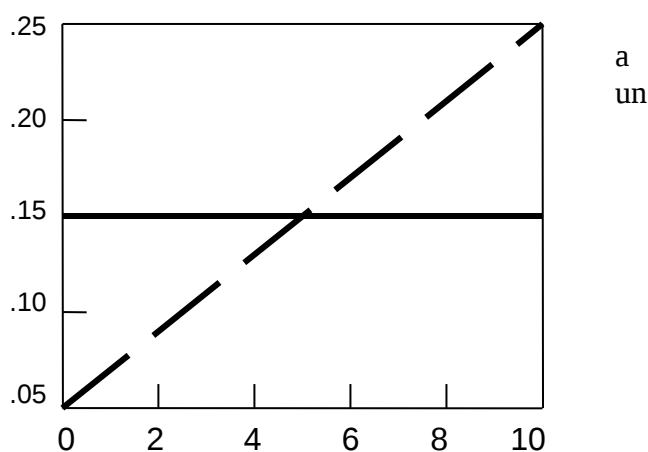


Fig. 6.1: Propina independiente de la calidad del servicio y con dependencia lineal

7.2. Una solución

Comenzamos con la sencilla posible. Podemos decidir entregaremos una propina del 15 cuenta, con lo que tendríamos un de la Fig. 6.1.

Para tener en cuenta la servicio introducimos un nuevo que el servicio varía entre cero y tendremos que la propina varía 5% si el servicio es pésimo a 25% Con esto tenemos (Fig. 6.2):

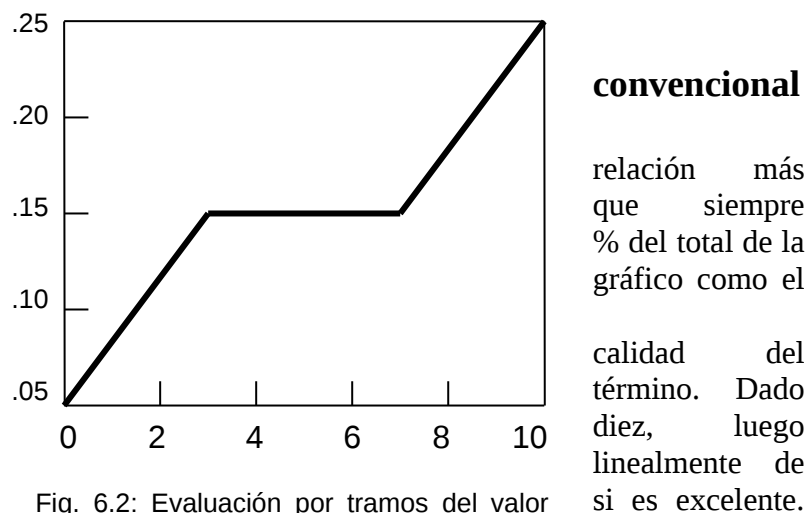


Fig. 6.2: Evaluación por tramos del valor de la propina

$$\text{Propina} = 0.20/10 * \text{servicio} + 0.05$$

Además podríamos querer una respuesta plana en el medio, es decir decidir una propina del 15% en general, y nos apartaríamos de este valor solo en casos en que el servicio o la comida son especialmente buenos o malos.

Podemos modelar este criterio empleando una función por trozos para lo cual planteamos una evaluación condicional:

si servicio < 3,

```
propina=(0.10/3)*servicio+0.05;
else si servicio<7,
    propina=0.15;
else si servicio<=10,
    propina=(0.10/3)*(servicio-7)+0.15;
```

Supongamos ahora que queremos también considerar la calidad de la comida para fijar el valor de la propina. Podríamos proponer la siguiente fórmula:

$$\text{Propina} = 0.20 / 20 * (\text{servicio} + \text{comida}) + 0.05 ;$$

Podríamos pretender que el servicio es un factor mas importante que la calidad e la comida. Esto lo consideraríamos asignado una ponderación del 80% para el servicio y el restante 20% para la comida. La representación de esta función de dos variables sería un plano inclinado.

La función es sorprendentemente complicada considerando el punto de partida y no parece sencilla de modificar. Esta llena de números mágicos y no hace aparente que es lo que esta representando.

7.3. La Solución Fuzzy

Sería deseable plantear la solución del problema en términos que permitan capturar la esencia del mismo. Particularmente sería deseable eliminar factores arbitrarios. La esencia del problema es que:

1. *si el servicio es pobre luego la propina es poca*
2. *si el servicio es bueno luego la propina es media*
3. *si el servicio es excelente luego la propina es generosa*

Si quisiéramos incluir la incidencia de la calidad de la comida solo agregamos algunas reglas:

4. *si la comida es rancia luego la propina es poca*
5. *si la comida es deliciosa luego la propina es generosa*

Podemos combinar las dos listas de reglas en una sola:

1. *si el servicio es pobre o la comida es rancia luego la propina es poca*
2. *si el servicio es bueno luego la propina es media*
3. *si el servicio es excelente o la comida es deliciosa luego la propina es generosa*

Si damos ahora significado matemático a las variables lingüísticas (que significa una propina "media" ?) obtendremos un sistema de inferencias basado en lógica difusa.

Algunas cosas quedan pendientes por el momento tales como el modo en que se combinan las reglas o como se obtiene el significado matemático de la variable lingüística.

Y se verán en los párrafos siguientes.

Con este método obtendríamos una función como la graficada en la Fig. 6.3.

Observaciones

Se ha mostrado una solución relación lineal por trozos para problema planteado. Si bien funciona, escrito su código no resulta entendible. Si la lógica difusa nos escribir un código para la solución mas fácil e interpretar para las entonces estamos diciendo que su radica en la posibilidad de escribir un ventajoso en algún aspecto.

Como el sistema difuso esta modo basado en sentencias de común”, podremos también agregar sentencias al final de la lista que mas apropiadamente la forma de la decir que entre las ventajas de este código estará la de una modificación mas sencilla.

También el mantenimiento del código será mas simple debido a que claramente se puede desacoplar en diferentes reglas sencillas. La noción de una propina alta, media o baja podrá cambiar día a día, ciudad por ciudad o país por país pero la lógica subyacente será siempre la misma. No importa en que parte del mundo apliquemos el algoritmo, será posible recalibrar rápidamente el sistema sin necesidad de reescribir reglas y código.

Pensemos como sería la solución con la función lineal por trozos, deberíamos comentar profusamente el código. Estudiarlo muy cuidadosamente en caso de modificaciones y comentarlo nuevamente. Ciertamente podría ser mucho mas engorroso.

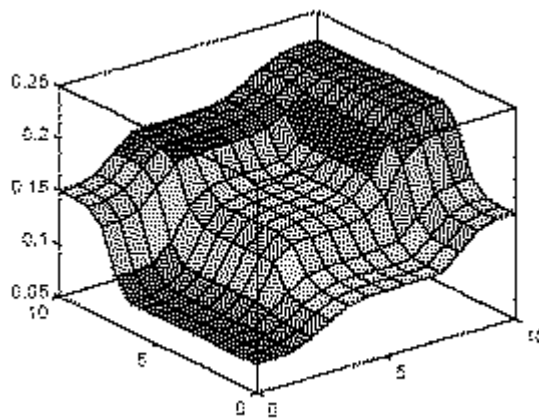


Figura 6.3: Solución Fuzzy para el problema e la propina con dos variables

mediante una resolver el una vez directamente permite que resulte personas, contribución código

de algún “sentido mas moldearán función. Es

7.4. Una visión general de un sistema de inferencia difusa

Comenzaremos con un general de un sistema basado en lógica para luego discutir sus componentes. La un sistema fuzzy es mapear un espacio en uno de salida. El mecanismo para lograrlo es la inferencia basada en

Todas las reglas se evalúan en su orden no tiene importancia. Las lógica difusa se refieren a variables y que califican a esas variables. Antes de construcción del sistema que interpreta emplean adjetivos, necesitamos definir términos que planeamos usar y los que los describen.

Esto quiere decir que si quisiéramos discutir acerca de “cuan agua” debemos considerar el rango de temperatura posible para el agua así significado de la palabra fría, lo que se los párrafos siguientes.

El diagrama de la Fig. 6.4 un proceso de inferencia difusa. A la se muestra un esquema general y a la caso específico del problema de la

La idea tras un sistema de difusa es interpretar los valores de un entrada y, basado en un conjunto de asignar valores al vector de salida.

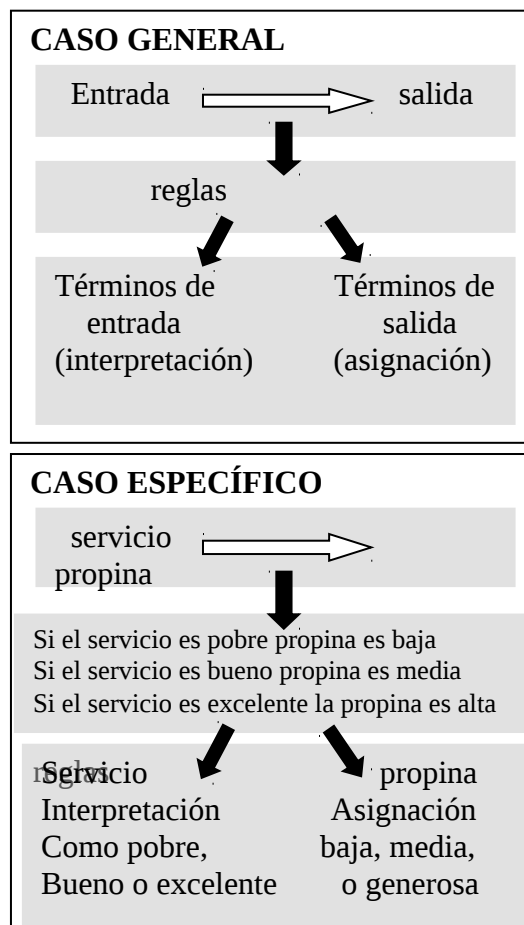


Figura 6.4: Reglas asociadas al proceso de difusión.

panorama difusa esencia de de entrada primario reglas. paralelo, reglas en adjetivos discutir la reglas que los adjetivos

fría esta el como el discute en

muestra izquierda derecha el propina. inferencia vector de reglas,

7.5. Reglas Si-entonces

Los conjuntos difusos y los operadores difusos son los verbos y sujetos de la lógica fuzzy. Como para lograr algo útil debemos construir sentencias completas, las sentencias condicionales, reglas si-luego, son los primeros elementos que permiten sacar provecho de la lógica difusa. Una regla difusa toma la forma:

si x es A luego y es B

donde A y B son valores lingüísticos definidos por conjuntos difusos sobre los rangos (Universos de discurso) X e Y , respectivamente.

La parte primera de la regla es el antecedente o premisa, mientras que la parte final " y es B " es el consecuente o conclusión. Un ejemplo sería:

Si el servicio es bueno luego la propina es media

Notar que el antecedente (servicio) es una interpretación que retorna un número entre 0 y 1, mientras que el consecuente es una sentencia que asigna el conjunto fuzzy completo B a la variable de salida y (media). De modo que la palabra "es" es empleada de dos modos diferentes dependiendo si aparece en el antecedente o en el consecuente.

La interpretación de estas reglas comprende distintas partes: 1) evaluación del antecedente (Lo que implica difusión de la entrada y aplicación de operadores difusos) y 2) aplicación del resultado al consecuente (implicación).

En el caso de lógica binaria, las reglas si-luego son sencillas. Si la premisa es verdadera, luego la conclusión es verdadera. Pero si relajamos las restricciones de la lógica binaria y permitimos que el antecedente sea una sentencia fuzzy, tendremos que considerar como influye esto en la conclusión. En general podríamos expresar que si el antecedente es verdadero en alguna medida de pertenencia, luego el consecuente es también verdadero en el mismo grado. Es decir,

en lógica binaria: $p \rightarrow q$ (p y q son o verdaderos o falsos)

en lógica difusa: $0.5 p \rightarrow 0.5 q$ (antecedentes implican parcialidad)

El antecedente de una regla puede estar compuesto de múltiples partes:

si el cielo es gris y el viento es fuerte y la presión esta descendiendo, luego ...

En cuyo caso todas las partes del antecedente son calculadas simultáneamente y resueltas como un número simple empleado operadores lógicos, El consecuente también puede tener múltiples partes.

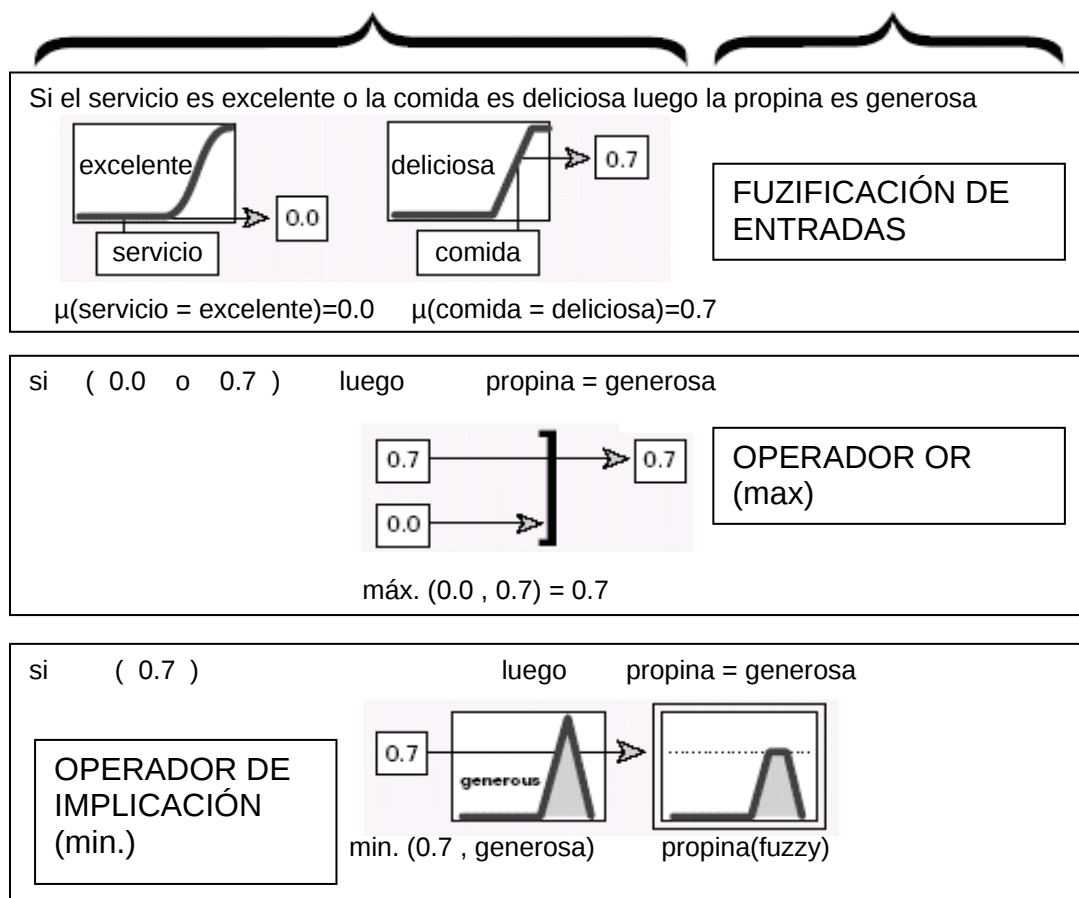
si la temperatura es baja luego la válvula de agua caliente esta abierta y la válvula de agua fría esta cerrada.

En cuyo caso todos los consecuentes son afectados igualmente por el resultado del antecedente.

Pero dijimos que debíamos considerar ahora como afecta el antecedente al consecuente. En Lógica Difusa el consecuente especifica un conjunto difuso para la salida.

La función de implicación modifica luego al conjunto difuso en el grado especificado por el antecedente.

Los medios mas usuales para modificar le conjunto de salida difuso son truncamiento (función mín.) (donde el conjunto difuso es recortado, Fig. 6.5) o escalación usando la función producto.



6.6. Sumario de reglas

La interpretación de las reglas aplicando lógica difusa es un proceso de varias etapas.

- I Difusión de entradas: Se deben resolver todas las sentencias en el antecedente en función de su grado de membresía entre 0 y 1. Si

el antecedente tiene solo un componente este es el grado de soporte del consecuente.

- II. Aplicación de operadores difusos: Si existen múltiples partes en el antecedente se aplican los operadores fuzzy y se resuelve el antecedente como un número entre 0 y 1. Esto es el grado de soporte de la regla
- III. Aplicación de la implicación: Se emplea el grado de soporte de la regla para conformar el conjunto fuzzy de salida. El consecuente de una regla fuzzy asigna un conjunto fuzzy completo a la salida. Si el antecedente es solo parcialmente cierto, luego el conjunto fuzzy de salida es truncado según el método de implicación.
- IV En general, una sola regla no es muy útil. Y por lo tanto se requieren dos o mas reglas que interactúen entre ellas. La salida de cada regla es un conjunto fuzzy, pero en general se necesita que la salida sea en definitiva un número. Para obtener este número, primero los conjuntos fuzzy de salida de cada regla se amalgaman en un único conjunto. Luego este conjunto resultante es defuzificado, o resuelto de modo de obtener un número como valor de salida.

La naturaleza paralela de las reglas es un importante aspecto de los sistemas de lógica difusa. Estos sistemas permiten variar suavemente desde regiones donde el comportamiento del

Figura 6.5: Operadores difusos

sistema es dominado ya por esta regla ya por aquella.

7.7. Paso 1. Fuzificación

El primer paso es tomar las determinar el grado en el cual ellos uno de los conjuntos difusos a través membresía o pertenencia. Esta muestra en la Fig. 6.7. La entrada es numérico limitado al universo del variable de entrada (en este caso el 10) y la salida es un grado difuso de (siempre intervalo entre 0 y 1). De fuzificación no difiere sustancialmente evaluación de una función. Estas pertenencia pueden ser vistas como la distribución de probabilidad del cual han sido definidas.

El ejemplo visto al comienzo consta para cada una de las cuales debe difusión inicial para cada adjetivo o conjunto difuso. Esta fuzificación inicial debe ser realizada antes de aplicar las reglas.

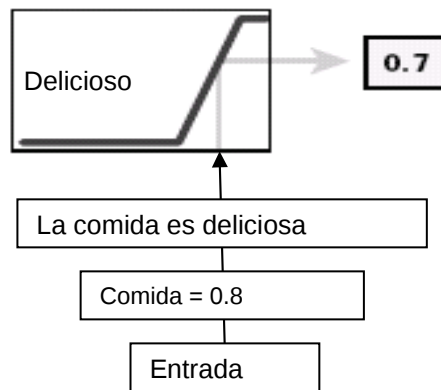


Figura 6.7: Empleo de la función de membresía o pertenencia para asignar valores difusos al adjetivo “delicioso”

entradas y pertenecen a cada de funciones de situación se siempre un valor discurso de la intervalo entre 0 y pertenencia modo que la de una LUT o funciones de curva de adjetivo para el

de tres reglas, realizarse la debe ser realizada

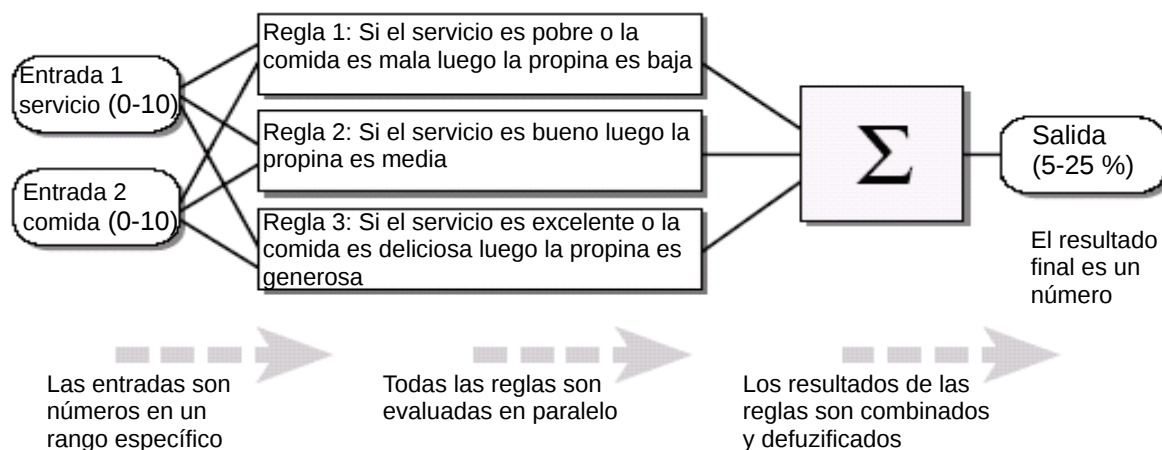


Figura 6.8: La red de inferencia difusa requiere de una fuzificación de las variables de entrada (servicio y comida) para que los números obtenidos a partir de ésta sean ingresados a la red.

7.8. Paso 2. Aplicación del operador difuso

Tabla 6.1 – Cálculo de valores de pertenencia

Oper	Método	Descripción
P y Q	Mínimo	$\min(XP, XQ)$
"	Producto	$XP * XQ$
"	truncamiento	$\max((XP + XQ - 1), 0)$
P o Q	Máximo	$\max(XP, XQ)$
"	Amplificación	$XP + XQ * (1 - XP)$
"	Adición	$\min(XP + XQ, 1)$
no P	complemento	$1 - XP$

Una vez que las entradas han sido fuzzificadas, conocemos el grado en el cual cada parte del antecedente ha sido satisfecho para cada regla. Si el antecedente de una regla dada ha sido dividido en mas de una parte, se debe aplicar un operador difuso para obtener un número que representa el resultado del antecedente para esa regla.

La entrada del operador difuso son dos o mas valores de pertenencia desde las variables de entrada fuzzificadas. La salida es un único valor de verdad.

Existen métodos que en conjunciones producto y disyunciones amplificación y negación (Tabla continuación se procesa para la (max) .

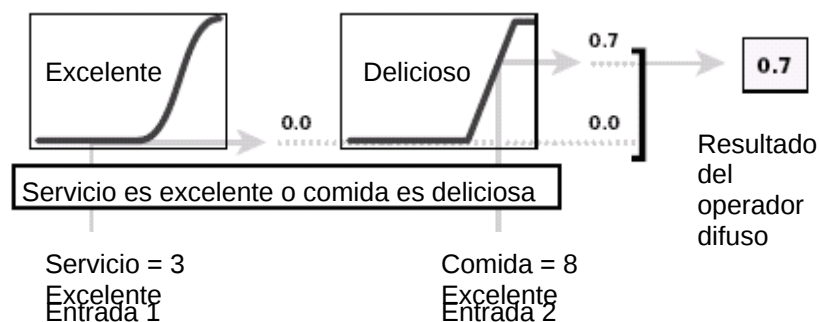


Figura 6.9: Fuzificación y operador or.

distintos general son (mínimo, truncamiento), (máximo, adición) y 6.1). A grafica el función OR Evaluamos el la regla 3 para propina. Los

antecedentes arrojan los valores de pertenencia 0.0 y 0.7 respectivamente. El operador Or selecciona el máximo de ambos valores, 0.7, y la operación para la regla esta completada (Fig. 6.9).

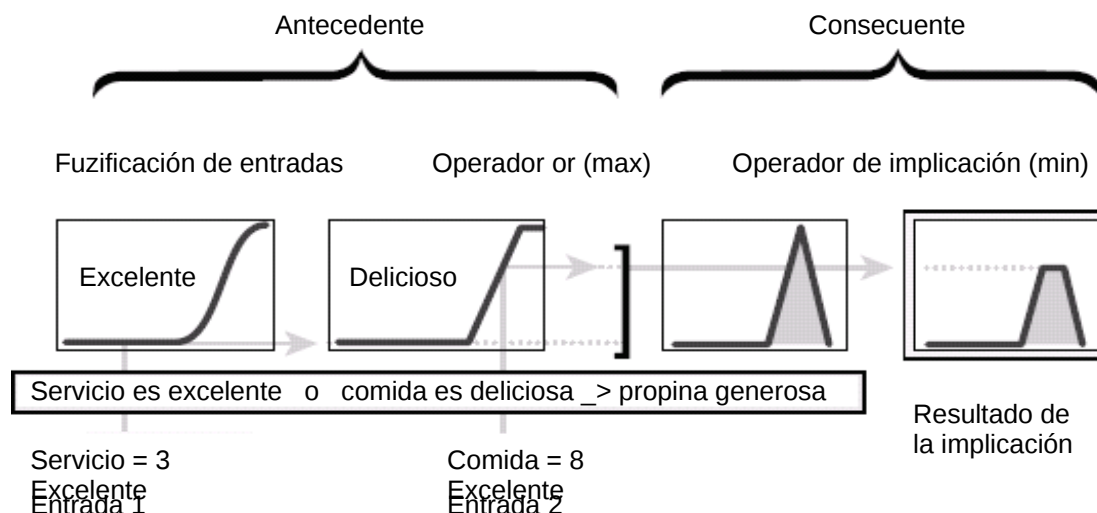


Figura 6.10: Fuzificación, disyunción e implicación

7.9. Paso 3. Aplicación del método de Implicación

El método de implicación se define como la conformación del consecuente (un conjunto fuzzy) basado en el antecedente (un número). La entrada para la implicación es un numero dado por el antecedente, y la salida es un conjunto fuzzy. La se obtiene para cada regla por separado. Algunos métodos podrían ser mín. (mínimo) que trunca el conjunto fuzzy, y prod. (producto) el que amplifica el conjunto de salida fuzzy.

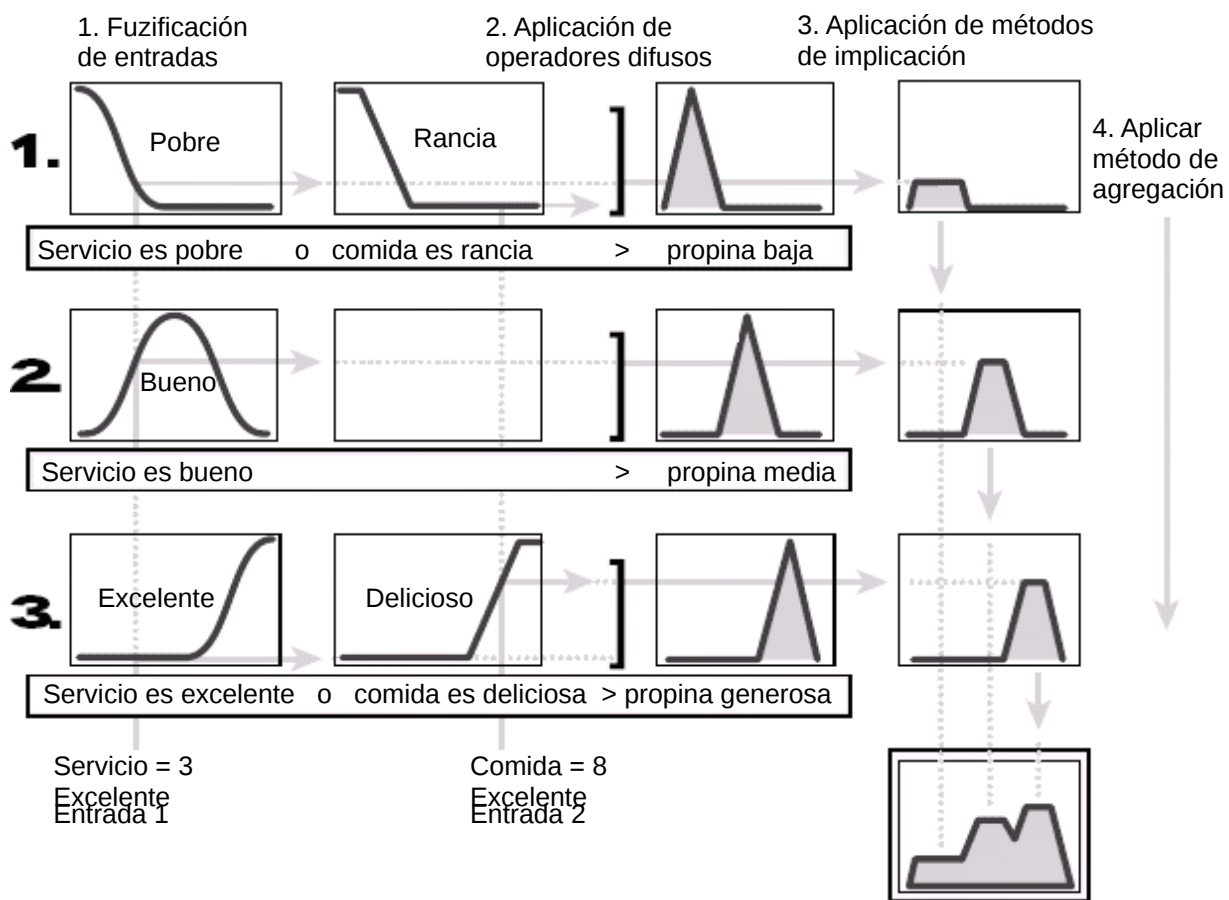


Figura 6.11: Las reglas se evalúan individualmente para obtener el valor de la implicación. Luego deben asociarse los conjuntos de salida obtenidos para cada regla.

7.10. Paso 4. Agregación

Agregación es el proceso de unificar las salidas para cada regla uniando los procesos paralelos. Es decir se combinan las salidas en un único conjunto difuso preparándolas para el paso final que es la defuzzificación.

La entrada para la agregación es la lista de salidas truncadas (o modificadas por la implicación) retornadas por el proceso de implicación para cada regla. La salida de este proceso era un conjunto difuso para cada variable de salida.

La agregación es conmutativa, luego el orden en que se ejecutan las reglas no es relevante.

Algunos métodos son: max (máximo), prob-or (or probabilística), y sum (la suma de cada conjunto de salida para cada regla).

En la Fig. 6.12, se observa como la salida de cada regla se combina, o agrega, en un único conjunto difuso para toda la salida.

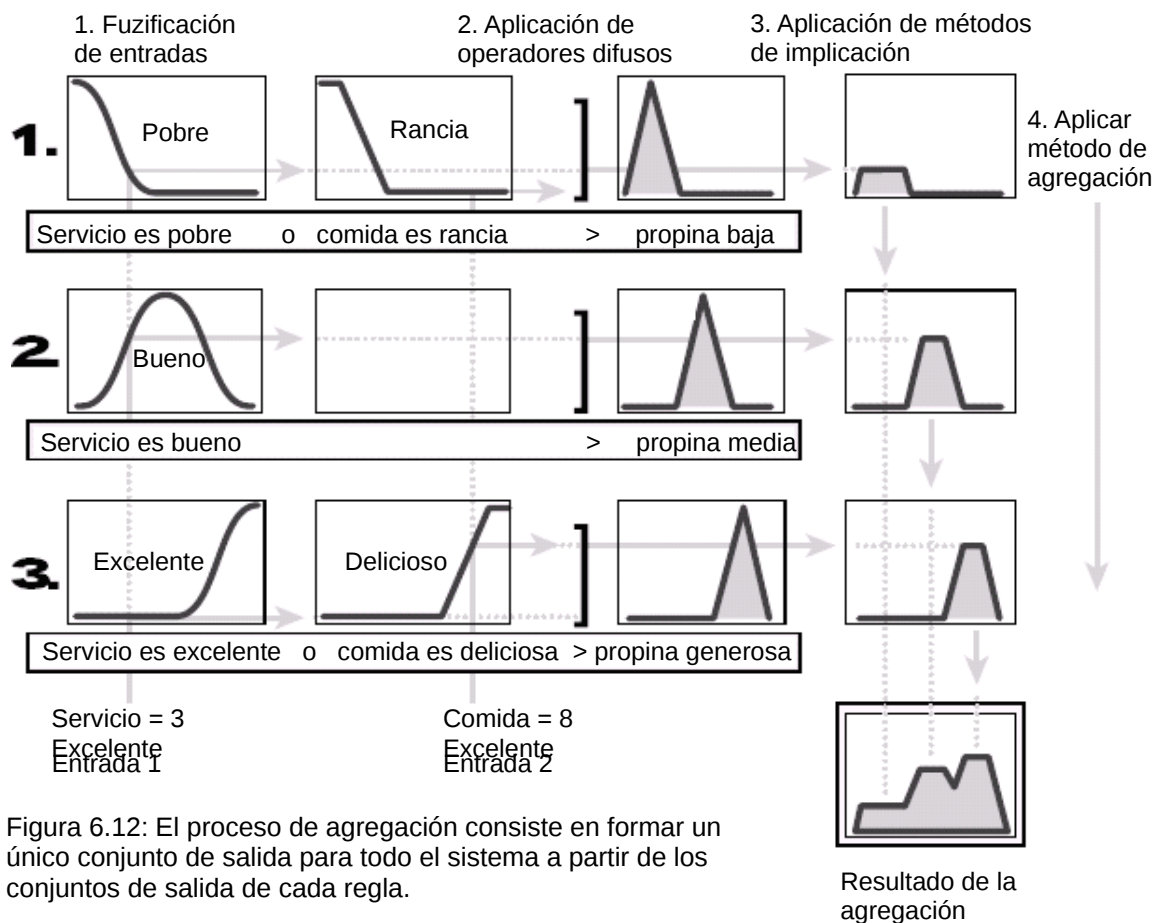


Figura 6.12: El proceso de agregación consiste en formar un único conjunto de salida para todo el sistema a partir de los conjuntos de salida de cada regla.

7.11. Paso 5. Defuzzificación

La entrada para el proceso de defuzzificación es el conjunto difuso agregado (el conjunto difuso unificado de salida) y la salida es un número.

Entonces dado un conjunto difuso que engloba un rango de valores de salida se requiere obtener un único número.

Un popular método de defuzzificación es el cálculo centroide, el cual retorna el centro de un área bajo una curva.

La defuzzificación tiene lugar en dos distintos pasos. Primero las funciones de pertenencia son escaladas de acuerdo a sus posibles valores, luego estas son usadas para calcular el centroide de los conjuntos difusos asociados.

Un diagrama de inferencia difusa completo quedaría formado por el ensamble de los diagramas vistos. En la Fig se muestra nuevamente un modelo simplificado del diagrama completo

Notar que las flechas muestran la evolución del flujo de información desde la fuzificación inicial hasta el resultado final.

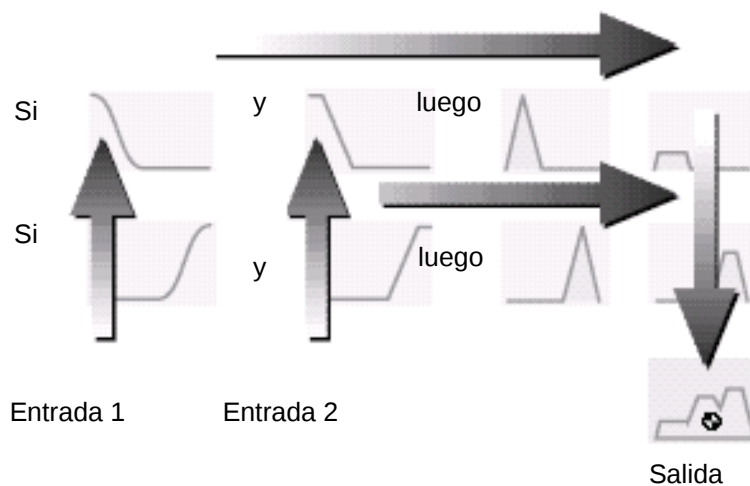
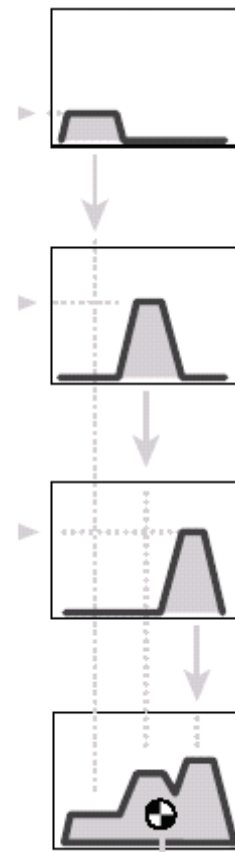


Figura 6.14: Diagrama simplificado del sistema de inferencia difusa. Las flechas muestran la evolución del flujo de información desde la fuzificación inicial hasta el resultado final.



Salida
Propina = 16.7

Figura 6.13: El método mas sencillo de defuzzificación es hallar el centroide.

7.12. Heurística y Lógica Difusa

En IA, se observa a veces cierta tendencia a asociar el concepto de heurística a los procedimientos de búsqueda. Sin embargo dado que su significado es general, puede obviamente aplicarse a otras áreas de la IA.

El caso de los Sistemas de Inferencia Difusa, hace particularmente evidente esta situación. Si nos detenemos a pensar en el ejemplo que se ha explicado en este capítulo, existen varios puntos oscuros en lo que hace a su validación a través de un método formal. Por ejemplo, El tipo de función de membresía de los adjetivos empleados en las reglas debe ser propuesta por el diseñador del sistema, los valores particulares también deben ser propuestos por el diseñador del sistema (digamos por ejemplo las pendientes y curvaturas particulares de cada curva), el método de defuzzificación ha de ser elegido por el diseñador también, y finalmente también los métodos de implicación a usar en un sistema.

Todo esto no solo muestra una aplicación particular de la heurística, sino que evidencia la fuerte naturaleza heurística de los Sistemas de Inferencia Difusa, lo que hace en algunos ámbitos, se cuestionen los fundamentos para el diseño y la construcción de estos sistemas. Por su parte los cultores de la Lógica Difusa contraponen a esta visión un tanto despectiva las aplicaciones exitosas logradas en diferentes ámbitos.