
UNIVERSIDAD POLITECNICA DE VALENCIA

Departamento de Ingeniería de Sistemas y Automática



Tesis Doctoral

**Diagnóstico de fallos
en sistemas industriales
basado en razonamiento
borroso y posibilístico.**

Autor: D. Julio César Ramírez Valenzuela

Director: Dr. D. Antonio Sala Piqueras

Valencia, 20 de abril de 2007

Diagnóstico de fallos
en sistemas industriales
basado en razonamiento
borroso y posibilístico.

D. Julio César Ramírez Valenzuela
Instituto Tecnológico de Nogales
Av. Tecnológico 911, Nogales, Sonora, México
E-Mail: juliorv@prodigy.net.mx

Preámbulo, agradecimientos.

Primeramente doy gracias a Dios que me ha permitido concluir esta tesis, desarrollada a lo largo de varios años de investigación doctoral. Agradezco también a mi director de tesis, Dr. Antonio Sala Piqueras quien me tuvo mucha paciencia y me fue guiando en los diferentes temas relevantes, que fueron conformando mi trabajo de investigación.

Doy gracias al Dr. Pedro Albertos, al Dr. Francisco Morant y a todos mis profesores del departamento de Ingeniería de Sistemas y Automática de la UPV, que me ayudaron en mi formación académica. Agradezco al personal administrativo del departamento por agilizar mis trámites para la lectura de mi tesis.

Debo también dar las gracias a los directivos del Instituto Tecnológico de Nogales México donde laboro, por todo el apoyo recibido, sin el cuál no hubiese podido culminar mis estudios de doctorado. Agradezco a la academia de electrónica de dicho instituto por su apoyo y amistad, también al MC. Pedro Ochoa.

Por último pero no menos importante a mi querida familia, que han sufrido mis ausencias. Gracias doy a mi esposa Sara por el ánimo y amor con el que siempre me motivó para seguir adelante, a mis hijos Josué, Joanna y Daniel por su amor.

Valencia, 20 de abril de 2007.

Resumen

Esta tesis aborda el diagnóstico de fallos en sistemas industriales por técnicas de Inteligencia Artificial, tratando en particular el razonamiento borroso y posibilístico. Inicialmente, se presentan los problemas a resolver en el diagnóstico de sistemas y después se plantean estrategias para abordarlos a partir de diferentes técnicas de Inteligencia Artificial, en donde destacamos los métodos relacionales borrosos que serán la base para nuestra aportación principal. También se han estudiado los sistemas expertos basados en lógica borrosa y que usan tablas de decisión, los sistemas expertos que combinan lógica borrosa con probabilidad y los sistemas de diagnóstico basados en redes Bayesianas. Se experimenta con varias técnicas de diagnóstico descritas en el estado del arte, haciendo combinaciones entre ellas. Una vez experimentadas y evaluadas las anteriores técnicas, vistos los inconvenientes que surgían, se decidió implementar una nueva metodología que diera una mejor solución al problema del diagnóstico. Esta metodología es el diagnóstico posibilístico borroso visto como un problema de optimización lineal. La metodología convierte los enunciados lingüísticos, que componen una base de reglas de un sistema experto borroso, en un conjunto de ecuaciones lineales a través de técnicas relacionales. Luego, estas ecuaciones se utilizan con algoritmos de programación lineal. Algunas modificaciones requieren programación cuadrática. Los resultados obtenidos en esta última aportación en una aplicación de análisis de aceites fueron satisfactorios, presentando al usuario una salida de diagnóstico fácil de interpretar, suficientemente exacta y teniendo en cuenta la incertidumbre en reglas y medidas.

Resum

La present tesi tracta el diagnòstic de fallades en sistemes industrials amb tècniques d'intel·ligència artificial, en particular el raonament borrós i possibilístic. Inicialment, es presenten els problemes a resoldre en el diagnòstic de sistemes i després es plantegen estratègies per a abordar-los a partir de diferents tècniques d'intel·ligència artificial, on destaquem els mètodes relacionals borrosos que seran la base de la aportació principal. També s'han estudiat els sistemes experts que combinen lògica borrosa amb probabilitat, i els sistemes basats en xarxes baessianes. S'experimenta amb diverses tècniques de diagnòstic descrites en l'estat de l'art fent combinacions entre elles. Una vegada experimentades i avaluades, vistos els inconvenients de les mateixes, es decideix implementar una nova metodologia per tractar de millorar les solucions existents. Aquesta metodologia és el diagnòstic possibilístic borrós vist com un problema d'optimització lineal. La metodologia converteix els enunciats lingüístics de les regles d'un sistema expert borrós en un conjunt d'equacions lineals a partir de tècniques relacionals. Aquestes equacions s'utilitzen amb algoritmes de programació lineal. Algunes modificacions requereixen programació quadràtica. Els resultats obtinguts en una aplicació d'anàlisi d'olis han segut satisfactoris i presenten una eixida fàcil d'interpretar, tenint en compte la incertesa en regles i mesuraments.

Abstract

This thesis discusses automatic fault diagnosis in industrial systems with artificial intelligence techniques, in particular fuzzy and possibilistic reasoning. Initially, the problems to be solved are presented and strategies to deal with them are suggested in the scope of artificial intelligence, with special emphasis in fuzzy relational models which will be the base of the main contribution. Expert systems which combine fuzzy logic and probability have been also studied, as well as Bayesian networks. Tests with the mentioned techniques have been carried out. After its evaluation, given the drawbacks that some of them had, a decision was made to implement a new methodology in order to improve the existing solutions. This methodology views possibilistic fuzzy diagnosis as an optimization problem. The methodology converts the linguistic assertions in the expert system rulebases in a set of linear restrictions from relational techniques. These equations are used with linear programming code. Some modifications require quadratic programming. The obtained results in a practical oil analysis application are promising, presenting easily interpretable outputs and taking into account uncertainty in rules and measurements.

Índice de contenidos

Índice general

I	Estado del arte	7
1.	Diagnóstico de fallos: generalidades	9
1.1.	Introducción	9
1.2.	Adquisición de la información	9
1.2.1.	Fiabilidad de la información	10
1.2.2.	Incertidumbre de la información	11
1.3.	Terminología usada en la supervisión, detección y diagnóstico de fallos . .	13
1.4.	Modelos del proceso	14
1.4.1.	Clases de procesos	14
1.4.2.	Modelos de conocimiento profundo o “ <i>Deep Knowledge</i> ”	15
1.4.3.	Modelos de conocimiento superficial “ <i>Shallow Surface</i> ”	16
1.5.	Supervisión	16
1.6.	Generación de síntomas	18
1.6.1.	Generación analítica de síntomas	18
1.6.2.	Generación heurística de síntomas	19
2.	Detección de fallos basado en modelos dinámicos	21
2.1.	Introducción	21
2.2.	Detección de fallos basado en el modelo	21
2.3.	Modelado de fallos	22

2.4.	Detección de fallos por estimación de parámetros	24
2.4.1.	Métodos de ecuaciones de error	24
2.4.2.	Métodos de error de salida	25
2.5.	Detección de fallos por medio de observadores	26
2.5.1.	Principios básicos para la construcción del generador residual en observadores	27
2.6.	Detección de fallos con ecuaciones de paridad	28
2.6.1.	Ecuaciones de paridad para procesos de una entrada y una salida (SISO)	29
2.6.2.	Ecuaciones de paridad para procesos de múltiples entradas y múltiples salidas (MIMO)	30
2.7.	Detección de fallos mediante el modelado de señales	31
3.	Lógica y sistemas borrosos: generalidades.	33
3.1.	Introducción.	33
3.2.	Lógica binaria.	34
3.3.	Sistemas basados en reglas.	35
3.3.1.	Verificación lógica. Anomalías.	35
3.4.	Cálculo proposicional borroso.	36
3.5.	Conjuntos borrosos.	37
3.6.	Significado de la borrosidad.	38
3.7.	Sistemas basados en reglas borrosas.	41
3.7.1.	Reguladores borrosos.	42
3.7.2.	Sistemas de diagnóstico borrosos	42
3.8.	Incertidumbre en el conocimiento	43
3.8.1.	Teoría de probabilidad	43
3.8.2.	Teoría de posibilidad	44
3.8.3.	Teoría de la evidencia	45

3.9.	Medidas intervalares borrosas	45
3.9.1.	Cantidades borrosas	46
3.9.2.	Intervalo borroso	46
3.9.3.	Interpretación	47
3.9.4.	Borrosificación	48
4.	Diagnóstico basado en técnicas de Inteligencia Artificial	51
4.1.	Introducción	51
4.2.	Métodos de clasificación para el diagnóstico	51
4.2.1.	Principios del método de clasificación	52
4.3.	Métodos basados en modelos	52
4.3.1.	Técnicas de los métodos basados en el modelo	52
4.3.2.	Método basado en el modelo de eventos discretos	53
4.4.	Métodos basados en relaciones	55
4.4.1.	Método relacional borroso aplicando FMECA	55
4.4.2.	Tasas de cobertura en los métodos relacionales	59
5.	Combinación de lógica borrosa y probabilidad	63
5.1.	Introducción	63
5.2.	Teoría probabilista Bayesiana	63
5.2.1.	Corolario de probabilidad	64
5.2.2.	Corolario de probabilidad condicional	64
5.2.3.	Teorema de Bayes	65
5.3.	Teoría de evidencia	66
5.3.1.	Aplicaciones de la teoría de Dempster-Shafer	68
5.3.2.	Combinación de evidencia	70
5.4.	Generalización de la teoría de Dempster-Shafer a conjuntos borrosos . . .	73
5.4.1.	Función de creencia generalizada	74
5.4.2.	Regla de combinación generalizada	80

6. Diagnóstico de fallos mediante redes Bayesianas	93
6.1. Introducción	93
6.2. Aplicaciones de redes Bayesianas	94
6.3. Redes Bayesianas	95
6.3.1. Representación del conocimiento en un dominio incierto	95
6.3.2. Representación de la distribución de probabilidad conjunta	98
6.3.3. Un método para construir redes Bayesianas	98
6.3.4. Representación de las tablas de probabilidad condicional	99
6.4. Inferencia en redes Bayesianas	101
6.4.1. Patrones de inferencia en redes Bayesianas	101
6.4.2. Separación dependiente de la dirección o separación-d	103
6.5. Inferencia probabilista en poliárboles	105
6.5.1. Apoyo causal	105
6.5.2. Apoyo evidencial	107
6.5.3. Apoyo causal y evidencial	109
6.5.4. Un ejemplo numérico	109
II Aportaciones	113
7. Aportaciones en técnicas de Inteligencia Artificial para el diagnóstico	115
7.1. Introducción	115
7.2. Diccionario de fallos borroso para el diagnóstico	116
7.2.1. Introducción a un sistema de diagnóstico basado en un diccionario de fallos	116
7.2.2. Diagnóstico lógico	117
7.2.3. Imprecisión en las medidas	118
7.2.4. Variables de contexto	119

7.2.5.	Diagnóstico predictivo a través de la tendencia de las variables . . .	120
7.2.6.	Simulación de un proceso	121
7.2.7.	Sumario sobre el sistema experto borroso	123
7.3.	Diagnóstico de sistemas mediante combinación de lógica borrosa y probabilidad	124
7.3.1.	Aplicaciones industriales de la teoría de evidencia	124
7.3.2.	Sistema que combina lógica borrosa y probabilidad	124
7.3.3.	Análisis de resultados por combinación evidencial	129
7.3.4.	Sumario de la combinación de lógica borrosa y probabilidad . . .	130
7.4.	Redes Bayesianas en el diagnóstico de sistemas con aprendizaje de parámetros	131
7.4.1.	Caso de estudio	131
7.4.2.	Descripción breve del programa	132
7.4.3.	Diagnóstico de fallos usando una red Bayesiana	134
7.4.4.	Análisis de resultados en el diagnóstico por redes Bayesianas . . .	136
8.	Diagnóstico posibilístico borroso como un problema de optimización lineal	139
8.1.	Introducción	139
8.2.	Optimización lineal	141
8.2.1.	Programación lineal	142
8.3.	Definiciones básicas	144
8.4.	Metodología de diagnóstico (caso binario)	145
8.4.1.	Incorporando observaciones	148
8.4.2.	Refinamientos en la metodología	149
8.4.3.	Diagnosis basada en la consistencia	151
8.5.	El enfoque borroso	153
8.6.	Marco de trabajo Posibilístico	156
8.7.	Bases de conocimiento posibilístico	157
8.7.1.	Resultados del diagnosis posibilístico	161
8.8.	Ejemplos adicionales	162
8.9.	Discusión comparativa	168

9. Aplicación del diagnóstico posibilístico borroso a la detección de fallos en motores de combustión interna mediante análisis de aceite	171
9.1. Introducción	171
9.2. Diagnóstico mediante análisis de aceite	172
9.3. Definición de la base de conocimiento	173
9.4. Preproceso de medidas y escenario de diagnóstico	175
9.4.1. Preproceso de medidas	175
9.4.2. Escenario de diagnóstico	175
9.5. Construcción de las reglas de inferencia	176
9.6. Diagnóstico de fallos	179
9.6.1. Adquisición de observaciones σ_j	179
9.6.2. Función de optimización lineal	183
9.6.3. Discusión sobre algunos resultados	186
9.7. Conocimiento incierto y posibilidad condicional	189
9.7.1. Conocimiento incierto	189
9.7.2. Posibilidad condicional	190
10. Conclusiones	199
10.1. Conclusiones del trabajo	199
10.2. Trabajos futuros	200

Índice de figuras

1.1. Definición de un conjunto borroso	12
1.2. Exactitud alcanzada por los modelos, dependiente del grado de cantidad o calidad del conocimiento del proceso	15
1.3. Detección y diagnóstico de fallos basado en el conocimiento	18
2.1. Detección de fallos basado en el modelo del proceso, donde N son las perturbaciones	22
2.2. Dependencia de los fallos con respecto al tiempo: a) abrupta, b)incipiente, c) intermitente	22
2.3. Modelos básicos de fallos: a) fallos aditivos, b) fallos multiplicativos . . .	23
2.4. Modelo de proceso estático no lineal	23
2.5. Modelo para la estimación de parámetros por ecuaciones de error	24
2.6. Modelo para la estimación de parámetros por el error de salida	25
2.7. Observador de orden completo para la generación de residuos	27
2.8. Observador de orden completo para la generación de residuos	28
2.9. Ecuaciones de paridad, método de error de salida	29
2.10. Ecuaciones de paridad, método de ecuación de error	29
2.11. Método de ecuaciones de paridad para modelos en espacio estado MIMO; D' : filtro de diferenciación	30
3.1. Temperatura agradable	37
4.1. Diagnóstico de fallos usando métodos de clasificación	52
4.2. Arquitectura para el diagnóstico	54

4.3.	d es una explicación posible: las observaciones son consistentes con sus manifestaciones predichas, presentes y ausentes	56
4.4.	Solapamiento de conjuntos borrosos	57
4.5.	d es una explicación relevante: algunas de las manifestaciones presentes, predichas con anterioridad, son realmente observadas en el presente . . .	58
4.6.	Modelo relacional que muestra los efectos de contexto no especificado $M(d)^{+?}$	60
5.1.	Particionamiento de un evento en dos subconjuntos mutuamente excluyentes	65
5.2.	Rueda de la fortuna dividida en tres áreas	68
5.3.	Descripción gráfica de cómo se obtiene $f_{B,A}(\alpha_i)$. Obtención de: a) $f_{B,A}(0,25)$, b) $f_{B,A}(0,5)$, c) $f_{B,A}(0,75)$, d) $f_{B,A}(1)$	80
5.4.	Descripción gráfica de los elementos focales de B y C	81
5.5.	Descripción de los subconjuntos borrosos	86
5.6.	Subconjuntos usados en la combinación de dos fuentes evidenciales	90
6.1.	Apilamiento de bloques por un brazo de robot	96
6.2.	Ejemplo de una red Bayesiana	97
6.3.	Ejemplo de RB que observa <i>separación-d</i>	103
6.4.	Un poliárbol típico	105
6.5.	Ejemplo de poliárbol	109
7.1.	Funciones de pertenencia	117
7.2.	Pendientes m_u y Ff en la tendencia de una variable	120
7.3.	Proceso industrial a simular	122
7.4.	Fallo 4 cuando $k = 14$	123
7.5.	Sistema de dos depósitos interconectados	125
7.6.	Funciones de pertenencia para q_2	125
7.7.	Errores entre la creencia y la regla	127
7.8.	Funciones de pertenencia no borrosas A y B	128

7.9. Gráfica de credibilidad y plausibilidad	129
7.10. Estructuras de red propuestas en el aprendizaje supervisado	133
7.11. Funciones de pertenencia que caracterizan al nivel de h_1 bajo y al sensor h_1	134
7.12. Gráfica de plausibilidad y credibilidad del fallo de V_i	136
8.1. Una base de conocimiento básica. (\longrightarrow) la flecha sólida: asociaciones ciertas; ($- - - \rightarrow$) flecha discontinua: asociaciones inciertas.	146
8.2. Descripción gráfica de un problema de diagnóstico. (\longrightarrow) la flecha sólida: asociaciones ciertas; ($- - - \rightarrow$) flecha discontinua: asociaciones inciertas; ($\cdots *$) línea punteada con estrella: desigualdades (la estrella denota a en $a \leq b$).	152
8.3. Conocimiento borroso incierto (la zona fiable aparece sombreada).	155
8.4. Distribución de posibilidad proveniente de restricciones de igualdades “suavizadas asimétricamente”.	159
8.5. Distribución de posibilidad con 4 variables artificiales: restricciones de sector suavizadas.	160
8.6. Posibilidades marginales en el caso 1. Abscisas: severidad (de izquierda a derecha: d_1, d_2, d_3); ordenadas: posibilidad	166
8.7. Posibilidades marginales en el caso 2	167
8.8. Posibilidades marginales en el caso 3	167
9.1. Diagrama de relación causa-efecto del sistema de diagnóstico por el análisis de aceites usados	181
9.2. Gráficas de posibilidad condicional de los desórdenes del caso <i>caso 1</i> . Abscisas: severidad; ordenadas: posibilidad	193
9.3. Gráficas de posibilidad condicional de los desórdenes del caso <i>medidas2</i> . Abscisas: severidad; ordenadas: posibilidad	194
9.4. Gráficas de posibilidad condicional de los desórdenes del caso <i>medidas3</i> . Abscisas: severidad; ordenadas: posibilidad	195
9.5. Gráficas de posibilidad condicional de los desórdenes del caso <i>medidas4</i> . Abscisas: severidad; ordenadas: posibilidad	196

Índice de cuadros

6.1. Clasificación de redes probabilistas	94
6.2. Ejemplo de parámetros de ruido	101
7.1. Diccionario de fallos (límites mínimos)	118
7.2. Diccionario de fallos (límites máximos)	118
7.3. Límites mínimos en el diccionario para diagnóstico del sistema de depósitos	126
7.4. Límites máximos en el diccionario para diagnóstico del sistema de depósitos	126
9.1. Relación desórdenes\síntomas	177
9.2. Casos de medidas de 5 muestras de aceite usado	180
9.3. Donde se muestran las observaciones σ_j	187
9.4. Resultado de desórdenes encontrados por el sistema	188

Introducción

Introducción

Motivación

Con las máquinas y los procesos industriales, el diagnóstico ha tomado importancia. En los últimos años la automática ha contribuido grandemente en estos avances, formalizando esta área de la investigación.

El primer problema que se encuentra en el diagnóstico de fallos en los sistemas industriales, es la adquisición de la información, problema que puede resultar complejo ya sea el caso de información analítica o heurística. En el caso de la información analítica o cuantitativa los instrumentos de medición suelen proveer información que dependiendo de la calidad de los sensores introducen mayor o menor grado de incertidumbre en las medidas. En el caso de la información cualitativa o heurística la incertidumbre es todavía mayor, pues depende del punto de vista del ser humano, cuyos criterios pueden ser subjetivos y varían de una persona a otra.

Tras la adquisición de la información la siguiente etapa, es la inferencia o razonamiento de diagnóstico. En esto, se han desarrollado muchos tipos de técnicas en el razonamiento de diagnóstico las del tipo analítico, que se basan en los modelos matemáticos de los procesos industriales y las del tipo de conocimiento superficial, que se basan en las relaciones causa-efecto de dichos sistemas.

Las dos técnicas citadas no están exentas de problemas y también suelen introducir incertidumbre en los resultados de diagnóstico que ofrecen. En las técnicas analíticas basadas en los modelos matemáticos, aunque son por su naturaleza robustas, tienen el problema de que los modelos matemáticos son difíciles de adquirir y una vez que se tienen, no hay garantía de que permanezcan sin cambios, además, pueden ser afectados por factores externos de manera que su fiabilidad disminuye.

El problema en las técnicas basadas en el conocimiento causa-efecto, son los algoritmos de inferencia o razonamiento de diagnóstico. La mayoría de estos algoritmos son desarrollados con técnicas de Inteligencia Artificial, como los sistemas expertos, la lógica borrosa, redes neuronales etc., aunque los resultados de estos algoritmos son buenos, no dejan de ser aproximaciones de optimización que implican un cierto grado de incertidumbre.

Como último problema en los sistemas de diagnóstico tenemos la diagnosis o interpretación de resultados que la inferencia de diagnóstico nos ofrece. Este último problema está relacionado con la toma de decisiones, tarea que puede resultar compleja, sobre todo si el sistema de diagnóstico no ha dado un buen trato a la incertidumbre en las diferentes etapas del diagnóstico.

Objetivos

Vista la anterior problemática, se plantea un estudio de las técnicas de Inteligencia Artificial que han probado ser más efectivas en el diagnóstico de fallos, como son: los sistemas expertos borrosos, basados en tablas de conocimiento, los sistemas expertos que combinan lógica borrosa con probabilidad, y las redes Bayesianas. En estas técnicas se dará un trato a la incertidumbre presente en la información, representando las medidas en forma de intervalo.

El objetivo principal de esta tesis es presentar una nueva metodología de diagnóstico, que consiste en un sistema experto posibilístico borroso visto como un problema de optimización lineal. Se explicará la fundamentación teórica de esta técnica, basada en métodos relacionales borrosos y en optimización lineal.

Otro objetivo, será probar esta técnica a través de la implementación de un sistema de diagnóstico de motores de combustión interna por el análisis de aceites usados. Donde se demostrará además de lo fácil que resulta la mecanización y aprendizaje de esta técnica, lo óptimo de sus resultados y lo fácil que es la interpretación de los mismos, con lo cual se disminuye la incertidumbre en el diagnóstico considerablemente.

Parte I

Estado del arte

Capítulo 1

Diagnóstico de fallos: generalidades

1.1. Introducción

En este capítulo se trata el primer problema que se encuentra en la detección y diagnóstico de fallos, que es la adquisición de la información, problema que puede resultar complejo ya sea el caso de información analítica o heurística. Por el lado de la información analítica, es sabido que a menor grado de calidad en los sensores, aumenta la incertidumbre al disminuir la exactitud y la precisión en las mediciones. Si consideramos la información heurística, se puede apreciar que la incertidumbre es algo inherente en ella, pues depende del punto de vista del operador humano, que frecuentemente utiliza un razonamiento aproximado para la solución de problemas complejos, con un grado aceptable de eficacia.

1.2. Adquisición de la información

El conocimiento que podemos adquirir acerca del comportamiento de un sistema físico se fundamenta principalmente en la adquisición y valoración de dos tipos de información (Chacón 2001):

- *Cuantitativa*, que se adquiere a través de los diferentes instrumentos de medición de las variables que caracterizan el funcionamiento del sistema.
- *Cualitativa*, adquirida por humanos a través de los órganos sensoriales y procesada por el cerebro, que suministra información generalmente en forma lingüística.

En el proceso de detección y diagnóstico, no se puede despreciar ninguna clase de información, debido a que las dos son fundamentales para la generación de los síntomas de fallos.

1.2.1. Fiabilidad de la información

Tanto en la vida cotidiana como en el mundo técnico, la información que se tiene y se manipula es habitualmente “incompleta”. Nunca se dispone de información en toda su integridad o con la “certeza” necesaria. Existen diferentes tipos de errores que pueden contribuir a la no fiabilidad de la información, como son:

- *Ambigüedad*. Una medición o apreciación de un efecto puede ser interpretado de diferentes maneras.
- *Ausencia de información*. Es cuando la información se ha perdido o no se dispone de ella.
- *Incorrección*. Es cuando la información es equivocada. Una hipótesis es una afirmación hecha inicialmente, pero que debe verificarse. Una hipótesis incorrecta puede ser falsa positiva o falsa negativa, en el caso falsa positiva, se afirma la hipótesis cuando no es verdadera, y en el caso falsa negativa, se rechaza cuando es verdadera.

Los siguientes son errores resultantes de las mediciones:

- *Exactitud*, esta relacionada con el valor verdadero de la magnitud medida. Se denomina exactitud a la capacidad de un instrumento de medir un valor cercano al valor de la magnitud real. Exactitud implica precisión. Pero no al contrario. Esta cualidad también se encuentra en instrumentos generadores de magnitudes físicas, siendo en este caso la capacidad del instrumento de acercarse a la magnitud física real.
- *Precisión*, es la capacidad de realizar medidas similares. Por ejemplo: varias medidas son como flechas disparadas hacia un objetivo. La exactitud describe la proximidad de las flechas al centro del objetivo. Las flechas que impactaron más cerca del centro se consideran más exactas. Cuanto más cerca están las medidas a un valor aceptado, más exacto es un sistema.

La precisión, en este ejemplo, es el tamaño del grupo de flechas. Cuanto más cercanas entre sí estén las flechas que impactaron el objetivo, más preciso será el sistema. Hay que notar que el hecho de que las flechas estén muy cercanas entre sí es independiente al hecho que estén cerca del centro del objetivo. En sí, se puede decir que la precisión es el grado de repetibilidad del resultado. La precisión es normalmente caracterizada en términos de desviación estándar de las medidas. La precisión no puede identificarse en medidas de diferentes magnitudes, en cambio, la exactitud sí lo puede.

- *Fiabilidad*, esta relacionada con medidas erráticas, es decir que no son constantes, fluctúan, algunas veces son correctas y algunas no.

- *Aleatorio*, fluctuaciones aleatorias alrededor del valor medio que induce incertidumbre en el valor medio. Otros tipos de errores aleatorios pueden ser causados por el movimiento Browniano, el ruido electrónico debido a efectos térmicos, etcétera.
- *Sistemático*, no es un error de naturaleza aleatoria pero se introduce a causa de una falta de exactitud. Por ejemplo una falta de calibración de una regla en donde sus graduaciones o divisiones son más pequeñas de lo normal producen un error sistemático que indica lecturas que son más grandes que lo normal.

1.2.2. Incertidumbre de la información

La incertidumbre se puede considerar como la falta de información adecuada para tomar una decisión. La incertidumbre es un problema debido a que puede impedir que se tome una decisión y al contrario, puede provocar que se tome una decisión incorrecta. La incertidumbre puede estar asociada a una sentencia no vaga: frecuentemente no existe un intermedio entre la verdad y la falsedad, pero no se puede estar seguro que la sentencia sea verdadera o sea falsa. A veces un hecho entre más preciso puede ser menos cierto. Existen ciertas representaciones de la incertidumbre, como son:

- *La teoría de las probabilidades*. El origen de las probabilidades se ha definido exclusivamente a partir de estadísticas, lo que ha conformado la escuela de los “objetivistas” o “frecuentistas”. Estos basan las probabilidades en la repetición de un evento. Existe también la escuela de los “subjetivistas”, que proponen que es posible definir una probabilidad que no sea necesariamente una frecuencia de aparición de un evento, sino más justamente una evaluación subjetiva de la aparición de este hecho.

Aunque las probabilidades se expresan en números, el mérito del cálculo de probabilidades se basa en que suministra la forma de articular y manipular relaciones cualitativas que se han encontrado útiles en el discurso normal. Se distinguen 4 relaciones principales que forman según Pearl (Pearl 1988) las bases primitivas del lenguaje de las probabilidades:

- *Condicionalidad*: Sí Tomás esta enfermo, el no puede correr.
- *Verdad o certeza*: Tomás es más probable que corra que camine.
- *Dependencia*: Cuando Tomás corre, dependiendo del clima, el puede enfermar.
- *Casualidad*: El estar enfermo causa que Tomás no pueda correr.

La mecánica de la teoría de las probabilidades permite extraer numerosos esquemas naturales de razonamiento como los basados en la no-monotonía que dependen uno a uno del contexto, la abducción o explicación, la noción de disculpa o “*explaining away*”, entre otras. Sin embargo, la teoría de la probabilidad se observa como una plataforma demasiado normativa y rígida para tomar en cuenta todos los aspectos del razonamiento y juicios cargados de incertidumbre del mundo real.

Un modelo basado en la teoría de las probabilidades se adapta muy bien para la expresión de información precisa pero dispersa. Una vez se ha perdido la precisión el modelo pierde validez.

- *La teoría de las funciones de creencia de Dempster-Shafer.* Intenta el modelado de la incertidumbre mediante un rango de probabilidades más que a través de un número de probabilidad único (Caro 1999). Las funciones de creencia tienen una base teórica sólida y propone una representación más sofisticada de la incertidumbre, conservando las bases sólidas de la teoría de las probabilidades. Sin embargo, a pesar de sus buenas características, esta teoría tiene algunas limitantes como que: las hipótesis deben ser mutuamente exclusivas y lo más completas posibles. No obstante esto, la estaremos aplicando en esta aportación aprovechando sus buenas características.
- *Conjuntos borrosos y razonamiento aproximado.* En esta teoría su autor (Zadeh 1965) trata principalmente de la cuantificación y el razonamiento utilizando el lenguaje natural en donde muchas palabras tienen significados “ambiguos”. En la siguiente gráfica (figura 1.1) se muestra el concepto de conjunto borroso.

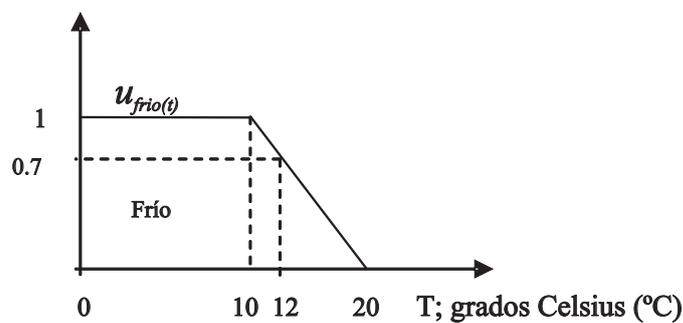


Figura 1.1: Definición de un conjunto borroso

En el ejemplo que se muestra (ver figura 1.1), una sentencia tendrá un grado de verdad más o menos grande según corresponda o no a la realidad. Por ejemplo: “Hace frío” es una afirmación parcialmente verdadera, el grado de verdad es 0.7 si la temperatura es de 12 grados Celsius (°C). Es necesario señalar que la definición de conjunto difuso o borroso puede variar sensiblemente de un individuo a otro y de un contexto a otro, es una definición que puede ser subjetiva. La lógica borrosa es la teoría más general de la incertidumbre que ha sido formulada. Tiene un amplio espectro de aplicación debido principalmente al principio de extensión.

- *La teoría de las posibilidades,* (Zadeh 1978) consiste en utilizar los conjuntos borrosos no solamente para representar el aspecto gradual de ciertas propiedades, sino igualmente para representar la incertidumbre presente en las informaciones incompletas. En esta teoría, la incertidumbre se describe a través de dos medidas duales: la posibilidad y la necesidad. La “Posibilidad (Π)” se define en lenguaje finito de proposiciones entre $[0,1]$ con los siguientes axiomas.

$$\forall p, \forall q, \Pi(p \vee q) = \max(\Pi(p), \Pi(q))$$

$$\max[\Pi(p), \Pi(\neg p)] = 1$$

La medida de “Necesidad (N)” esta asociada por dualidad con la medida de posibilidad π a través de la siguiente expresión.

$$\forall p, N(p) = 1 - \Pi(\neg p)$$

Y también usa los siguientes axiomas:

$$\forall p, \forall q, N(p \vee q) = \min(N(p), N(q))$$

$$\min[N(p), N(\neg p)] = 0$$

1.3. Terminología usada en la supervisión, detección y diagnóstico de fallos

A medida que nos introducimos en el campo del diagnóstico y detección de fallos encontramos que la terminología no es consistente. Por ejemplo ¿Cuál es la diferencia entre monitorizado y supervisión? o ¿Cuál es la diferencia entre fallo, mal funcionamiento y error? Podemos ver en (Isermann 1997b) algunas definiciones que se obtuvieron del diccionario RAM (*Reliability, Availability and Maintainability*) y en otros estándares de la automática. A continuación se enumeran algunas definiciones que el comité de automática sugiere:

- *Fallo*: Una desviación no permitida de por lo menos una propiedad característica o parámetro de las condiciones aceptables / usuales / normales del sistema.
- *Fuera de servicio*: Una interrupción permanente de la habilidad de un sistema para realizar las funciones específicas que de él se requirieren.
- *Funcionamiento defectuoso*: Una irregularidad intermitente en la habilidad de un sistema para realizar las funciones específicas que de él se requirieren.
- *Error*: Una desviación entre un valor medido o computado (de una variable de salida) y el valor verdadero, especificado o teóricamente correcto.
- *Perturbación*: Una entrada que actúa en un sistema produciendo una desviación temporal del estado actual, que no es considerado fallo.
- *Residuo*: Un indicador de fallo, basado en una desviación, entre las medidas y cálculos basado en modelo de ecuaciones.
- *Síntoma*: Cambio de una cantidad respecto a un comportamiento normal.
- *Detección del fallo*: Determinar la presencia de un fallo en un sistema y el tiempo de detección del mismo.

- *Aislamiento del fallo*: Después de la detección del fallo, determinar la clase y ubicación, pero sin identificar que fallo es.
- *Identificación del fallo*: Después del aislamiento del fallo, determinar el tamaño y variación en tiempo del comportamiento de un fallo. Se especifica claramente si es un fallo conocido y cuál es. En caso de ser desconocido se somete a análisis para determinar si es o no fallo. En caso de ser un fallo nuevo se le identifica y clasifica.
- *Diagnóstico del fallo*: Después de la detección del fallo, se determina clase, tamaño, localización y tiempo de detección del fallo. Incluye aislamiento del fallo e identificación. Es la etapa donde se toma una decisión con respecto a la veracidad, gradualidad y gravedad del fallo.
- *Monitorizado*: Una tarea en tiempo-real continua, que determina las condiciones de un sistema físico, almacenando información, reconociendo e indicando anomalías en el comportamiento.
- *Supervisión*: Monitorizado de un sistema físico, tomando acciones apropiadas para mantener el funcionamiento en el caso de fallos.
- *Protección*: Medios por los que, si es posible, un comportamiento potencialmente peligroso del sistema se suprime, o medios por los que, las consecuencias de un comportamiento peligroso se evitan.

1.4. Modelos del proceso

Un modelo es una descripción de un sistema que captura propiedades de interés para un cierto propósito. Dependiendo de dicho propósito podrán hacerse determinadas aproximaciones o simplificaciones, y seleccionarse uno u otro tipo de modelo. Debe tenerse en cuenta que un modelo perfecto no existe y que cualquier modelo representa una imagen parcial del comportamiento del sistema (Sala 2001a).

1.4.1. Clases de procesos

En los procesos industriales las variables o parámetros que nos muestran el comportamiento del sistema pueden ser de tipo cuantitativo o cualitativo, numéricas o lingüísticas. Estas variables tanto las internas como las de entrada o salida, generalmente observan cierta imprecisión e incertidumbre. Ante esto podemos clasificar el tipo de proceso por el grado de conocimiento que se tenga de las variables del sistema. A medida que el grado de precisión crece, el grado de exactitud en el modelo también crece, según podemos observar en la clasificación en la figura 1.2 propuesta por R. Isermann (Isermann 1998).

1.4.2. Modelos de conocimiento profundo o “*Deep Knowledge*”

Proporcionan información de gran riqueza y calidad, y que por supuesto son difíciles y costosos de adquirir. Los más atractivos debido a su carácter sintético y a su poder de predicción son los modelos físicos, donde la descripción de su comportamiento estático y dinámico por medios matemáticos juega un papel importante en el diseño enfocado tanto para control como para diagnóstico y detección de fallos. Se pueden distinguir los siguientes casos:

- A.1 Ecuaciones con parámetros conocidos.
- A.2 Ecuaciones con parámetros parcialmente conocidos.
- A.3 Ecuaciones con parámetros desconocidos.

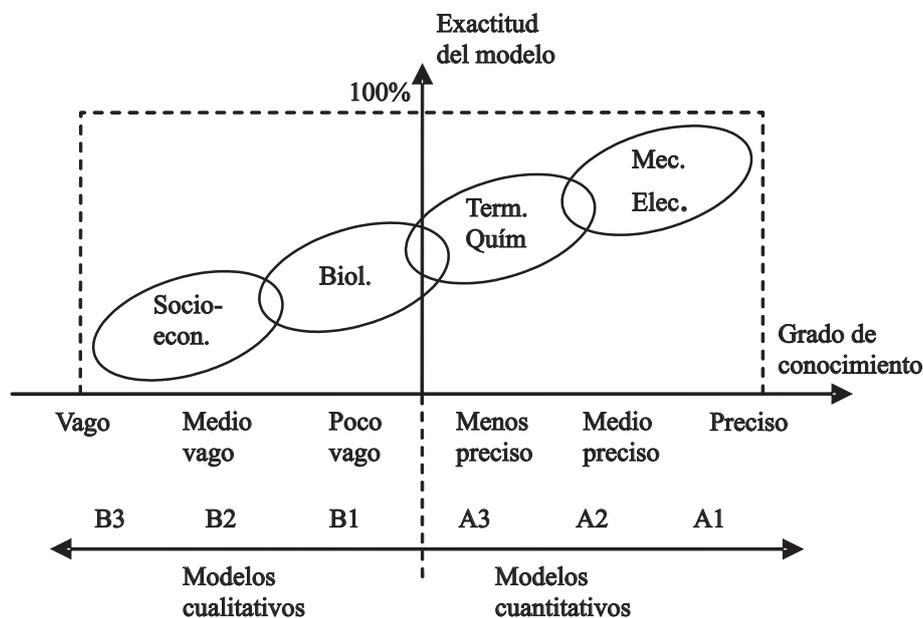


Figura 1.2: Exactitud alcanzada por los modelos, dependiente del grado de cantidad o calidad del conocimiento del proceso

En la práctica, el caso A.2 aparece más frecuentemente que A.1 y A.3 si la estructura es conocida, los parámetros desconocidos pueden ser determinados experimentalmente por estimación de parámetros y métodos de identificación. En el campo de modelado de procesos cuantitativos se ha alcanzado un estado bastante maduro y se encuentran en el mercado atractivas herramientas de software para este propósito. Sin embargo, si el comportamiento interno no puede expresarse por leyes físicas en forma de ecuaciones, en cuanto a un “no bien definido” proceso como el secado, vulcanizando, reacciones químicas especiales, o procesos biológicos, una parte de la información cualitativa, en las causalidades puede ser expresada como reglas:

SI{condición}ENTONCES{CONCLUSIÓN}

1.4.3. Modelos de conocimiento superficial “*Shallow Surface*”

Preferido por algunos expertos en diagnóstico (más que en detección) por su relativa simplicidad de adquisición. Sin embargo, no describen el sistema a diagnosticar de una manera rica, estableciendo las relaciones causa-efecto con sentido “profundo”, por lo que la calidad del conocimiento que suministran no es muy elevada, es de naturaleza empírica, compilada o heurística. De esta clase de modelos podemos identificar los siguientes:

- *Modelo relacional FMECA* (Cayrac 1996). A cada anomalía se le asocia un conjunto de manifestaciones o síntomas, es decir, se establece la relación “Causa-Efecto” entre las anomalías y sus síntomas. Sus siglas FMECA significan: *Failure Mode Effects and Criticality Analysis*. Sólo los fallos previstos “*a priori*” se toman en cuenta (lo cual es una desventaja). Para que tenga una aplicabilidad práctica, se debe escoger un conjunto de fallos lo suficientemente amplio, para que pueda cubrir un gran número de problemas potenciales. Aunque es de naturaleza binaria se puede borrosificar haciéndolo apto para manejar la incertidumbre.
- *Árboles de fallos o Fault Tree Analysis (FTA)*. Es una técnica de análisis de diseño de la fiabilidad y/o seguridad que parte de la hipótesis de los efectos de los fallos, referidos a la parte de arriba (“*top events*”). Es también una forma de conocimiento estructurado de los diferentes tipos de fallos que se reagrupan en forma jerárquica siguiendo en cada caso particular árboles de descomposición AND/OR. La raíz del árbol es un problema de naturaleza muy general, que se refina a medida que se va descendiendo en la jerarquía.

1.5. Supervisión

Dentro del control automático de sistemas técnicos, las funciones de supervisión sirven para indicar los estados indeseables o no permitidos de los procesos, y para tomar acciones apropiadas que permitan mantener el funcionamiento y evitar daños o accidentes. En supervisión podemos distinguir las siguientes funciones:

- a) *monitorizado*: las variables medibles son revisadas con respecto a sus tolerancias, y en caso de ser necesario se generan alarmas para el operador.
- b) *proteccion automatica*: en el caso de procesos peligrosos, la función de monitorizado automáticamente inicia una acción apropiada para contrarrestar los efectos del fallo. Lo anterior con el fin de evitar en la medida de lo posible los daños.
- c) *supervisión con diagnóstico de fallos*: basándose en las variables medidas, se generan los síntomas vía detección de fallos. Se lleva acabo la tarea de diagnóstico y se toman las decisiones para corregir los fallos o contrarrestar sus efectos nocivos.

Los métodos clásicos a) y b) son adecuados para la supervisión completa de los procesos. Para configurar las tolerancias (en los valores de umbral) se tiene que hacer un compromiso entre el tamaño de las desviaciones anormales y las alarmas innecesarias. Esto por las fluctuaciones normales de las variables.

Generalmente se usan los valores de umbral, los cuales trabajan bien si el proceso se encuentra en estado estable. Sin embargo, el problema aparece, cuando el punto de operación del proceso cambia rápidamente. En el caso de bucle cerrado, los cambios en el proceso son cubiertos por las acciones de control y no pueden ser detectados por las señales de salida. Esto también dependerá de que las entradas de control del proceso permanezcan en un nivel normal. En consecuencia, se puede decir, que los sistemas en bucle cerrado esconden la detección temprana de fallos en un proceso.

La ventaja de los métodos de supervisión clásicos de monitorizado y protección automática (a, b), es su simplicidad y confiabilidad. Sin embargo, ante un cambio repentino de una característica, o ante un fallo que imperceptiblemente se incrementa a través de un largo periodo de tiempo, estos métodos pueden fallar. Lo cuál significa que son poco sensibles a cambios que incrementan su gradualidad a través de un largo periodo de tiempo. En las técnicas clásicas no se cuentan con las herramientas suficientes para profundizar en el análisis de un fallo. Por lo tanto, un diagnóstico del fallo en profundidad normalmente no es posible

Por consiguiente, son necesarios métodos avanzados de supervisión y diagnóstico de fallos (c), que se basen en variables medidas y características calculadas (residuos, varianzas, medias, historia del fallo, frecuencias de aparición, etc.), adecuadas para la generación de síntomas, de manera que cuando un cambio en ellas es detectado, se realice un diagnóstico apropiado que genere acciones correctivas adecuadas. Los requerimientos que los métodos avanzados de supervisión y diagnóstico deben cumplir son:

- Detección temprana de fallos pequeños con abrupto o incipiente comportamiento temporal.
- Diagnóstico de fallo en el accionador, componentes del proceso o sensores.
- Detección de fallos en lazo cerrado.
- Supervisor de procesos en estado transitorio.

La meta de la detección temprana y el diagnóstico, es tener tiempo suficiente para contrarrestar las consecuencias que pudiera traer un fallo a través de operaciones, reconfiguración, mantenimiento o reparación. La detección temprana puede ser lograda adquiriendo más información, sobre todo, usando la relación entre las cantidades medidas en la forma de modelos matemáticos. Para el diagnóstico de fallos, el conocimiento de las relaciones de causa-efecto debe ser usado.

1.6. Generación de síntomas

En la figura 1.3 se muestra un esquema global sobre la detección y diagnóstico de fallos basado en el conocimiento. Según Isermann (Isermann 1997a) las tareas principales pueden subdividirse en generación analítica de síntomas, generación heurística de síntomas y diagnóstico de fallos. En la generación de síntomas se encuentra incluida la detección de fallos. En el capítulo de detección de fallos basado en modelos dinámicos, se estudiará con respecto a los modelos matemáticos de los procesos, de donde derivan los síntomas analíticos.

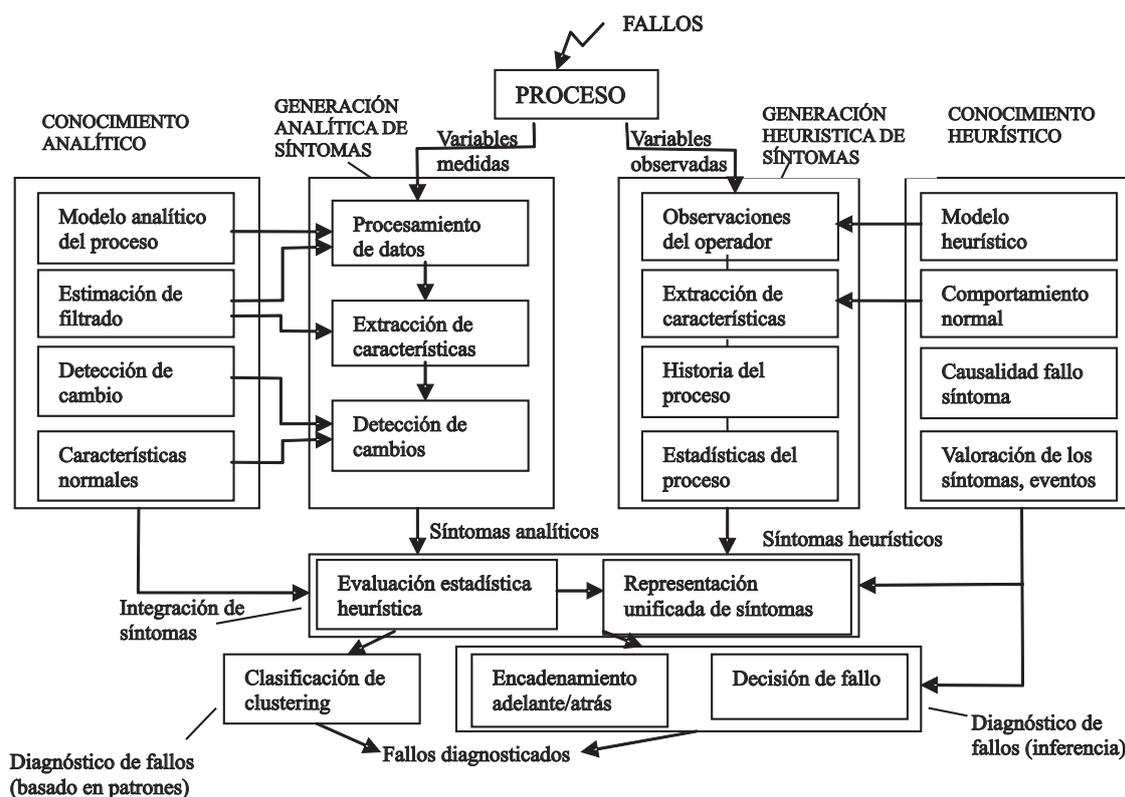


Figura 1.3: Detección y diagnóstico de fallos basado en el conocimiento

1.6.1. Generación analítica de síntomas

En la generación analítica de síntomas, el conocimiento analítico del proceso debe producir información cuantificable. Para lograr esto, se deben procesar datos basados en variables medidas del proceso. Primero se generan los valores característicos por:

- *Revisión de los valores de umbral*, de las señales directamente medibles. Los valores característicos son las tolerancias excedidas por las señales.

- *Análisis de señales*, directamente medidas por el uso de modelos de señales como las funciones de correlación, espectros de frecuencia, *AutoRegressive Moving Average* (ARMA), valores característicos (variancias, amplitudes, frecuencias, o parámetros de modelo).
- *Análisis de proceso*, donde se usan los modelos matemáticos del proceso junto con la estimación de parámetros, estimación de estados y los métodos de ecuaciones de paridad. Los valores característicos son parámetros, variables de estado o residuos.

En algunos casos, pueden extraerse características especiales de estos valores, características como son: coeficientes del proceso físicamente definidos o residuos transformados o filtrados. Estas características son después comparadas con las características del proceso sin fallos. Luego se aplican métodos de detección de cambios y clasificación. Los cambios resultantes (discrepancias) en las señales medidas descritas anteriormente, se consideran como los síntomas analíticos.

1.6.2. Generación heurística de síntomas

En la detección de fallos, tanto los síntomas analíticos como los síntomas heurísticos tienen importancia, aunque en la detección temprana, los síntomas analíticos juegan un papel más relevante. Hasta ahora hemos descrito los síntomas analíticos, que son determinantes para la detección de fallos. A continuación se hace la descripción de los síntomas heurísticos, que junto con los síntomas analíticos, completan el conjunto necesario de síntomas para el diagnóstico.

Los síntomas heurísticos, que son los efectos causados por los fallos, son producidos por el uso de información cualitativa de parte de un operador humano. A través de la observación e inspección humana, se obtienen valores característicos heurísticos en la forma de ruidos especiales, colores, olores, vibraciones, gasto y desperdicio, entre otros.

Los datos históricos del proceso en la forma de mantenimiento realizado, reparaciones, fallos anteriores, datos de pruebas de vida y pruebas de carga, constituyen una fuente completa de información heurística. También se pueden usar datos estadísticos adquiridos de la experiencia, con el mismo proceso o con otro similar.

De esta manera, son generados los síntomas heurísticos, que pueden ser representados por variables lingüísticas como pequeño, medio, grande o como números vagos que estén alrededor de cierto valor. No sólo los síntomas heurísticos pueden representarse en forma borrosa, sino que también los síntomas analíticos pueden representarse en funciones de membresía, junto con síntomas heurísticos de manera unificada, con un rango de $0 \leq \mu(s_i) \leq 1$, en las funciones de pertenencia de los síntomas unificados s_i .

Conclusiones del capítulo

En este capítulo se ha dado una introducción al diagnóstico de fallos, donde especificamos la terminología usada y las definiciones básicas del diagnóstico de fallos que estaremos empleando a lo largo de la tesis. También se mencionaron las diferentes herramientas teóricas relacionadas con este campo; que tipo de técnicas son usadas para auxiliar en el diagnóstico de fallos. Se trataron los tipos de modelos de sistemas para el diagnóstico disponibles y las diferentes etapas del diagnóstico. Por último se hizo mención del esquema general de diagnóstico, haciendo énfasis en la generación analítica y heurística de síntomas.

Capítulo 2

Detección de fallos basado en modelos dinámicos

2.1. Introducción

En este capítulo se tratan los métodos de detección fallos por el análisis de procesos industriales que se representan por medio de un modelado dinámico. La tarea del diagnóstico de fallos consiste en determinar el tipo, tamaño y localización de los fallos, así como su tiempo de detección, basado en los síntomas analíticos y heurísticos.

En este capítulo se estudia la detección de fallos más que su diagnóstico. Los métodos analíticos se consideran más precisos debido a que se fundamentan en un modelo matemático, por lo que resultan muy útiles en la detección de fallos, sobre todo cuando los síntomas son difíciles de percibir por instrumentos de medición convencionales. En esto último, las ecuaciones de residuos y paridad resultan bastante útiles como se estudiará a continuación.

2.2. Detección de fallos basado en el modelo

En la figura 2.1 según R. Isermann (Isermann 1997a), se muestra la estructura básica de un sistema para la detección de fallos basado en el modelo. Se detectan los fallos en el proceso, actuadores, y sensores, usando la dependencia entre las diferentes señales a medir.

Se basan en señales de entrada U y señales de salida Y , los métodos de detección generan residuos r , estimaciones de parámetros Θ o estimaciones de estado \hat{x} , las cuales son llamadas características, que se comparan con las características del sistema en estado normal. Los síntomas S son los cambios detectados en las características con respecto a las características del estado normal.

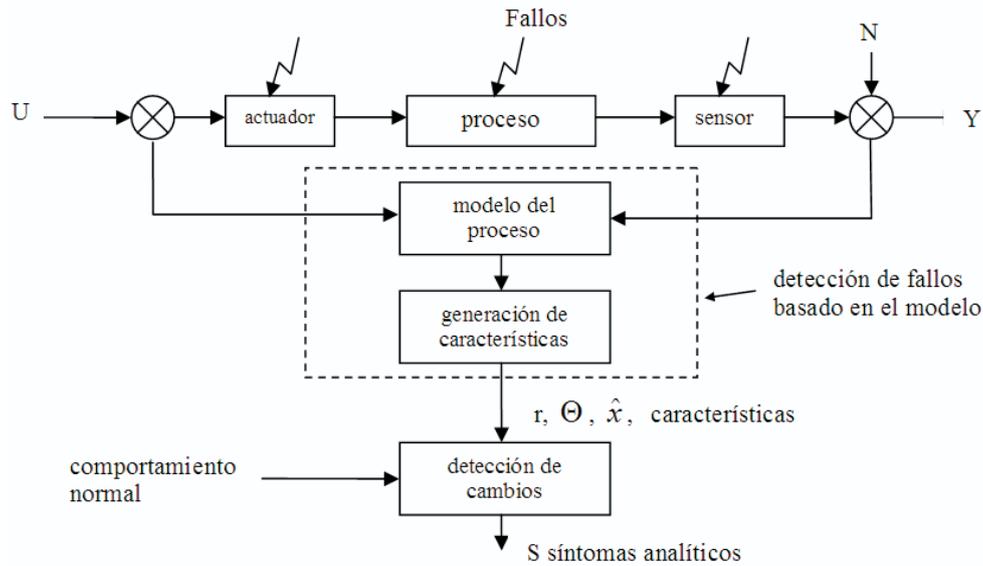


Figura 2.1: Detección de fallos basado en el modelo del proceso, donde N son las perturbaciones

2.3. Modelado de fallos

Un fallo, es la desviación indebida de una variable característica en un sistema que puede causar el mal funcionamiento o avería del sistema. La dependencia de los fallos con respecto al tiempo, se puede distinguir como se muestra en la figura 2.2.

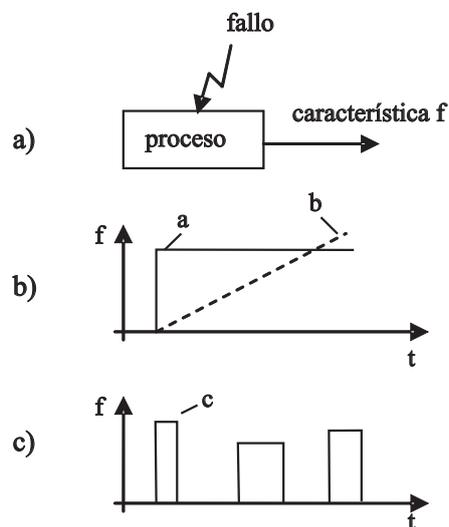


Figura 2.2: Dependencia de los fallos con respecto al tiempo: a) abrupta, b)incipiente, c) intermitente

- Fallo abrupto (fallo completo).
- Fallo incipiente (como tendencia).

- Fallo intermitente.

Con respecto al modelo del proceso, el fallo puede clasificarse completamente de acuerdo con la siguiente figura (ver figura 2.3).

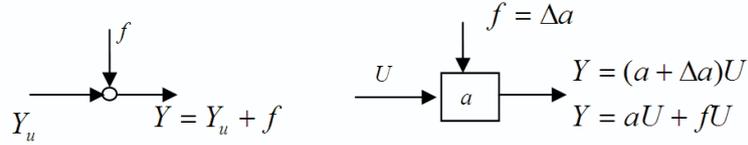


Figura 2.3: Modelos básicos de fallos: a) fallos aditivos, b) fallos multiplicativos

Los fallos aditivos influyen en la variable “Y”, por la adición del fallo f y el fallo multiplicativo por el producto de la variable U con f . Los fallos aditivos pueden resultar del *offset* (desviación) de los sensores, considerando que los fallos multiplicativos son los cambios internos del proceso.

Ahora se consideran los parámetros globales del proceso, los cuales operan en lazo abierto. El comportamiento estático (estado estable) puede expresarse frecuentemente por características no lineales.

$$\begin{aligned}
 Y &= \beta_0 + \beta_1 U + \beta_2 U^2 + \dots + \beta_q U^q \\
 Y &= \psi_s^T \Theta_s \\
 \Theta_s &= [\beta_0 \beta_1 \dots \beta_q] \\
 \psi_s^T &= [1 U U^2 \dots U^q]
 \end{aligned}$$

Los cambios de los parámetros β_i son fallos multiplicativos. Los fallos de señales de entrada f_U y los fallos de señales de salida f_Y son fallos aditivos. Compare la figura 2.4.

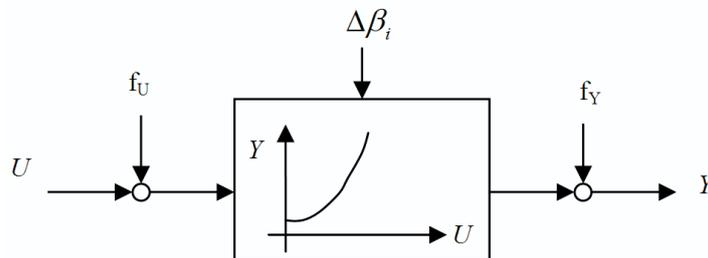


Figura 2.4: Modelo de proceso estático no lineal

Si se consideran desviaciones de pequeña señal alrededor de un punto de operación ($Y(x)/U(x)$), el comportamiento de entrada/salida del proceso, puede ser descrito frecuentemente por ecuaciones diferenciales lineales ordinarias.

2.4. Detección de fallos por estimación de parámetros

A veces no se tienen determinados en su totalidad los parámetros de un proceso. Es entonces, cuando a través de métodos de estimación de parámetros pueden ser obtenidos, donde las señales de entrada y salida nos pueden servir para este propósito, si se conoce la estructura básica del modelo.

2.4.1. Métodos de ecuaciones de error

El modelo del proceso es escrito en forma vectorial con los parámetros y el vector de datos de tiempo continuo. Para la estimación de parámetros se introduce la ecuación de error, ver figura 2.5.

$$e(t) = y(t) - \psi^T(t)\Theta$$

ó

$$e(s) = \hat{A}(\hat{\Theta}, s)y(s) - \hat{B}(\hat{\Theta}, s)u(s)$$

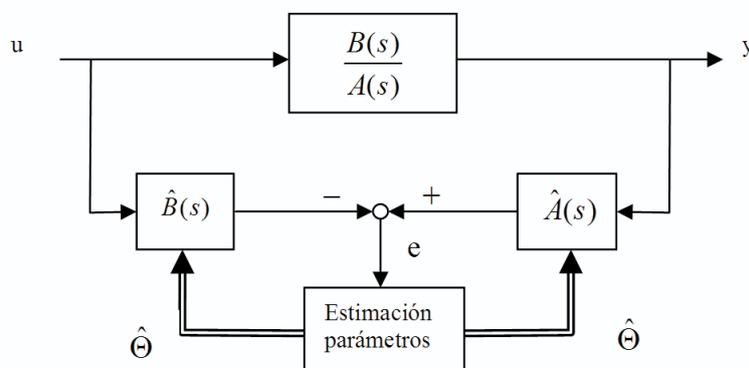


Figura 2.5: Modelo para la estimación de parámetros por ecuaciones de error

Después de muestrear con el tiempo discreto $k = t/T_0 = 0, 1, 2, \dots$ (sólo para la estimación de parámetros) Con T_0 como el (pequeño) tiempo de muestreo, minimización de la suma de mínimos cuadrados,

$$V = \sum_{k=1}^N e^2(k) = e^T e$$

$$dV/d\Theta = 0$$

lleva a la estimación de los mínimos cuadrados (LS)

$$\hat{\Theta}(N) = [\Psi^T \Psi]^{-1} \Psi^T y$$

o en su forma recursiva (RLS)

$$\hat{\Theta}(k+1) = \hat{\Theta}(k) + \gamma(k)y[k+1] - \psi^T(k+1)\hat{\Theta}(k)$$

$$\gamma(k) = \frac{1}{\psi^T(k+1)P(k)\psi(k+1) + 1} P(k)\psi(k+1)$$

$$P(k+1) = [I - \gamma(k)\psi^T(k+1)]P(k)$$

2.4.2. Métodos de error de salida

En lugar de ecuaciones de error se puede usar el error de salida (OE),

$$r(t) = y(t) - y_M(\hat{\Theta}, t)$$

donde

$$y_M(\hat{\Theta}, s) = \frac{\hat{B}(\hat{\Theta}, s)}{\hat{A}(\hat{\Theta}, s)} u(s)$$

es el modelo de salida, ver figura 2.6. Pero entonces es posible efectuar cálculos indirectos de la estimación de parámetros Θ , porque en los parámetros, $r(t)$ es no lineal. Por ende la función V iterativa, es minimizada por métodos de optimización numérica.

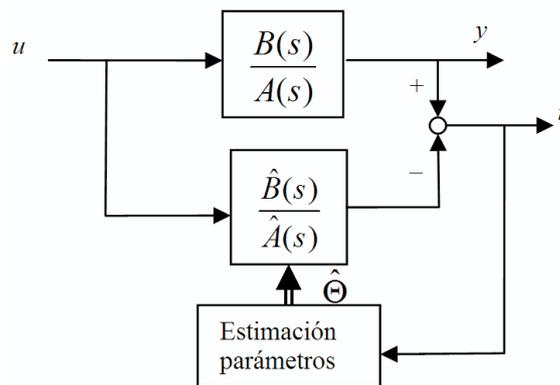


Figura 2.6: Modelo para la estimación de parámetros por el error de salida

El esfuerzo de cómputo es más grande y por lo general no se pueden obtener aplicaciones en tiempo real. De cualquier manera se generan parámetros con relativa precisión. Si un fallo dentro del proceso cambia uno o varios parámetros debido a un diferencial $\Delta\Theta_j$, la señal de salida cambia con las pequeñas desviaciones de acuerdo a

$$\Delta y(t) = \psi^T(t)\Delta\Theta(t) + \Delta\psi^T(t)\Theta(t) + \Delta\psi^T(t)\Delta\Theta(t)$$

y el estimador de parámetros indica un cambio en $\Delta\Theta$. Por lo general, el parámetro del proceso Θ depende de coeficientes físicos del proceso p (como factor de amortiguamiento, rigidez y resistencia)

$$\Theta = f(p)$$

vía una ecuación algebraica no lineal. Sí la inversión de esta relación

$$p = f^{-1}(\Theta)$$

existe, los cambios de los coeficientes del proceso Δp_i se pueden calcular. Estos cambios en los coeficientes del proceso en muchos casos están directamente relacionados con los fallos. De aquí que, el conocimiento de Δp_i facilita el diagnóstico de fallos, lo que quiere decir que no sólo es necesario para la detección de fallos, sino para el diagnóstico también.

2.5. Detección de fallos por medio de observadores

Una diferencia entre el método de estimación basado en observadores, y el método basado en la estimación de parámetros (ecuaciones de error), es que este último, puede hacer la estimación de parámetros en línea (Frank 1997).

La teoría existente para estimación paramétrica se puede aplicar directamente al diagnóstico de fallos sin mayores modificaciones. Los observadores para diagnóstico son diferentes a los bien conocidos observadores para control (Ogata 1998). Las diferencias esenciales son:

- Los observadores de diagnóstico son primeramente observadores de salida, mas bien que observadores del espacio estado, como se requiere para propósitos de control.
- Los observadores de diagnóstico operan en bucle abierto, mientras los de control en bucle cerrado.

Tanto en los métodos de estimación paramétrica como los basados en observadores, existirá siempre el riesgo de contaminación del fallo por el control en bucle cerrado (Isermann 1997a). El esfuerzo de control en bucle cerrado procura mantener las variables de salida en ciertos límites. Este esfuerzo de control, cuando el fallo es incipiente, lo suele ocultar y sólo es detectable, cuando el controlados no puede evitar que los residuos crezcan.

2.5.1. Principios básicos para la construcción del generador residual en observadores

En la generación de residuos capaces de reflejar los fallos, son usadas las salidas del proceso y la estimación de error r como residuo, ver figura 2.7 donde f es el vector de fallos a detectar, representado por funciones de tiempo desconocidas, d es el vector de entradas desconocidas (perturbaciones, ruidos, errores de modelado) para las cuales el sistema de detección debe ser inmune. Consideremos un proceso *lineal* descrito por

$$y_L(s) = G_u(s)u_L(s) + \Delta y_L(s) + G_f(s)f_L(s)$$

donde G_u y G_f son matrices de transferencia del vector de entrada $u \in R^p$, vector de fallos $f \in R^q$ para el vector de salida $y \in R^m$, y el subíndice L denota la transformada de Laplace en función del tiempo. $\Delta Y_L(s)$ Es un vector desconocido que representa los perturbaciones imprevistos y la incertidumbre del modelo.

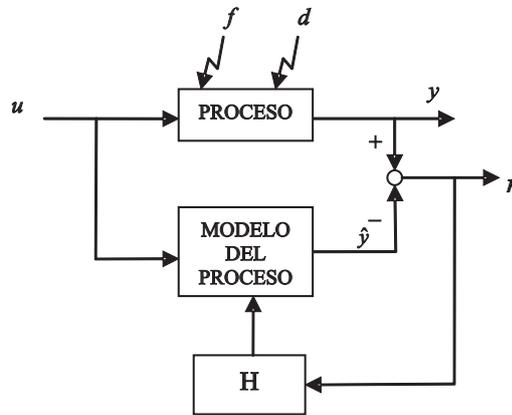


Figura 2.7: Observador de orden completo para la generación de residuos

El generador residual se construye a través de una técnica de factorización, la cual es aplicable a todos los procesos (estables e inestables), entonces el generador residual se expresa en la siguiente forma:

$$r_L = \hat{M}_u(s)[y_L(s) - G_u(s)u_L(s)]$$

$$r_L = \hat{M}_u(s)y_L - \hat{N}_u(s)u_L(s)$$

donde $\hat{M}_u(s)$ y $\hat{N}_u(s)$ son factores coprimos de $G_u(s)$ que satisfacen $\hat{M}_u^{-1}(s)\hat{N}_u(s) = G_u(s)$. Denotando la realización del espacio estado de la matriz de transferencia nominal $G_u(s)$ con

$$\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t) + Du(t)$$

se puede calcular $\hat{M}_u(s)$ y $\hat{N}_u(s)$ como sigue:

$$\hat{M}_u(s) = I - C(sI - A + LC)^{-1}L$$

$$\hat{N}_u(s) = D + C(sI - A + LC)^{-1}(B - LD)$$

donde L asegura la estabilidad de la matriz $A - LC$. Según los autores de este tema (Frank 1997) r_L se puede expresar así

$$r_L(s) = Q(s)[\hat{M}_u(s)y_L(s) - \hat{N}_u(s)u_L(s)]$$

donde $Q(s) \in RH_\infty$ denota una matriz de parametrización que aun se encuentra libre, apta a ser seleccionada. El significado físico de $Q(s)$ es el de un filtro posterior como se observa en la figura 2.8, el cual provee grados adicionales de libertad para el aislamiento de fallos y para la generación robusta de residuos.

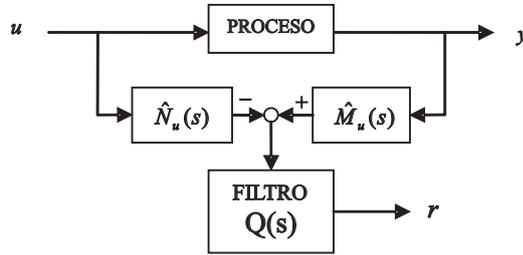


Figura 2.8: Observador de orden completo para la generación de residuos

Se ha demostrado que

$$\hat{M}_u(s)y_L(s) - \hat{N}_u(s)u_L(s) = y_L(s) - \hat{y}_L(s)$$

donde $\hat{y}_L(s)$, es una estimación de salida enviada por un observador

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + L[y(t) - \hat{y}(t)]$$

$$f(t) = C\hat{x}(t) + D(t)$$

de lo anterior podemos concluir que

$$r_L(s) = Q(s)(y_L(s) - \hat{y}_L(s))$$

2.6. Detección de fallos con ecuaciones de paridad

Este método tiene la propiedad de poder operar en línea. La configuración de estos modelos puede ser en serie o en paralelo, lo que nos lleva a obtener residuos con diferente sensibilidad. Se comparan los residuos de estos dos modelos y se escoge el que tenga mayor sensibilidad para la detección de los fallos.

2.6.1. Ecuaciones de paridad para procesos de una entrada y una salida (SISO)

Un método directo basado en el modelo para la detección de fallos es obtener un modelo G_M (equivalente a G_P) y operarlo en paralelo con G_P y de esta manera formar el error de salida, como se observa en la figura 2.9.

$$r'(s) = [G_p(s) - G_M(s)]u(s)$$

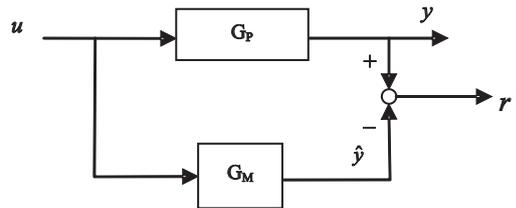


Figura 2.9: Ecuaciones de paridad, método de error de salida

Como para los observadores los parámetros del proceso deben ser conocidos a priori, si $G_P(s) = G_M(s)$, entonces el error de salida será un fallo aditivo de entrada y salida, como se vio anteriormente y tendremos

$$r'(s) = G_p(s)f_u(s) + f_y(s)$$

Otra posibilidad es generar un error polinomial como se muestra en la figura 2.10.

$$r(s) = A_M(s)y(s) - B_M(s)u(s)$$

$$r(s) = B_p(s)f_u(s) + A_p(s)f_y(s)$$

En ambos casos se obtienen diferentes respuestas de tiempo para un fallo aditivo de entrada y salida. $r'(s)$ es equivalente al error de salida de la estimación paramétrica vista anteriormente y $r(s)$ a la ecuación de error también vista en ese mismo punto, ambas ecuaciones son llamadas ecuaciones de paridad. Para procesos SISO solamente un residuo puede ser generado y por lo tanto es difícil distinguir entre diferentes errores.

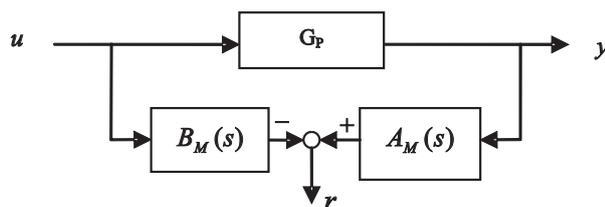


Figura 2.10: Ecuaciones de paridad, método de ecuación de error

2.6.2. Ecuaciones de paridad para procesos de múltiples entradas y múltiples salidas (MIMO)

Con este tipo de procesos es apropiado el uso de modelos de ecuaciones de estado, como se muestra en (Gertler 1995) para modelos en tiempo discreto. Para modelos en tiempo continuo

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}$$

la ecuación de salida anterior se inserta en la ecuación de estado, de lo cual resultan tiempos diferenciados p en el sistema de ecuaciones

$$\begin{bmatrix} y(t) \\ \dot{y}(t) \\ \ddot{y}(t) \\ \vdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 & 0 & \dots \\ CB & 0 & 0 \\ CAB & CB & 0 \\ \vdots \end{bmatrix} \begin{bmatrix} u(t) \\ \dot{u}(t) \\ \ddot{u}(t) \\ \vdots \end{bmatrix}$$

$$Y_F(t) = Tx(t) + QU_F(t)$$

Con el propósito de eliminar los estados no medibles $x(t)$, se multiplica por W de manera que

$$WT = 0$$

de aquí obtenemos los residuos

$$r(t) = WY_F(t) - WQU_F(t)$$

y esto se esquematiza de la siguiente forma

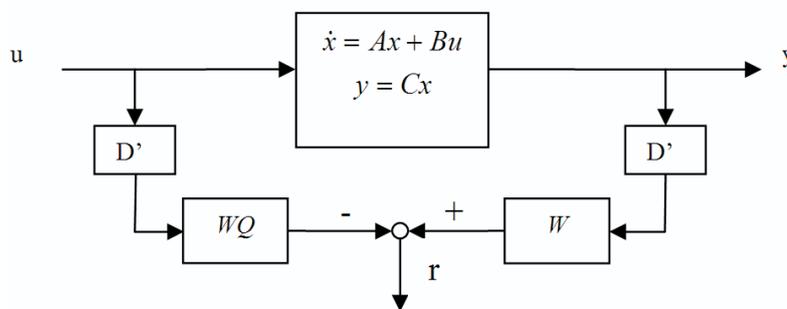


Figura 2.11: Método de ecuaciones de paridad para modelos en espacio estado MIMO; D': filtro de diferenciación

Las derivadas en $Y_F(t)$ y $U_F(t)$ se pueden obtener por medio de filtros digitales de variables de estado con un orden $n \leq 3$.

2.7. Detección de fallos mediante el modelado de señales

Señales medidas de salida $y(t)$, pueden mostrar desviaciones que son, ya sea de naturaleza armónica o estocástica o una combinación de las dos. Si estas señales están relacionadas con los fallos del proceso, de los elementos de accionamiento o de los sensores, el análisis de estas señales es una nueva fuente de información que se puede utilizar en el proceso de detección de fallos.

Por lo general, se usan sensores de desplazamiento, velocidad o aceleración para detectar las vibraciones en máquinas, el fallo de alineamiento y fallos de rodamientos en maquinaria rotativa (Hui 2003), el golpeteo en motores Diesel o las vibraciones en máquinas de corte de metales. Pero también las señales de otros sensores que miden variables como la corriente eléctrica, la posición, la velocidad, la fuerza, el flujo y la presión pueden mostrar variaciones en una gama amplia de frecuencias diferentes a las nominales y que corresponden a las respuestas dinámicas del sistema.

La extracción de las características relevantes de la señal y su relación con los fallos pueden en algunos casos restringirse a las amplitudes $y_0(\omega)$, frecuencias ω , o a las densidades de las amplitudes $|y(i\omega)|$, dentro de cierto ancho de banda $\omega_{min} \leq \omega \leq \omega_{max}$ de la señal, para esto, se utilizan filtros pasa bajas o “*wavelet transforms*” (Hui 2004). Otra posibilidad es la determinación de las funciones de correlación o la transformada de Fourier, especialmente la FFT (Rai 2006; Jin-Fu 2002) o la densidad espectral (Spurbeck 2006).

Las funciones de correlación (Avazpour 2006) y las densidades espectrales (Bonte 2007) son muy útiles para separar los componentes periódicos y estocásticos de la señal. Si se conocen las frecuencias se puede establecer la correlación cruzada de los armónicos principales y de grado superior. Sin embargo, la función de correlación o el espectro de frecuencias son representaciones no paramétricas de la señal y su uso no es muy práctico en la detección Automática de fallos si las respectivas frecuencias y su relación con los fallos se desconocen.

Por consiguiente, se recomienda el modelado paramétrico de la señal, el cual permite estimar directamente las frecuencias principales y sus amplitudes, además, que es esencialmente sensible a pequeños cambios de frecuencia. Esto se logra modelando las señales como una superposición de señales en forma de sinusoides amortiguadas en tiempo discreto, y habitualmente conocido, como modelos ARMA (Pan 2005) (*Auto-Regresive Moving Average*). También se pueden emplear otras técnicas de tratamiento elaborado de las señales como el método de la envolvente, el método de convolución y las “*Wavelet transforms*” (Smith 2007; Zarei 2006).

Conclusiones del capítulo

En este capítulo se estudio la detección de fallos por métodos analíticos como son: el modelado de sistemas dinámicos a través de sus ecuaciones diferenciales, representadas en diferentes formas como son: las funciones de transferencia en el dominio de Laplace, el modelado en el espacio estado. Este tipo de representación tiene como ventaja su exactitud, cualidad que es muy útil en la detección de fallos, sobre todo cuando los fallos son enmascarados por los sistemas de control automático que poseen los procesos industriales. Otra ventaja, es que algunas de las técnicas vistas pueden operar en línea, característica muy deseada en los sistemas de diagnóstico en tiempo real. Sin embargo, estas técnicas tienen como desventaja, que es difícil la adquisición de los modelos dinámicos de ecuaciones diferenciales y una vez que se han adquirido, difícil de garantizar que estos modelos permanecerán sin cambios; que en la práctica es lo más usual, pues los procesos industriales son susceptibles a desgaste y son afectados por variables externas de todo tipo.

Capítulo 3

Lógica y sistemas borrosos: generalidades.

3.1. Introducción.

Desde los tiempos de la antigua Grecia se comenzó a utilizar la lógica binaria como herramienta para formalizar el *razonamiento*. Durante el pasado siglo los trabajos de De Morgan y Boole la formalizaron matemáticamente, intentando usarla como base de las matemáticas. La aparición de los computadores (internamente binarios) ha llevado a la utilización de la lógica binaria para la resolución de problemas de *inteligencia artificial*, que culmina en la aparición de los sistemas expertos, aplicados a gran variedad de tareas industriales, entre ellas el control, supervisión y diagnóstico de fallos.

Un importante aspecto presente en el modelado y control de procesos industriales es la incertidumbre parcial y la aparición de magnitudes *cualitativas* no claramente formalizables mediante el modelo binario de la lógica. El modelado que distintas escuelas de la inteligencia artificial hacen de ellas ha dado origen a aplicaciones de ingeniería entre las que destaca, el *control inteligente*. En particular, el razonamiento ambiguo modelado mediante la lógica borrosa ha sido aplicado a los reguladores borrosos, cuya estructura genérica será descrita en la sección 3.7 de este capítulo. Otros modelos de la incertidumbre de naturaleza estadística o de la incertidumbre en forma de cotas de variación paramétrica han sido aplicados a control (control estadístico (Isermann 1989), robusto (Freeman 1996), etc.) pero requieren de un modelo matemático relativamente complejo. El enfoque basado en la lógica borrosa ofrece, sin embargo, posibilidades de descripción cualitativa de los procesos industriales para capturar su naturaleza inexacta y aproximada. Otros enfoques del tratamiento de la incertidumbre son los de razonamiento evidencial (Dempster-Shafer) (Shafer 1976), inferencia bayesiana (Shafer 1990), factores de certeza (Buchanan 1984), los *rough sets* (Pawlak 1982; Czogala 1995) etc.

A continuación se introducen las definiciones básicas de la lógica borrosa utilizadas en el diseño de controladores borrosos. Las diferentes interpretaciones tanto de lo que

será definido como concepto borroso y de las conectivas para operar con ellos pueden dar lugar a confusión en el usuario poco experto que trate de comprender el proceso de razonamiento en una determinada aplicación.

La lógica borrosa se ocupa del *grado* en el que ocurren los sucesos y no en la *probabilidad* de que ocurran. Se aplica a conceptos con *gradación del significado*, en particular a variables de control de procesos sobre las que se establece una serie de valores lingüísticos cualitativos. Introducida modernamente por Lofti A. Zadeh (Zadeh 1965), el problema de la inexactitud de la bivalencia para describir el mundo real ha preocupado a filósofos como Stuart Mill (1843) en el siglo pasado y matemáticos como Gödel y Lukasiewicz en los años 20 (Lukasiewicz 1920) y Black (Black 1937) en los años 30. De hecho, la lógica de Lukasiewicz provee de muy interesantes ideas a la teoría de conjuntos borrosos (Giles 1997).

La lógica borrosa es, hoy en día, una herramienta de razonamiento aproximado ampliamente utilizada y referenciada en diversos campos de aplicación. Se remite al lector, para información detallada, a abundantes referencias bibliográficas como (Wang 1994; Pedrycz 1993; Kosko 1992; Hellendoorn 1993; Albertos 1992; Jager 1995) etc.

3.2. Lógica binaria.

La lógica binaria es una disciplina que presenta un conjunto de métodos, válidos para llegar a conclusiones correctas a partir de unos conocimientos iniciales basándose en determinados axiomas.

Si se divide el lenguaje humano en imperativo (acciones), declarativo (conocimientos) e interrogativo, la lógica clásica simboliza el lenguaje humano de tipo declarativo, descriptor del conocimiento, mediante dos enfoques:

- **Cálculo proposicional.** El elemento básico de la formulación es la frase declarativa simple o *proposición* combinada por operadores lógicos o *conectivas* (negación, conjunción, disyunción e implicación, entre otras).
- **Cálculo de predicados.** Opera con *términos a los que se aplican predicados*. En esta formulación, aparecen los *cuantificadores de existencia* (\exists) y *universalidad* (\forall).

El cálculo proposicional puede abordarse mediante el enfoque denominado *teoría de la demostración*, basado en *axiomas* del sistema y *reglas de demostración* o bien mediante el enfoque denominado *teoría semántica*, basado en *tablas de verdad*. Las proposiciones de valor lógico conocido se denominan *premisas*.

El cálculo de predicados asume una estructura de las proposiciones de tipo *sujeto-predicado*, junto con cuantificadores de existencia y universalidad, tratando con proposiciones de tipo “Existe sujeto x para el que el predicado P es aplicable” o “Para

todo x , P es aplicable”, como por ejemplo, “Algunos hombres son rubios”, “Todos los mamíferos son animales”.

La *lógica modal* es una extensión a los anteriores sistemas lógicos que formaliza los conceptos de posibilidad y necesidad. La necesidad y la posibilidad plantean formalmente la diferencia entre concluir “hay una sobreintensidad” y “puede haber una sobreintensidad”. Los conceptos en cuestión y su generalización a sistemas lógicos no binarios son de utilidad en aplicaciones de la inteligencia artificial como la detección y diagnóstico de fallos (Cayrac 1996). Para tratar el problema de la ambigüedad e incertidumbre, diversos autores (Rescher 1969; Walker 1994) han propuesto sistemas lógicos *multivaluados* en los que los posibles valores lógicos son más de dos. El enfoque más generalizado es el de la lógica borrosa, que se describirá a continuación.

3.3. Sistemas basados en reglas.

Aunque anteriormente se ha dado a entender que toda expresión lógica binaria de la cual se derivan otras se considera como premisa, en el contexto particular de las aplicaciones de la lógica (cálculo proposicional) en sistemas expertos (Gupta 1988; Buchanan 1984; Cuenca 1985), el conjunto de premisas se dividirá en:

- **Premisas simples.** Las proposiciones atómicas de valor lógico conocido,
- **Reglas.** Aquellas expresiones (proposiciones) complejas en las cuales intervengan varias variables de las cuales no puedan deducirse trivialmente premisas simples¹.

En general, la forma usual de representar una regla es mediante una expresión equivalente donde intervenga un operador de implicación o doble implicación (aunque esta equivalencia plantea problemas en generalizaciones multivaluadas). El conocimiento, en ese caso, se agrupa en *bases de reglas* de tipo:

R1: “Si p , entonces q ”, R2: “Si y sólo si r , entonces s ”,...

3.3.1. Verificación lógica. Anomalías.

Dado un sistema lógico, con el conocimiento expresado en forma de premisas y reglas según descrito anteriormente, en (Preece 1994) se enumeran una serie de posibles anomalías, a saber:

- **Regla no satisfacible.** Una regla R es *no satisfacible* si no puede deducirse su antecedente de ninguna premisa admisible.

¹Esto es, aquéllas cuya tabla de verdad no sea falsa excepto en una única combinación de valores lógicos de las proposiciones atómicas que la integran: $A \wedge B$ será considerado como dos premisas simples, $A \vee \neg B$ sería una *regla*.

- **Regla inútil.** Se dice que una regla es inútil si su consecuente no es ni una conclusión final ni antecedente de otra regla.
- **Regla contenida en otra.** Si existe una regla más general de la que la regla en cuestión es caso particular (p. ej, $A \vee B \rightarrow C, A \rightarrow C$).
- **Regla redundante.** Generaliza la anterior anomalía a conjuntos de reglas: una regla es redundante si para toda premisa admisible, la regla no es esencial para el cálculo de ninguna conclusión, esto es, si su eliminación de la base de reglas no produce alteración en las conclusiones del sistema sean cuales sean las premisas.
- **Par de reglas inconsistentes.** Si siendo las dos aplicables, puede derivarse una restricción semántica².
- **Conjunto de reglas inconsistente.** Generalización de la anterior anomalía a la deducción de restricciones semánticas tras encadenamiento de varias reglas.
- **Conjunto de reglas circular.** Si el antecedente de una regla del conjunto no puede ser derivado salvo que se suponga su consecuente como premisa.
- **Premisa inútil.** Si pueden llegarse a las mismas conclusiones sin suponer la existencia de dicha premisa.
- **Base de reglas incompleta.** Si no puede determinarse el valor lógico de una determinada conclusión.

Existe la necesidad de incluir en la validación y verificación lógica de sistemas expertos basados en lógica binaria (Chang 1990; Nguyen 1987; Suwa 1982; Preece 1994), aspectos particulares relacionados con la interpretación de la implicación en lógica borrosa.

3.4. Cálculo proposicional borroso.

La lógica borrosa, desde un punto de vista formal, es la extensión de la teoría semántica del cálculo proposicional, cuando la veracidad de las proposiciones puede tomar valores en el intervalo $B = [0, 1]$. Estas proposiciones se denominan *proposiciones borrosas*, y al valor lógico que toman se le denomina *grado de verdad* o *veracidad*. Por generalización de la lógica binaria, se define como *cierto* a la veracidad 1, y *falso* a la veracidad 0.

La extensión de las conectivas debe ser compatible con las tablas de verdad de las conectivas binarias. No obstante, existen diferentes interpretaciones de las mismas que dan lugar a varios modelos de inferencia.

Del mismo modo que en lógica binaria coexisten los enfoques de teoría de la demostración y de teoría semántica, en lógica borrosa también se pueden establecer sistemas

²esto es, un par de valores lógicos de proposiciones incluidos en una *lista de contradicciones*, como por ejemplo {" x es hombre", " x tiene 6 patas"}. Esto generaliza la definición de contradicción *formal* ($A \wedge \neg A$) a contradicción *semántica* para lógica binaria.

axiomáticos (Lukasiewicz 1920; Frank 1996) o sistemas basados en generalizar la tabla de verdad de las conectivas. Este segundo enfoque es, con diferencia, el más utilizado.

Las conectivas conjuntiva (\wedge) y disyuntiva (\vee) son representadas por normas y conormas triangulares respectivamente. La negación (\neg) se define usualmente como el complemento a uno.

3.5. Conjuntos borrosos.

Del mismo modo que en la teoría semántica binaria las proposiciones con predicados atributivos (“ x es A ”) pueden interpretarse como aserción de la pertenencia a un conjunto, la lógica borrosa también da lugar a una interpretación basada en *conjuntos borrosos*.

Los conjuntos borrosos se aplican a elementos x de un universo de discurso U para describir conceptos *graduables* (esto es, aquéllos a los que los elementos pueden pertenecer en mayor o menor medida). Como ejemplo, en una aplicación de climatización, el conjunto “temperatura ambiente agradable” vendría definido por una función sobre un universo de temperaturas representada en la figura 3.1. Consideraciones adicionales sobre el significado de los conjuntos borrosos se realizarán en la sección 3.6.

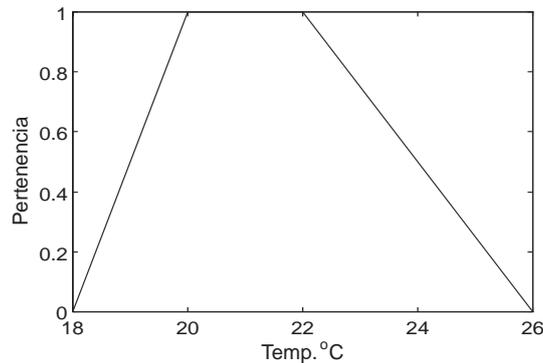


Figura 3.1: Temperatura agradable

Definición 3.1 *Un conjunto borroso A se define mediante una función de pertenencia:*

$$\mu_A : U \rightarrow [0, 1] \tag{3.1}$$

de manera que $\mu_A(x)$ denota el grado de pertenencia de x a A .

Definición 3.2 *El conjunto vacío es aquél en que $\forall x \in U \mu_\emptyset(x) = 0$. El conjunto universo es el definido por $\forall x \in U \mu_U(x) = 1$.*

Definición 3.3 (Borrosificación) *Dado un conjunto borroso A , definido por una función de pertenencia $\mu_A(\cdot)$, se define como valor lógico de la proposición atómica “ x es A ” al valor $\mu_A(x)$.*

Una notación alternativa equivalente definiendo conjuntos borrosos es:

$$A = \sum_{i=1}^n \mu_A(x_i)/x_i = \mu_A(x_1)/x_1 + \cdots + \mu_A(x_n)/x_n \quad (3.2)$$

para conjuntos borrosos sobre un universo numerable y :

$$A = \int_U \mu_A(x)/x \quad (3.3)$$

para universos de discurso no numerables. De todas formas, esa notación puede ser confusa en cuanto a que determinados conceptos sobre los conjuntos borrosos necesitan de verdaderas sumas o integrales para calcular su cardinalidad o el centro de masas en las fórmulas de deborrosificación (en inglés, *defuzzification*) , por lo que se preferirá en adelante la notación funcional descrita según (3.1).

Aunque, en teoría, no existe ninguna restricción a la forma particular de los conjuntos borrosos, las funciones de pertenencia, en la práctica, suelen ser o bien monótonas o bien funciones crecientes hasta determinado punto, constantes en un intervalo, y decrecientes a partir del extremo superior del intervalo. Las funciones de pertenencia borrosas con múltiples máximos relativos (multimodales) no tienen interpretación lingüística clara excepto en términos de unión de varias funciones más simples.

3.6. Significado de la borrosidad.

El concepto de *conjunto borroso* plantea, en principio, problemas de significado y asignación intuitiva de pertenencias, ya que los conceptos “ambiguos” son transformados en una función matemática unívoca, en absoluto ambigua: del mismo modo que la teoría de la probabilidad, un conjunto borroso describe la ambigüedad mediante números en el intervalo $[0,1]$, y existen choques y duda en la distinción entre ambos conceptos (Kosko 1992), llegando algunos autores a concluir la inutilidad de la lógica borrosa como sistema de representación de la incertidumbre (Lindley 1987), considerándolo un caso particular de la interpretación bayesiana de la probabilidad *a priori*.

Aclarando esta situación, el significado intuitivo que en muchos casos se da a los conceptos borrosos que un experto define y utiliza cae dentro de una serie de casos que se detallan a continuación, en algunos casos muy relacionados con la probabilidad, lo cual justifica el empleo, dentro de esa interpretación, de fórmulas de estadística bayesiana.

Como ejemplo, la frase “la temperatura T es *Alta* 0.7” se puede interpretar según:

Enfoque probabilístico. El 70 % de los expertos diría que esa temperatura es alta, el 30 % que no (Hisdal 1995). Nótese que, en este caso, muchos expertos dirían “medio-alta / medio-baja” y plantearían serias objeciones al ingeniero del conocimiento si se les obligase a definirse por “alta” o “baja”.

Enfoque de conjuntos aleatorios. El 70 % de los expertos da como intervalo de temperaturas que consideran *altas* un intervalo conteniendo a T .

Similaridad. La distancia “conceptual” a una temperatura alta “prototipo” es 0.3 (Zimmermann 1985; Sala 1995). La distancia conceptual igual a 1 tiene el significado de “todo lo contrario”.

Incertidumbre (*posibilidad, factibilidad*). Dada una afirmación lingüística pura “La temperatura T es alta” (sin cuantificación) por un experto *distinto* al que procesa la información, la función de pertenencia sería una distribución de las posibilidad o factibilidad de que T tenga un determinado valor. Esta interpretación es el resultado de una comunicación puramente lingüística ambigua (Zadeh 1978).

Utilidad o preferencia. 0.7 es la utilidad de afirmar que T es alta, en el sentido de una función de coste (implícita en la mente del experto) (Giles 1988). En este contexto, el significado y definición de las conectivas admite gran flexibilidad, sin necesidad de ceñirse a la interpretación algebraica estricta de la teoría semántica.

Ordinalidad. T es una temperatura más alta que unas temperaturas y menos alta que otras. En una ordenación mediante una escala entre 0 y 1 más o menos objetiva/subjetiva, le corresponde el lugar 0.7. Así, en el espacio de pertenencias sólo tendrían sentido los operadores de comparación ($<$, $>$, $=$) pero no las agregaciones (unión, intersección) o ponderaciones entre los valores. Del mismo modo, las conectivas deberán generalizar la lógica binaria y cumplir restricciones de monotonía. De hecho, el intervalo $[0,1]$ puede sustituirse bajo esa interpretación con pertenencias definidas sobre un conjunto finito, linealmente ordenado, de niveles cualitativos.

Las dos visiones probabilísticas y la posibilística enunciadas consideran el origen de la borrosidad en los errores de medida, la información incompleta y las contradicciones interpersonales.

El entender las funciones de pertenencia según unas u otras interpretaciones da lugar a diferentes formas de manipularlas, y a determinar qué operaciones sobre ellas tienen significado, lo cual es, en cierto modo, un inconveniente importante de cara a la validación y generalización de la lógica como una herramienta *independiente del significado de las proposiciones* que demuestra, como lo es en lógica binaria.

Los anteriores enfoques también pueden estar originados al considerar que la raíz de la “borrosidad” en los conceptos empleados por los sistemas inteligentes artificiales pueda ser:

- objetiva vs. subjetiva
- individual vs. grupal

Así, la ordinalidad sería el resultado de una interpretación subjetiva, individual, de la gradación de conceptos; la similaridad vendría de una interpretación individual objetiva. Los enfoques probabilísticos responden a interpretaciones subjetivas grupales.

El enfoque inicial del creador de la lógica borrosa (L. A. Zadeh) era que la asignación de pertenencias responde a un concepto subjetivo, dependiente del contexto (Zadeh 1965).

Un estudio en profundidad de las cuestiones aquí abordadas puede encontrarse en (Bilgic 1997) y (Dubois 1997).

De hecho, dada la diversidad de interpretaciones, y la dificultad de asignar en los conjuntos borrosos los valores de pertenencia³, cabe preguntarse “*cuánta de la parafernalia de la lógica borrosa es necesaria para expresar los conceptos, y si acaso solamente una descripción basada en intervalos o rangos sería suficiente*” (Mamdani 1984).

Granulación.

La unión de la lógica borrosa y los conceptos borrosos tratan de reproducir un modelo conceptual de análisis mental basado en la *granulación*: el tratamiento de la información mediante la agrupación en *clases* de las cuales la etiqueta lingüística no es más que un símbolo.

Así, el concepto clave de granulación en el análisis de la información engloba tanto a los sistemas borrosos como a otros sistemas de tipo *modelo local*. Los gránulos pueden tener límites claramente definidos o bien presentar superposición (borrosos). La mente humana utiliza granulación primero, interpolación después.

La granulación permite manejar la información con una capacidad finita de memoria expresada, por ejemplo, en una base de reglas. La creación de reglas y *excepciones* no son más que mecanismos para reducir el número de gránulos. No obstante, las reglas no tienen por qué ser los únicos modelos de expresar la granulación.

La granulación adecuada es un elemento de *reducción de complejidad* clave en el tratamiento inteligente de los datos.

Otros tipos de proposiciones.

Nótese que en un sistema basado en el conocimiento deberían poder considerarse varios tipos de proposiciones:

- **Binarias.** No admiten gradación (“La velocidad de giro es mayor de 2 rpm”).
- **Borrosas.** Admiten gradación (“La temperatura es *alta*”).

³En realidad, los expertos, muchas veces sólo saben expresar sin excesivas dudas los “prototipos” de un concepto ($\mu(x) = 1$) y los “prototipos” de no serlo ($\mu(x) = 0$). Casi cualquier curva monótona de 0 a 1 que “interpole” ambos prototipos sería válida para el experto, en muchos casos, ya que su imagen mental del concepto ambiguo es bastante “ambigua” (indefinida), valga la redundancia.

Por una parte, esto puede dar lugar a pensar en extensiones del concepto de conjunto borroso o por otra parte a reducir la complejidad de la representación a únicamente esos dos prototipos positivo y negativo, en un enfoque de tipo *rough sets* (Pawlak 1982).

- **Con incertidumbre por desconocimiento.** $3 < v < 5$, “La temperatura es *alta* o *muy alta*”.
- **Con incertidumbre por aleatoriedad.** La temperatura de entrada es normal de media 45°C y varianza 2.

Pueden estar combinados varios de los tipos de proposición, como desconocimiento de las funciones de densidad de probabilidad, o probabilidades de conjuntos difusos como en el caso de “La probabilidad de que la temperatura sea *alta* es 0.4”, o desconocimiento de veracidades binarias (“No existen suficientes datos para determinar si la anomalía es debida o no a un fallo en el componente B”). Asimismo, puede hablarse de una lógica modal borrosa en donde se incorporen los conceptos de posibilidad y necesidad, incluso que estos conceptos se cuantifiquen mediante un valor en el intervalo $[0, 1]$. Como se intuye, el tratamiento de algunos de los tipos complejos de proposiciones puede resultar muy problemático.

3.7. Sistemas basados en reglas borrosas.

Del mismo modo que en lógica binaria, un sistema de lógica borrosa basado en reglas deduce hechos (conclusiones) a partir de una serie de proposiciones. Estas proposiciones se dividen en premisas y reglas.

Hay que hacer notar que en lógica binaria se puede llegar a expresiones equivalentes de una misma regla con implicaciones, conjunciones o disyunciones. No obstante, en lógica borrosa la equivalencia ya no es trivial y única, de manera que el cómo el experto expresa una proposición puede influir en el resultado final.

En general, se supone que las reglas están dadas en forma de *implicación*. A una estructura del tipo:

SI A_1 , ENTONCES C_1

SI A_2 , ENTONCES C_2

...

SI A_n , ENTONCES C_n

donde A_1, \dots, A_n son proposiciones borrosas antecedentes y C_1, \dots, C_n son proposiciones borrosas consecuentes se le denominará base de reglas. Si los consecuentes de una base son antecedentes de otra, se dice que las bases de reglas están *concatenadas*.

En la práctica, los consecuentes suelen ser proposiciones *atómicas*, y los antecedentes suelen ser grupos de proposiciones atómicas conectados por conectivas conjuntivas o disyuntivas.

Asimismo, el tipo de proposición atómica más común es “ x es M ”, donde M es un conjunto borroso definido sobre el universo al que pertenece x .

3.7.1. Reguladores borrosos.

Una de las aplicaciones de los sistemas basados en reglas más extendida en la práctica es la utilización de los mismos en control de procesos (Albertos 1992; Langari 1992; Astrom 1986), bien como modelos del proceso a controlar, bien como reguladores directos (Sanchez 2006; Jianbin 2006; Guang 2004). También existen sistemas expertos de ayuda al diseño de reguladores clásicos (Larsson 1990). Los sistemas de lógica borrosa son generalmente utilizados en este campo para aproximar funciones estáticas no lineales. La dinámica es introducida en pre- y post-procesado mediante líneas de retardo, derivadores o integradores (normalmente con aproximaciones numéricas discretas de derivadas e integrales a un determinado período de muestreo).

Cuando los sistemas borrosos son usados como aproximadores funcionales $f : U \rightarrow Y$, las bases de reglas suelen agrupar el conocimiento sobre relaciones entre una serie de variables, en el sentido de que los antecedentes constituyen proposiciones borrosas que incluyen conjuntos borrosos sobre las variables de U (por ejemplo, error, derivada del error) y los consecuentes suelen ser conjuntos borrosos cuyo dominio es Y (por ejemplo, acción de control), y se suelen referir a una *única* variable. Las ventajas inherentes al enfoque basado en reglas para la aproximación funcional son la localidad (granularidad) y la “conmutación” suave (interpolación) entre las distintas situaciones particulares que plantea cada regla (Pedrycz 1997).

La integración de varias bases de reglas (concatenadas o no) produce controladores para sistemas MIMO. Los sistemas borrosos con conclusiones no triviales⁴ sobre varias variables no son comunes en la práctica.

3.7.2. Sistemas de diagnóstico borrosos

En estos, los síntomas son representados como conjuntos borrosos (ver figura 3.1) con significados lingüísticos como “pequeño”, “mediano”, “grande”. El razonamiento lógico borroso es aplicado siguiendo el principio de los sistemas de reglas SI-ENTONCES borrosas, con los siguientes pasos: borrosificación, inferencia, acumulación y deborrosificación (Isermann 1997a). Al encadenar las reglas borrosas, es posible considerar varios niveles de reglas. El resultado es entonces la posibilidad de la medición de un fallo, considerando también su tamaño. Para más detalles en las técnicas de diagnóstico borroso ver (Isermann 1994; de Miguel 2005). Aplicaciones resientes en esta campo se pueden encontrar en (Jian-Da 2006; Lo 2007; Pawar 2003).

Debido a lo laborioso que es la tarea de adquisición de la base de reglas, se han desarrollado los “sistemas borrosos adaptables” (en inglés, *adaptive fuzzy systems*). En estos se usan las similitudes entre los sistemas borrosos y las redes neuronales. Ayoubi

⁴obviamente, “ x es $A \rightarrow (y$ es $c \wedge z$ es $D)$ ” puede descomponerse trivialmente (proposición) en dos reglas ordinarias “ x es $A \rightarrow y$ es c ” y “ x es $A \rightarrow z$ es D ”, pero no así la regla “ x es $A \rightarrow (y$ es $c \vee z$ es $D)$ ”, que requiere un tratamiento más sofisticado.

(Ayoubi 1996) ha propuesto un “sistema neuro-borroso adaptable”, donde las funciones de pertenencia de los síntomas y el grado de las conectivas \wedge , \vee , son ajustadas experimentalmente. Aplicaciones resientes en el diagnóstico por sistemas neuro-borrosos adaptables se encuentran en (Sainz 2005; Korbicz 2006).

3.8. Incertidumbre en el conocimiento

Existen diferentes metodologías para tratar la incertidumbre en las medidas, en los sistemas expertos. La teoría de conjuntos, la teoría de conjuntos aproximados (en inglés, *rough sets*), la teoría de la posibilidad, la teoría de las medidas borrosas, la de los conjuntos borrosos y la teoría de evidencia, todas estas son una alternativa, cuando la teoría de la probabilidad no es un marco apropiado.

3.8.1. Teoría de probabilidad

Esta teoría es la base de la teoría de evidencia, que se puede usar en los sistemas expertos, es por eso que aquí se da una introducción.

Con frecuencia, es útil cuantificar la posibilidad de que se presente un resultado de un experimento aleatorio. La probabilidad de un resultado se cuantifica asignándole un número del intervalo $[0, 1]$, o un porcentaje del 0 al 100%. Entre más grande sea el número, mayor es la probabilidad del resultado.

La probabilidad de un resultado puede interpretarse de dos formas: bien objetiva (frecuentista) o bien como la probabilidad subjetiva, o grado de creencia, de que ocurra el resultado. Personas distintas no dudan en asignar probabilidades diferentes a los mismos resultados. Otra interpretación de la probabilidad, se basa en el modelo conceptual (Cooper 2007) de la repetición del experimento aleatorio, este es el punto de vista objetivo de la probabilidad. La probabilidad del resultado, se interpreta como el valor límite de la proporción de veces, que el resultado aparece en n repeticiones del experimento aleatorio, a medida que n crece sin fin. Las probabilidades se eligen de modo que la suma de las probabilidades, de todos los resultados de un experimento sea uno (Montgomery 1996). Esta convención facilita la interpretación de frecuencia relativa para la probabilidad. La probabilidad relativa, es el número de veces que se presenta un evento distintivo, con respecto a un determinado conjunto de eventos.

Una de las características distintiva de la probabilidad, son sus funciones de probabilidad, entre las que se encuentra la distribución de probabilidad (Davis 2007) o distribución de una variable aleatoria X , que es una descripción del conjunto de valores de X (rango de X), junto con la probabilidad asociada con cada uno de estos valores. A menudo la distribución de probabilidad de una variable aleatoria, es el resumen más útil de un experimento aleatorio.

3.8.2. Teoría de posibilidad

La posibilidad es considerada como una confianza objetiva de alcanzar algo, o desde un punto de vista subjetivo, es la alternativa que con más confianza se prefiere (Dubois 1988). En la teoría de evidencia la plausibilidad y la credibilidad son complementarias, esto no sucede en la teoría de posibilidad. Así la expresión, no es probable que “no A ” ($\neg P(\neg A)$) significa que “ A ” es probable (en probabilidad la doble negación es una afirmación porque es complementaria (3.4)). La expresión, no es posible que “no A ” no significa que “ A ” es posible (la posibilidad no es complementaria (3.6)), de esta forma es necesario que “ A ” tenga un significado más fuerte. De manera contraria, es posible que “ A ”, no asegura algo acerca de la posibilidad ni la imposibilidad de que “ A ” no es posible.

$$P(A) + P(\neg A) = 1 \quad (3.4)$$

$$N(A) + N(\neg A) \leq 1 \quad (3.5)$$

$$\Pi(A) + \Pi(\neg A) \geq 1 \quad (3.6)$$

En las ecuaciones anteriores se puede ver la diferencia entre la posibilidad y la probabilidad: la probabilidad de un evento determina completamente la probabilidad del evento contrario (3.4). La posibilidad (o necesidad (3.5)) de un evento, con respecto a un evento contrario, se encuentran débilmente enlazadas. Para poder caracterizar la incertidumbre de un evento A son necesarios ambos números $\Pi(A)$ y $N(A)$.

La necesidad objetiva se refiere a las leyes de la naturaleza: si dejamos caer un objeto es necesario que este caiga. La necesidad subjetiva vista como el dual de preferencia significa prioridad (Dubois 2004). La necesidad objetiva se refiere al límite más bajo de medidas de frecuencia, mientras que la credibilidad subjetiva significa creencia, aceptación y probabilidad probada. Una proposición es creíble sí y solo sí su lógica proviene de un conocimiento disponible. A la vez que la necesidad también con lleva un significado lógico en términos de deducción, así como la posibilidad es un asunto de consistencia lógica.

Una distribución de posibilidad π , en un conjunto de posibles situaciones U que se relacionan de U a un conjunto L , que tiene una valuación linealmente ordenada. Este conjunto se define en el intervalo $[0, 1]$ y se encuentra equipado por las operaciones mínimo y máximo. Una distribución de posibilidad representa un estado de conocimiento, distinguiendo qué es plausible de lo que es menos plausible. Distinguiendo cuál es el curso normal de las cosas, de lo que no es. Distinguiendo qué es lo que sorprende de lo que no. La función π_x representa una restricción flexible en el valor de x con las siguientes convenciones

$$\begin{aligned} \pi_x(u) = 0 & \text{ significa que } x = u \text{ se rechaza como imposible;} \\ \pi_x(u) = 1 & \text{ significa que } x = u \text{ es totalmente posible (=plausible)} \end{aligned}$$

pero simultáneamente distintos valores pueden tener un grado de posibilidad igual a 1. Aplicaciones de la teoría de posibilidad en el diagnóstico de fallos se encuentran en (Detroja 2006), en otros campos se encuentran en (Wei-Guo 2007; Henna 2006)

3.8.3. Teoría de la evidencia

En esta teoría existen dos medidas que tienen su paralelo con la teoría de posibilidad. La posibilidad y la necesidad equivalen a la plausibilidad y a la credibilidad respectivamente en la teoría de evidencia. La plausibilidad y la credibilidad, son interpretadas como la probabilidad superior (pl) y la probabilidad inferior (bel) respectivamente.

Denotemos a E_1, E_2, \dots, E_p como subconjuntos, de pares distintos (pero no necesariamente disjuntos), de Ω (considerado finito), que transportan respectivamente masas de probabilidad $m(E_1), \dots, m(E_p)$ las cuales son tal que

$$\sum_{i=1}^p m(E_i) = 1$$

y

$$\forall i, \quad m(E_i) > 0$$

La masa $m(E_i)$ puede ser interpretada como una localidad global de probabilidad para el conjunto completo de eventos elementales que conforman a E_i , sin especificar como esta masa se encuentra distribuida entre los eventos elementales. Los eventos E_i son llamados “elementos focales” (Shafer 1976) y pueden ser usados para modelar observaciones imprecisas. En esta situación, la probabilidad de un evento A será impreciso, esto es, estará contenido en un intervalo de probabilidad $[bel(A), pl(A)]$, definido por

$$bel(A) = \sum_{E_i \subseteq A} m(E_i)$$

$$pl(A) = \sum_{E_i \cap A \neq \phi} m(E_i)$$

$bel(A)$ se calcula considerando todos los elementos focales, lo cuál hace la ocurrencia de A necesaria. $pl(A)$ se obtiene considerando todos los elementos focales lo cuál hace la ocurrencia de A posible. Notemos que así como en la posibilidad también existe una relación dual en la teoría de evidencia. bel y pl se relacionan de la siguiente forma:

$$\forall A, \quad pl(A) = 1 - bel(\neg A)$$

Según Dubois y Prade (Dubois 1988; Shafer 1976) es posible comprobar que las funciones pl y bel son medidas de posibilidad y necesidad respectivamente, si y sólo si, los elementos focales forman una secuencia de conjuntos anidados ($E_1 \subset E_2 \subset \dots \subset E_p$). Aplicaciones de la teoría de evidencia a diagnóstico de fallos se encuentran en (Bo-Suk 2006; Basir 2005).

3.9. Medidas intervalares borrosas

La idea de que las cantidades borrosas pueden ser aritméticamente combinadas, de acuerdo con las leyes de la teoría de conjuntos borrosos, se debe a Zadeh (1975). Después

se reconoció que las matemáticas de las cantidades borrosas son una aplicación de la teoría de las posibilidades, una extensión del análisis de intervalos así como del álgebra multivaluada. Esta última fue propuesta por Young en 1931, que es en sí misma una extensión del análisis de intervalos, para el cálculo con conjuntos de números. Hoy día existe un extenso cuerpo de resultados disponibles acerca de cantidades borrosas, y se han desarrollado métodos computacionales en muchas clases de aplicaciones (Dubois 2000b).

3.9.1. Cantidades borrosas

La simple definición de cantidades borrosas consiste en establecer que existe un conjunto borroso de números, generalmente números reales, pero algunas veces también enteros. En el conjunto de números reales, no todo el subconjunto de números reales es de interés. Los más útiles son los números intervalares. Con el fin de computar con cantidades borrosas como podrían ser las medidas, se deben definir principios básicos tales que nos permitan extender funciones con argumentos reales, o enteros a argumentos borrosos.

La teoría de la posibilidad ofrece marco de trabajo natural para interpretar el principio de extensión. En este marco de trabajo, las cantidades borrosas son interpretadas como distribuciones de posibilidad, describiendo los valores del conocimiento incierto o impreciso de las variables, y las reglas para computar cantidades borrosas son las usadas en la teoría de posibilidad (Moazeni 2006). Una cantidad borrosa Q , es cualquier conjunto borroso de los números reales \mathfrak{R} . Como cualquier conjunto borroso, una cantidad borrosa puede ser representada ya sea por su función de pertenencia $Q(\cdot)$, que asocia a cualquier valor r el grado $Q(r)$ de la función de pertenencia o por su nivel de corte $Q_\alpha = (a \in \mathfrak{R} : Q(a) \geq \alpha)$ (Abbasbandy 2006).

La importancia práctica del análisis del intervalo borroso estriba en la necesidad existente en áreas como: problemas de diseño, toma de decisiones o análisis de riesgo (Facchinetti 2006), de funciones que computan variables con incertidumbre, o propagando restricciones flexibles, así como el cómputo de distribuciones de incertidumbre de parámetros de interés, o el perfil preferente de las variables de decisión (Abbasbandy 2004). Los principales campos de aplicación de los intervalos borrosos parecen ser: procesamiento de datos (Kreinovich 2007), diseño (Giordani 2006), análisis de decisiones (Zeshui 2005), y estadística

3.9.2. Intervalo borroso

Hay una clase especial de cantidades borrosas las cuales extienden el concepto de un intervalo: un intervalo borroso M es una cantidad borrosa con una función de pertenencia quasi-cóncava, como podría ser un subconjunto borroso convexo de los números reales \mathfrak{R}

$$\forall a, b \forall c \in [a, b], M(c) \geq \min(M(a), M(b)).$$

Note que una cantidad es convexa si y solo si su nivel-corte es convexo, de igual manera lo son los intervalos (limitados o no). Los intervalos cerrados son generalizados por los intervalos borrosos con funciones de pertenencia semicontinuas superiores, también se aplica por definición para aquellos intervalos cuyo nivel-corte sean intervalos cerrados. Nótese que M es un intervalo borroso semicontinuo superior sí y solo sí se cumplen las tres siguientes condiciones

- El núcleo ($x : M(x) = 1$) puede ser cualquiera de los dos, ya sea un intervalo cerrado $[\underline{m}, \bar{m}]$ o bien un intervalo de la forma $[\underline{m}, +\infty)$, $(-\infty, \bar{m}]$, $(-\infty, +\infty)$.
- La restricción $M^-(.)$ de $M(.)$ para $(-\infty, \underline{m})$ (cuando sea apropiada) es una función continua-derecha no-decremental.
- La restricción $M^+(.)$ de $M(.)$ para $(\bar{m}, +\infty)$ (cuando sea apropiada) es una función continua-izquierda no-incremental.

Nótese que en un intervalo borroso semicontinuo superior puede existir un soporte limitado (el soporte de Q es el conjunto $supp(Q) = \{a \in \mathfrak{R} : Q(a) > 0\}$) o puede no existir. Subconjuntos compactos de los números reales (podrían ser cerrados o limitados) que son extendidos por los intervalos borrosos semicontinuos superiores con un soporte limitado. Un número borroso M se define como un soporte limitado de un intervalo borroso semicontinuo superior con un único valor modal m ; esta puede ser una representación de la imprecisamente cantidad especificada “aproximadamente m ”.

3.9.3. Interpretación

Un intervalo borroso es una representación conveniente de rangos de variables especificados imprecisamente, que transportan más información que un usual intervalo no borroso (Bondia 2006). Por ejemplo, cuando uno tiene subjetivamente que estimar el valor de un parámetro desconocido e incontrolable, la representación del intervalo tiende a ser una herramienta pobre con respecto al modelado: ¿se deben de proveer límites pesimistas? (computaciones del procesador producirían resultados demasiado imprecisos para ser de alguna ayuda); ¿O límites optimistas? (entonces se obtendrán resultados más precisos, ¿pero son confiables?). Un intervalo borroso refleja al mismo tiempo límites optimistas y pesimistas: su soporte se construye sobre bases confiables y al mismo tiempo su núcleo contiene sólo sus valores más plausibles. Similarmente cuando se especifican requerimientos acerca de una variable de decisión, una meta, una restricción que debe ser respetada, la naturaleza todo o nada de los intervalos enmascara la presencia de los valores preferidos “más o menos”. Las cantidades borrosas entonces también ofrecen modelos más expresivos con respecto a los requerimientos flexibles.

Así, las cantidades borrosas se pueden interpretar en dos maneras radicalmente diferentes: las “distribuciones de incertidumbre”, que son la contraparte posibilística de las densidades de probabilidad; o “perfiles de preferencia”, que hacen que las cantidades borrosas sean afines a funciones de utilidad. Existen técnicas para la estimación de funciones de pertenencia que se adaptan particularmente para la definición de los intervalos borrosos.

Un intervalo borroso, es una representación más precavida de una cantidad o medida con incertidumbre, que una distribución de probabilidad cuando la única fuente de información es el ser humano (por lo tanto subjetivo). Una razón puede ser, que las reglas de la teoría de posibilidad que gobernará el cálculo de las cantidades borrosas, envuelve sólo la comparación entre los grados de posibilidad. Esto es especialmente verdadero, si una cantidad borrosa es vista como un conjunto de intervalos anidados.

3.9.4. Borrosificación

Encontrar el valor escalar representativo de un intervalo borroso, es frecuente llamado deborrosificación en la literatura del control borroso (Amirkhazadeh 2005; Patel 2002; Yu 2001). Existen varias propuestas:

- **Media de la máxima (MOM)** (*mean of maxima*), la cual es el punto medio en el núcleo del intervalo borroso M ,
- **Centro de gravedad** $GR(M) = \frac{\int_{-\infty}^{+\infty} xM(x)dx}{\int_{-\infty}^{+\infty} M(x)dx}$ este es el centro de gravedad del soporte M , ponderado por el grado de pertenencia.
- **Centro del área (media)**: el valor $COA(M) = m$ tal que $\int_{-\infty}^m M(x)dx = \int_m^{+\infty} M(x)dx$. Este es el punto del soporte M que divide en partes iguales el área bajo la función de pertenencia.
- **Punto medio de la media del intervalo**

$$\bar{E}(M) = \int_0^1 \frac{(\inf M_\alpha + \sup M_\alpha)}{2} d\alpha = \frac{E_*(M) + E^*(M)}{2}$$

La MOM suena natural como representación del intervalo borroso M desde el punto de vista de la teoría de la posibilidad, donde los valores de la posibilidad más alta son considerados valores plausibles por defecto. Esto es en el particular caso cuando el máximo es único. De cualquier manera, esto tiene sentido especialmente si el proceso de revisión esta ordenado de tal forma, que podría afectar el MOM al borrar los valores más plausibles, antes de la llegada de nueva información. En una perspectiva de la estadística, la MOM claramente no explota toda la información contenida en M , debido a que descuida la función de pertenencia.

Sólo la “media del intervalo” tal y como se describe en la ecuación que define a $\bar{E}(M)$, cuenta para definir la específica naturaleza posibilística de un intervalo borroso. La

elección del punto medio expresa una actitud neutral del usuario y extiende la media de la máxima, a una media promedio de niveles-corte. Este valor característico tiene una interpretación natural, en términos de simulación de variables borrosas.

Conclusiones del capítulo.

Se analizaron las ideas más importantes en cuanto a la lógica, los tipos de sistemas basados en reglas. Se dio una introducción a la utilización en ingeniería de sistemas y control. Para la aplicación de las técnicas descritas, el usuario debe tener clara la interpretación de la borrosidad y los tipos de proposiciones más adecuados. De esta forma los resultados obtenidos por el sistema de razonamiento serán *consistentes* con el modelo conceptual del referido usuario o ingeniero. Las anomalías o errores que pueden cometerse en el diseño de las reglas, la incertidumbre en las medidas, los intervalos borroso y la deborrosificación de intervalos borrosos todo esto se ha descrito.

Capítulo 4

Diagnóstico basado en técnicas de Inteligencia Artificial

4.1. Introducción

En el trato con la incertidumbre en el conocimiento, son usuales las técnicas de Inteligencia Artificial. Una de las fuentes de incertidumbre en los sistemas de diagnóstico es la incertidumbre en las medidas o síntomas, usados como nuestra información primaria para el diagnóstico, otra es la incertidumbre en las propias reglas que constituyen el conocimiento. Entre las técnicas más usuales en el diagnóstico de fallos por métodos de Inteligencia Artificial, tenemos los métodos de clasificación, métodos basados en el modelo y los métodos basados en relaciones.

4.2. Métodos de clasificación para el diagnóstico

La propagación de los fallos a características observables o síntomas, en general siguen una relación física causa-efecto (fallo-síntoma), donde las propiedades físicas y variables, están conectadas una a la otra cuantitativamente, y también como una función del tiempo. Sin embargo, si no existe información disponible en las causalidades fallo-síntoma (causa-efecto), entonces es posible usar métodos de clasificación, entrenados experimentalmente para el diagnóstico de fallos. Lo que nos conduce a un método basado en el conocimiento no estructurado.

Un objeto que se inspecciona para un proceso de “reconocimiento” se llama rasgo o patrón. Habitualmente nos referimos a un patrón como a una descripción de un objeto que se quiere reconocer. El proceso de reconocimiento conlleva a un proceso de clasificación. Para determinar a que clase un objeto pertenece, se debe inicialmente encontrar los rasgos característicos o patrones necesarios para poder realizar esta clasificación.

4.2.1. Principios del método de clasificación

Los síntomas se representan en un vector de características $\underline{S}^T = [S_1, S_2, \dots, S_n]$. Los fallos correspondientes son conocidos en la forma vectorial $\underline{F}^T = [F_1, F_2, \dots, F_n]$, los elementos \underline{F} pueden ser binarios $F_j \in [0, 1]$ o también pueden representar medidas graduales del tamaño de los fallos $F_j \in [0 \dots 1]$.

Si no se dispone de un conocimiento completo en la relación de las características (causas) con los fallos (efectos), entonces es necesario el uso de métodos para el reconocimiento de patrones. En la figura 4.1 el vector de referencia \underline{S}_n es determinado por el comportamiento normal. Después de esto el vector correspondiente de entrada \underline{S} de las características es determinado experimentalmente para ciertos fallos F . La relación entre \underline{F} y \underline{S} es aprendida (o entrenada) experimentalmente y almacenada, formando una base de conocimiento explícito. Comparando el \underline{S} observado con el \underline{S}_n normal de referencia es posible identificar para su clasificación el fallo \underline{F} .

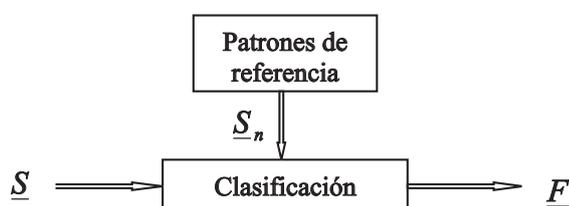


Figura 4.1: Diagnóstico de fallos usando métodos de clasificación

En base a los principios expuestos, existen métodos de clasificación estadísticos, como el clasificador de Bayes (Wang 2007) y otros más. También existen métodos de clasificación geométricos, basados en polinomios (Cairó 2007), en redes neuronales (Manevitz 2007) o los clasificadores borrosos (*fuzzy clustering*) (Özyer 2007; Belacel 2007) los cuales son aplicados en el siguiente apartado.

4.3. Métodos basados en modelos

Este enfoque es radicalmente diferente a los enfoques tradicionales. Se proponen nuevas técnicas de Inteligencia Artificial que emplean un modelo de propósito general de la estructura interna y del comportamiento del sistema a diagnosticar. Estos modelos se pueden construir empleando técnicas tradicionales de Inteligencia Artificial como la lógica de predicados, estructuras, condicionantes y reglas.

4.3.1. Técnicas de los métodos basados en el modelo

En la literatura se distinguen dos técnicas o corrientes que explotan formas de conocimiento sensiblemente diferente: la técnica basada en la coherencia y la basada en la abducción o explicación.

- *La técnica basada en la coherencia*, donde el objetivo es encontrar un conjunto mínimo de fallos, de tal manera que las observaciones sean “coherentes” con el funcionamiento nominal de todos los demás componentes. Esta técnica se fundamenta en la idea de que es posible comparar el funcionamiento de un sistema con una referencia, el modelo. En este método un síntoma es un “desfase o discrepancia” entre el comportamiento observado y el comportamiento previsto durante la operación por el modelo, o sea, el comportamiento de referencia.
- *La técnica basada en la abducción o explicación*, donde el objetivo es encontrar un conjunto mínimo de hipótesis de fallos que asociadas al conocimiento profundo del sistema “impliquen” a las observaciones. Los métodos de diagnóstico basados en la abducción utilizan un modelo de estructura y comportamiento o simplemente de comportamiento que describe las relaciones de causa-efecto entre los fallos y sus manifestaciones, sin tener en cuenta el funcionamiento nominal del sistema real como en los modelos basados en la coherencia.

Estas dos técnicas también son usadas en los métodos relacionales (FMECA), que se verán en la sección 4.4.

4.3.2. Método basado en el modelo de eventos discretos

La metodología propuesta por (García 1999) se basa en sistemas de eventos discretos (DES) y en el concepto de diagnosticador.

El modelo de eventos discretos así como el de diagnosticador (ver figura 4.2), son máquinas de estado finito (FSM). El modelo incorpora el comportamiento operacional de varios componentes, e incorpora estados adicionales que representan el comportamiento de fallos de los mismos. La fase de modelado inicia con el modelado individual de todos los componentes que forman el sistema, incluyendo su comportamiento en estado normal así como de fallo. El conjunto de posibles fallos que se incorporen a los modelos, se obtiene del conocimiento empírico y de los criterios de ingeniería. Cada componente individual es modelado por su expresión generadora

$$G_i = (X_i, \Sigma_i, \delta_i, x_{0i})$$

donde X_i es el espacio de estado finito, Σ_i es el conjunto de eventos, δ_i es la función de transición, y x_{0i} es el estado inicial de G_i . Los estados en X_i y los eventos en el alfabeto Σ_i , reflejan el comportamiento normal y de fallo del i -ésimo componente. Algunos de los eventos en δ_i son observables, mientras que el resto son no observables, $\Sigma_i = \Sigma_0 \cup \Sigma_{u0}$.

- *El operador de composición síncrono* se puede usar para modelar la operación de unión de cualquier número de DES. En la composición síncrona se encuentran dos casos:

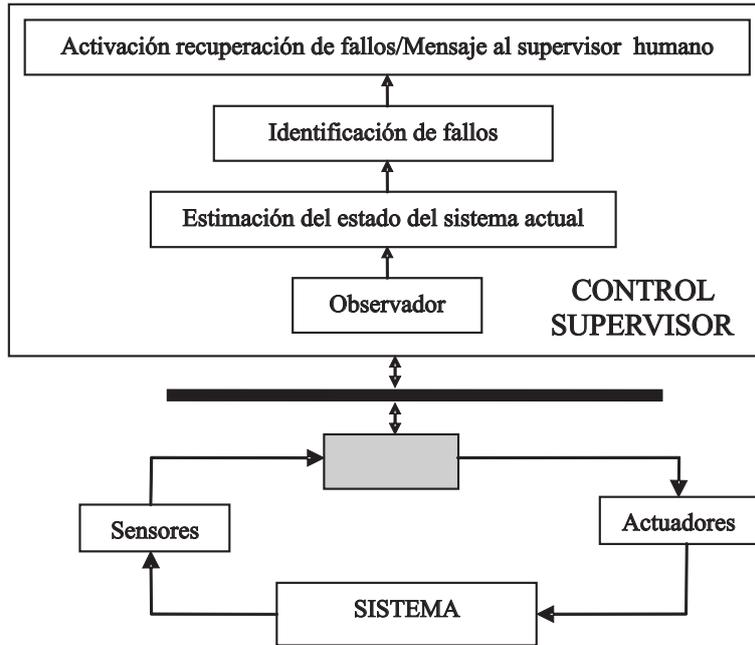


Figura 4.2: Arquitectura para el diagnóstico

a) $\Sigma_1 = \Sigma_2$, donde la composición síncrona es reducida a otros operadores entre FSMs (*Finite States Machines*), y es llamada composición síncrona completa.

b) $\Sigma_1 \cap \Sigma_2 = 0$, donde en este caso, la composición síncrona es en realidad una composición asíncrona, debido a que las transiciones síncronas no existen. Este caso concreto recibe el nombre de “producto mezclado”.

A través del operador de composición síncrono, se modela la operación de “unión” (denotada por el símbolo \parallel) de todos los generadores individuales que participan en el modelo global. Por ejemplo, para el caso de un par de generadores $G1$ y $G2$, el modelo global será $G = G1 \parallel G2$.

El DES resultante deberá ser determinado por la acción concurrente de $G1$ y $G2$, bajo la suposición de que estas acciones son asíncronas e independientes. Los estados de G son pares ordenados (X, Y) , donde X es estado de $G1$ e Y es estado de $G2$, las transiciones de G son también de la forma $(X, Y) \rightarrow (X', Y)$ donde $X \rightarrow X'$ es una transición de $G1$, o de la forma $(X, Y) \rightarrow (X, Y')$ donde $Y \rightarrow Y'$ es una transición en $G2$.

- *El concepto de diagnosticador.* El diagnosticador Gd es una FSM, obtenida del modelo de eventos discretos del sistema. Cada estado del diagnosticador q_d es de la forma

$$q_d = \{(x_1, l_1), \dots, (x_n, l_n)\}$$

este tiene información e indica el conjunto de estados x en aquellos que el sistema se puede encontrar. De hecho cada estado del diagnosticador q_d , se forma al incluir todos los posibles estados de destinación x_i , que deberían ser alcanzados como consecuencia de la ocurrencia de cierto evento observable σ . Estos estados q_d , son formados por pares x, l donde l puede ser una etiqueta normal N , una

etiqueta ambigua A o una etiqueta de fallo F . El objetivo consiste en hacer que el diagnosticador, sea capaz de diagnosticar de manera segura la ocurrencia de algún tipo de fallo, con un retardo finito como consecuencia de la ocurrencia de un evento de fallo F_i . Se puede pensar del diagnosticador Gd como un observador extendido en el cual, nosotros colocamos a cada estado estimado de la forma mencionada arriba.

El algoritmo de diagnóstico que constituye al diagnosticador, es ejecutado en el nivel de supervisión. Este recibe información en-línea del controlador local a bajo nivel y lleva a cabo la tarea de diagnóstico. Sin embargo, en otra etapa de diseño, se establecieron condiciones de diagnosticabilidad del modelo fuera-de-línea. Como las condiciones de diagnosticabilidad son alcanzadas al incorporar nuevos sensores, se puede iniciar analizando las condiciones de diagnosticabilidad del sistema fuera-de-línea, inicialmente, con esos que necesariamente son sensores estrictos, usados para llevar a cabo las tareas de control y monitorizado.

El hecho que un modelo sea no diagnosticable, significa que existirá cierta incertidumbre con respecto a si ha existido o no, un fallo en el sistema, alguna cosa que impida afirmar con toda seguridad que el fallo mencionado ha ocurrido. Es necesario para reestablecer la fase de modelado, introducir los instrumentos de detección necesarios, que permitan cambiar las condiciones de diagnosticabilidad. Por otro lado, para una correcta operación el diagnosticador debe ser capaz de establecer, que un fallo del tipo F_i se ha producido.

Específicamente, se dice que el sistema es diagnosticable, si el diagnosticador correspondiente entra en un bucle de estados F_i -cierto, donde todos los estados q_i del modelo del diagnosticador, contienen solamente una etiqueta de fallo F_i aislado.

Una explicación más amplia, sobre los métodos de diagnóstico basados en modelos de eventos discretos, se encuentra en (García 2000).

4.4. Métodos basados en relaciones

Los métodos basados en relaciones, utilizados ampliamente en el diagnóstico médico, utilizan una relación directa R entre los fallos posibles, el conjunto $\mathcal{D} = \{d_1, \dots, d_n\}$ y sus manifestaciones, el conjunto $\mathcal{M} = \{m_1, \dots, m_n\}$ y emplean como herramienta la teoría de los conjuntos convencionales o borrosos y en algunos casos la teoría de las probabilidades. A continuación se describen las características generales del método relacional borroso, que será usado en la aportación principal de esta tesis.

4.4.1. Método relacional borroso aplicando FMECA

La incompletitud es inherente de FMECA (*Failure Mode, Effects and Criticality Analysis*) (Cayrac 1996). La meta es proponer en una representación flexible los efectos de

los modos de fallo, que faciliten la expresión del conocimiento. El conjunto de los posibles modos de fallo, o desórdenes es $D\{d_1, \dots, d_i, \dots, d_p\}$ obtenido de una base de conocimientos causales y de un conjunto de efectos observados $M\{m_1, \dots, m_j, \dots, m_n\}$ relacionados a los modos de fallo.

En este modelo, los desórdenes y manifestaciones se pueden encontrar ausentes o presentes. Para un desorden dado d expresamos sus manifestaciones (más o menos) certeras, definidas en un conjunto borroso $M(d)^+$, y aquellas que son (más o menos) imposibles, en un conjunto borroso $M(d)^-$. Así el conjunto $M(d)^+$ contiene manifestaciones las cuales son (más o menos) vinculadas, o si lo preferimos causadas o producida por un fallo aislado d . En términos de funciones de pertenencia $\mu_{M(d)^+}(m) = 1$ lo cual significa que m esta siempre presente cuando un fallo aislado d se presenta. De manera contraria, el conjunto $M(d)^-$ contiene manifestaciones que con (más o menos) seguridad no pueden ser causadas por un fallo aislado d . De modo que $\mu_{M(d)^-}(m) = 1$ lo que significa que m nunca se presentará cuando un fallo aislado d ocurra.

Una manifestación m se determina al momento de ser “observada” como presente o como ausente, M^+ y M^- son conjuntos borrosos de manifestaciones las cuales se encuentran respectivamente (mas o menos cierto) presentes y (mas o menos cierto) ausentes. $\mu_{M^+}(m) = 1$ significa que con toda seguridad m esta presente; $\mu_{M^-}(m) = 1$ significa que con toda seguridad m esta ausente. Una completa ignorancia con respecto a m es representada por $\mu_{M^+}(m) = \mu_{M^-}(m) = 0$. Se condiciona a que $M^+ \cap M^- = 0$.

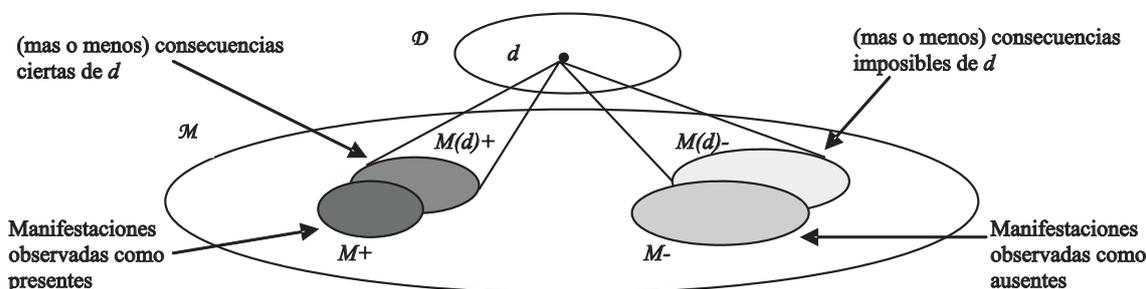


Figura 4.3: d es una explicación posible: las observaciones son consistentes con sus manifestaciones predichas, presentes y ausentes

Diagnóstico basado en la consistencia

Dados los conjuntos M^+ y M^- en determinado momento del proceso de diagnóstico, el problema es encontrar el conjunto singular de desórdenes \hat{D} que sea consistente con estos síntomas (ver la figura 4.3), que son todos los desórdenes que no contradicen las observaciones.

En el caso de los conjuntos binarios (*crisp*), un desorden singular es relacionado con los síntomas, si, ninguna de sus consecuencias que con toda seguridad se esperaban es realmente ausente y ninguna de sus consecuencias consideradas imposibles es presente

$$\hat{D}_{crisp} = \{d \in D, M(d)^+ \cap M^- = 0 \quad \text{y} \quad M(d)^- \cap M^+ = 0\}$$

\hat{D}_{crisp} Contiene todos los desórdenes potenciales.

La extensión de la ecuación anterior a conjuntos borrosos de fallos, puede ser hecha usando la consistencia de min-max de dos conjuntos borrosos. La consistencia de dos conjuntos borrosos F y G evalúa a qué extensión la intersección de ellos no resulta vacía. Al grado que $F \cap G \neq 0$. Y es definida por

$$cons(F, G) = \max_u (\min(\mu_F(u), \mu_G(u)))$$

el grado para lo cual $M(d)^+ \cap M^- = 0$ es $1 - cons(M(d)^+, M^-)$. \hat{D} puede así ser definida por

$$\forall d \in D \quad \mu_{\hat{D}}(d) = \min(1 - cons(M(d)^+, M^-), 1 - cons(M(d)^-, M^+))$$

$$\mu_{\hat{D}} = 1 - \max_{i=1, n} \max(\min(\mu_{M(d)^+}(m_i), \mu_{M^-}(m_i)), \min(\mu_{M(d)^-}(m_i), \mu_{M^+}(m_i)))$$

Note que los más consistentes son $M(d)^+$ y M^- el más inconsistente es d con el conjunto M^- de las manifestaciones ausentes. $cons(M(d)^+, M^-)$ es así un grado de conflicto entre d y M^- . La última ecuación que describe a $\mu_{\hat{D}}$, claramente expresa que un fallo d es al menos una explicación candidata de cómo los conjuntos borrosos, a partir de sus más o menos ciertos efectos, solapan al conjunto borroso de manifestaciones que (más o menos) ciertamente se encuentran ausentes, o cómo el conjunto borroso de efectos los cuales podemos considerar con más o menos certeza, están ausentes cuando d se encuentra presente y traslapan al conjunto borroso de manifestaciones (ver figura 4.4) las cuales son más o menos observadas con certeza como presentes.

Así \hat{D} contiene los fallos considerados más o menos posibles con respecto a las observaciones actuales. Un desorden d tal que $\mu_{\hat{D}}(d) = 0$ puede con toda seguridad descartarse por considerarse imposible de suceder.

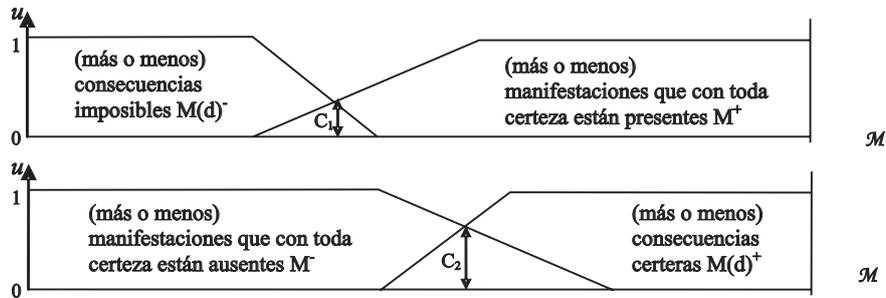


Figura 4.4: Solapamiento de conjuntos borrosos

Diagnóstico basado en la abducción

El conjunto \hat{D} agrupa las posibles causas de los síntomas observados conteniendo todos los desórdenes que están (más o menos) vinculados con los síntomas. Debe ser notado que los desórdenes irrelevantes, los cuales no pueden ser relacionados con las observaciones disponibles (desórdenes cuyas consecuencias son completamente desconocidas), permanecen como perfectos candidatos con respecto a nuestro acercamiento basado en la consistencia. Por ende es interesante calcular subconjuntos de \hat{D} que contengan desórdenes que sean directamente relevantes para las observaciones actuales.

Un desorden es relevante, si alguna de sus consecuencias (mas o menos) ciertas se encuentran verdaderamente presentes o si alguna de sus consecuencias (mas o menos) imposibles, se encuentran verdaderamente ausentes (la segunda parte de esta definición en la práctica no se usa).

En el caso de la lógica binaria (*crisp*). El subconjunto de desórdenes relevantes es entonces

$$\hat{D}_{crisp}^* = \{d \in \hat{D}_{crisp}, M(d)^+ \cap M^+ \neq 0 \text{ ó } M(d)^- \cap M^- \neq 0\}$$

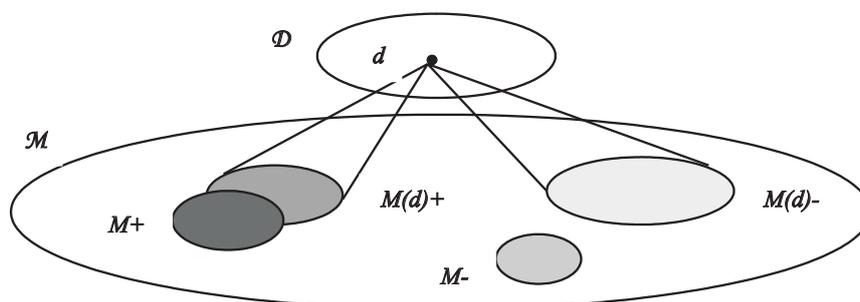


Figura 4.5: d es una explicación relevante: algunas de las manifestaciones presentes, predichas con anterioridad, son realmente observadas en el presente

Este conjunto sólo contiene los desórdenes que no son reglamentados por los síntomas y que son sugeridos por ellos (ver figura 4.5). Esto puede ser fácilmente extendido al caso borroso usando el índice de consistencia borrosa. Es decir

$$\mu_{\hat{D}}^*(d) = \min(\mu_{\hat{D}}(d), \max(\text{cons}(M(d)^+, M^+), \text{cons}(M(d)^-, M^-)))$$

Un desorden d es completamente relevante tan pronto como una de sus consecuencias seguras es ciertamente observada como presente (o tan pronto como una de sus imposibles consecuencias es ciertamente observada ausente, si es apropiado). En el caso de fallos múltiples, los desordenes relevantes pueden ser usados como “bloques de construcción” para construir la explicación de fallos-múltiples.

4.4.2. Tasas de cobertura en los métodos relacionales

Estas explican hasta que medida un fallo elemental es la causa del conjunto de efectos observados.

Tasa de coherencia o de consistencia de un fallo “ d ”

Es la medición que indica, hasta que punto sus efectos son compatibles con las observaciones

$$Cob(d) = \frac{|M^+ \cap M(d)^-|}{|M^*|}$$

Si las manifestaciones presentes y ausentes, deben ser cubiertas, entonces la tasa es

$$Cob(d) = \frac{|M^+ \cap M(d)^-| + |M^- \cap M(d)^-|}{|M^*| + |M^-|}$$

donde M^* , confirma que al menos uno de los efectos ciertos de “ d ”, se observa presente cuando “ d ”, es una explicación coherente. La tasa de cobertura de un fallo “ d ”, es igual a 1, cuando todas las observaciones son previstas por “ d ”, y es igual a 0, cuando ninguna de las observaciones es prevista por “ d ”. Es necesario señalar, que cuando $Cob(d) \geq 0$, es decir, que una observación es prevista por “ d ”, entonces “ d ”, es una explicación pertinente.

La tasa de cobertura, es una medida gradual intermedia, entre la pertinencia y la cobertura, y evalúa, hasta que medida un fallo elemental, explica el conjunto de efectos observados.

La tasa de observación de los efectos de un fallo “ d ”

Es la medida de la proporción de los efectos de “ d ”, que son confirmados

$$Obs(d) = \frac{|M(d)^+ \cap M^+| + |M(d)^- \cap M^-|}{|M(d)^+| + |M(d)^-|}$$

la tasa de observación de los efectos de un fallo “ d ”, es igual a 1, cuando todos sus efectos se observan, y a 0, cuando ninguno de sus efectos es observado. En el caso cuando “ d ” es incoherente, la tasa de observaciones de “ d ”, es necesariamente más pequeña que uno. De hecho, por definición, ciertos de sus efectos previstos no son confirmados, lo contrario es observado.

Cuando la tasa de observación de los efectos, es igual a 1, ninguna observación suplementaria, puede volver a los efectos del fallo incoherentes. Sin embargo, su tasa de cobertura se fundamentará en que otras observaciones estén presentes. De hecho, es fácil observar que la tasa de cobertura de un fallo, no se “fija”, hasta que todas las manifestaciones de \mathcal{M} , hayan sido evaluadas.

Tasa de cobertura debil

Ahora, si consideramos el caso más general, en el cuál los conocimientos de los efectos de un fallo, (consecuencias) son incompletos: $\exists d \in \mathcal{D}, M(d)^+ \cup M(d)^- \neq \mathcal{M}$, y la información de las manifestaciones presentes o ausentes, es igualmente incompleta: $M^+ \cup M^- \neq \mathcal{M}$. Entonces, existirá un conjunto de consecuencias $M(d)^{+?}$, que reúne ciertas condiciones no especificas suplementarias, relacionadas a su contexto de aparición. $M(d)^{+?}$ Representa las variables de contexto no especificadas y se representan en la figura 4.6. Dado que $M(d)^+ \subseteq M^+ \subseteq M(d)^{+?} \subseteq \overline{M^-}$ entonces $M^- \cup M(d)^{+?} = \mathcal{M}$.

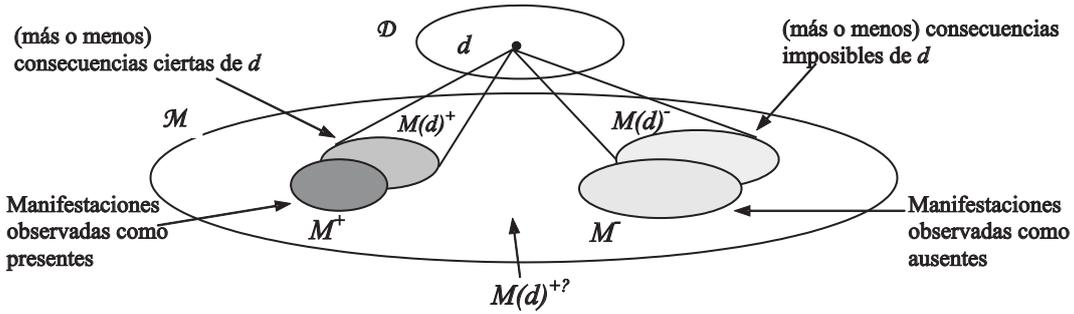


Figura 4.6: Modelo relacional que muestra los efectos de contexto no especificado $M(d)^{+?}$

Cuando los efectos de un fallo dado, aparecen solamente en ciertos contextos no especificados, estos efectos no hayan su lugar ni en $M(d)^+$, y lógicamente ni en $M(d)^-$, debido a que no son ni ciertos, ni imposibles (Cayrac 1995).

Sea $M(d)^{+?}$ el conjunto de manifestaciones presentes, cuando el fallo “ d ” esta presente y cuando eventualmente, se reúnen ciertas condiciones no específicas suplementarias, relacionadas a su contexto de aparición.

Se tiene por lo tanto que $M(d)^+ \subseteq M(d)^{+?} \subseteq \overline{M^-}$: cuando “ d ” ha ocurrido, sus consecuencias ciertas existen, sea cuál fuese el contexto, y ninguna de las consecuencias imposibles, esta presente (excluidas), sea cuál fuere el contexto.

Estas informaciones suplementarias, no cambiaran nada la evaluación de la coherencia de un fallo, en relación a las observaciones: la ausencia de una consecuencia “dependiente del contexto” “ m ”, del fallo “ d ”, no permite rechazar a “ d ”. Es posible que no este en el contexto en el cuál “ m ” aparece. Al contrario, la observación de “ m ” puede sugerir la presencia de “ d ”, es decir, suministrar una explicación pertinente de “ d ”.

Parece en efecto razonable preferir una explicación, donde se observa la presencia de al menos un efecto posible, que aparece en ciertos contextos, a las hipótesis del fallo, por las cuales ninguna relación con “ m ” se ha precisado. También se puede razonar lo mismo para la evaluación de la cobertura. Una explicación, en donde los efectos son ciertos y los efectos aparecen en ciertos contextos, cubre todas las anomalías observadas y es una explicación suficiente.

Se puede entonces, “suministrar” los criterios de pertenencia y de cobertura en un sentido más débil o laxo, y una tasa de cobertura “débil”, reemplazando $M(d)^+$ por $M(d)^{+?}$. La tasa de cobertura “débil” se define por ejemplo por

$$Cob'(d) = \frac{|M^+ \cap M(d)^{+?}|}{|M^+|} = \frac{\sum_{m \in M} \min [\mu_{M^+}(m), \mu_{M(d)^{+?}}(m)]}{\sum_{m \in M} \mu_{M^+}(m)}$$

la anterior tasa de cobertura se extendió a los conjuntos borrosos, así mismo, se pueden extender las tasas de coherencia y observación.

Conclusiones del capítulo

Se han estudiado algunas técnicas de Inteligencia Artificial en el diagnóstico de fallos. La primera de clasificadores borrosos que requiere de algunas otras técnicas como el reconocimiento de patrones, donde también son usuales técnicas de Inteligencia Artificial como las redes neuronales, además de los sistemas expertos borrosos como el estudiado en este capítulo. También el método de diagnosticador por el uso de máquinas de estado finito. Y por último los métodos relacionales, que según lo expuesto, podemos ver que es una herramienta muy útil para modelar el entorno general de un sistema de diagnóstico. En este esquema, podemos visualizar fácilmente las relaciones causa y efecto haciendo el modelado del conocimiento una tarea sencilla para el experto en diagnóstico, que no se complicará con conceptos muy difíciles, como los usados en los métodos analíticos. Esta herramienta es muy adecuada para la interpretación lingüística del conocimiento usado en el diagnóstico de fallos, que trata conceptos como efectos (síntomas o manifestaciones) y causas (desordenes, anomalías o fallos), donde las gradualidades de los mismos siempre están presentes. Es por esto que en la aportación principal de esta tesis es la técnica que se ha utilizado y extendido.

Capítulo 5

Combinación de lógica borrosa y probabilidad

5.1. Introducción

El razonamiento evidencial se basa principalmente en tres marcos de trabajo teórico. La teoría probabilística Bayesiana, la teoría de evidencia de Dempster-Shafer (D-S) (Shafer 1976), y la teoría de conjuntos borrosos (Zadeh 1965). La teoría probabilística Bayesiana es una teoría bien desarrollada en lo referente a toma de decisiones, pero requiere de juicios de probabilidad precisos (Yen 1990). Por lo que, es débil en la representación y manejo de información imprecisa. La teoría de D-S se basa en la teoría de la probabilidad, también ésta necesita juicios probabilísticos en la captura de la naturaleza imprecisa de la evidencia. Como resultado, el grado de probabilidad es medido por intervalos de probabilidad, lo cual es opuesto a las probabilidades de punto de la teoría Bayesiana. La teoría de conjuntos borrosos se enfoca en la representación y manejo de información vaga tal como “la temperatura es alta” o “el misil se encuentra como a 50 Km. de la base”. Una de sus fortalezas es que se encuentra fundada en la teoría de la posibilidad, útil en el manejo de datos imprecisos.

5.2. Teoría probabilista Bayesiana

La probabilidad trata con la aleatoriedad de las variables medidas en un proceso físico, esta aleatoriedad es producida por pequeñas variaciones en las variables no controladas dentro del proceso (por ejemplo, movimiento browniano y efecto térmico sobre los electrones). En algunos casos, las variaciones aleatorias observadas son tan pequeñas en relación con las metas perseguidas, que pueden ignorarse. Sin embargo, la variación casi siempre está presente y su magnitud puede ser tan importante a tal grado, que las conclusiones no sean muy evidentes (las perturbaciones causan incertidumbre en las medidas). En estos casos, son muy valiosos los métodos probabilísticos para el

modelado y análisis de datos (Montgomery 1996). Las características con las que la probabilidad y estadística modelan el comportamiento de las variables en los sistemas son: probabilidad, probabilidad condicional, media, varianza, desviación estándar, etc.

5.2.1. Corolario de probabilidad

- a) Un experimento aleatorio es aquel que proporciona diferentes resultados aun cuando se repita siempre de la misma manera.
- b) El conjunto de los posibles resultados de un experimento aleatorio recibe el nombre de espacio muestral del experimento. El espacio muestral se denota con la letra S .
- c) Un evento es un subconjunto del espacio muestral de un experimento aleatorio.
- d) Dos eventos, E_1 y E_2 tales que $E_1 \cap E_2 = \phi$, se dice que son mutuamente excluyentes, donde ϕ significa conjunto vacío.
- e) Cada vez que un espacio muestral esté formado por N posibles resultados igualmente probables, la probabilidad de cada uno de ellos será $1/N$.
- f) Para un espacio muestral discreto, la probabilidad de un evento E , denotada como $P(E)$, es igual a la suma de las probabilidades de los resultados en E .
- g) Si para los eventos A y B su intersección es $A \cap B \neq \phi$, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- h) Se dice que una colección de eventos, E_1, E_2, \dots, E_k , es mutuamente excluyente si, para todos los pares, $E_i \cap E_j = \phi$.
- i) Para una colección de eventos mutuamente excluyentes,

$$P(E_1 \cup E_2 \cup \dots \cup E_k) = P(E_1) + P(E_2) + \dots + P(E_k)$$

5.2.2. Corolario de probabilidad condicional

- a) Es necesario volver a evaluar la probabilidad a medida que se tiene más información disponible.
- b) La probabilidad condicional de un evento A dado un evento B , denotada por $P(A|B)$, es

$$P(A|B) = P(A \cap B)/P(B)$$

- c) La probabilidad total para cualquier par de eventos A y B , es

$$P(B) = P(B \cap A) + P(B \cap \bar{A}) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$$

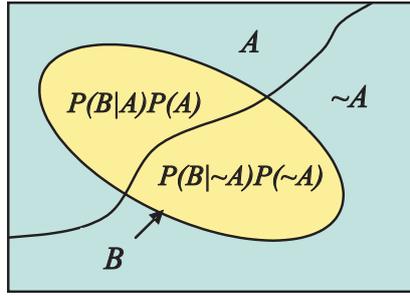


Figura 5.1: Particionamiento de un evento en dos subconjuntos mutuamente excluyentes

- d) Supóngase que E_1, E_2, \dots, E_k , son k conjuntos mutuamente excluyentes y exhaustivos. Entonces,

$$P(B) = P(B \cap E_1) + P(B \cap E_2) + \dots + P(B \cap E_k)$$

$$P(B) = P(B|E_1)P(E_1) + P(B|E_2)P(E_2) + \dots + P(B|E_k)P(E_k) \quad (5.1)$$

- e) Se dice que dos eventos son independientes si, y sólo si, cualquiera de las siguientes proposiciones es verdadera,

$$\begin{aligned} P(A|B) &= P(A) \\ P(B|A) &= P(B) \\ P(A \cap B) &= P(A)P(B) \end{aligned}$$

de tal manera, que si A y B son dos eventos independientes y B no es imposible, entonces la probabilidad de A no es modificada, cuando se recibe una nueva información en B .

- f) Los eventos E_1, E_2, \dots, E_n son independientes, si, y sólo si, para cualquier subconjunto $E_{i_1}, E_{i_2}, \dots, E_{i_k}$,

$$P(E_{i_1} \cap E_{i_2} \cap \dots \cap E_{i_k}) = P(E_{i_1}) \times P(E_{i_2}) \times \dots \times P(E_{i_k})$$

5.2.3. Teorema de Bayes

De la probabilidad condicional,

$$P(A \cap B) = P(A|B)P(B) = P(B \cap A) = P(B|A)P(A) \quad (5.2)$$

Ahora, si se consideran los términos segundo y último de la expresión anterior (5.2), entonces

$$P(A|B) = P(B|A)P(A)/P(B) \quad (5.3)$$

Éste es un resultado útil que permite determinar $P(A|B)$ en términos de $P(B|A)$. En general, si se escribe la $P(B)$ del denominador de la ecuación (5.3) utilizando la regla de la probabilidad total, ecuación (5.1), se obtiene el siguiente resultado general, conocido como “teorema de Bayes”.

Si E_1, E_2, \dots, E_k son k eventos exhaustivos y mutuamente excluyentes, y B es cualquier evento, entonces

$$P(E_1|B) = \frac{P(B|E_1)P(E_1)}{P(B|E_1)P(E_1) + P(B|E_2)P(E_2) + \dots + P(B|E_k)P(E_k)}$$

Es de notar que cuando se tienen datos con cierta precisión y exactitud, el teorema de Bayes proporciona resultados que nos pueden ayudar en la toma de decisiones, con un aceptable nivel de confianza.

5.3. Teoría de evidencia

La teoría de evidencia o teoría de Dempster-Shafer de la función de creencia, nació durante una lectura sobre inferencia estadística dada por Arthur Dempster en la Universidad de Harvard durante los años sesentas. Glenn Shafer contribuyó reinterpretando la idea principal de Dempster con su *teoría de evidencia matemática* (Shafer 1976).

Consideremos dos espacios Ω y Θ , y una relación multivaluada Γ asociando el subconjunto $\Gamma(\omega) \subset \Theta$ para cada elemento $\omega \in \Omega$. Asumiremos que P es una medida de probabilidad definida en (Ω, A) hecha de la σ -álgebra A de un subconjunto de Ω . Considerando que P representa la probabilidad de ocurrencia de un evento incierto $\omega \in \Omega$, y si se establece que este evento ω se encuentra en correspondencia con el evento $\theta \in \Gamma(\omega)$, ¿Qué juicio probabilístico podríamos hacer, con respecto a la ocurrencia de eventos inciertos $\theta \in \Theta$?

El punto de vista de Dempster es que las consideraciones anteriores llevan al concepto de medidas de probabilidad compatibles. Entonces, él define la existencia de un área delimitada por la probabilidad inferior y la probabilidad superior de esta familia probabilista (Brucq 2002).

El espacio probabilístico (Ω, A, P) es la fuente de información que permite la cuantificación del (imperfecto) estado de conocimiento sobre el nuevo referencial Θ por medio de Γ .

Se define el espacio probabilístico formado por $(\Omega, P, \Gamma, \Theta)$ como la estructura de creencia. A través del uso de estas herramientas matemáticas, Shafer propuso otra interpretación del trabajo de Dempster, esta interpretación, identifica la probabilidad inferior y superior de la familia de medidas compatibles, como auténticas medidas de confianza.

Si Θ es un espacio finito y $2^\Theta (= P(\Theta))$ el conjunto de las partes de Θ . Una función de credibilidad *bel* (*belief function*) es una aplicación de 2^Θ con valores en $[0, 1]$ que satisfacen las siguientes condiciones:

- $bel(\phi) = 0$,
- $bel(\Theta) = 1$,

– Para todo entero n y toda la familia de subconjuntos A_1, \dots, A_n de Θ

$$bel(A_1 \cup \dots \cup A_n) \geq \sum \left\{ (-1)^{|I|+1} bel(\cap_{i \in I} A_i); I \subset \{1, \dots, n\}, I \neq \emptyset \right\} \quad (5.4)$$

La condición en (5.4) es llamada condición general de subaditividad. Cuando $n = 2$ entonces,

$$bel(A_1 \cup A_2) \geq bel(A_1) + bel(A_2) - bel(A_1 \cap A_2)$$

La función de credibilidad bel permite cuantificar la información parcial en Θ . Un mapeo multivaluado del espacio Ω al espacio Θ asocia cada elemento en Ω con un conjunto de elementos en Θ , esto es $\Gamma : S \rightarrow 2^\Theta$. La imagen de un elemento ω en Ω bajo este mapeo es llamado el granulo de ω , denotado como $G(\omega)$.

El mapeo multivaluado puede también ser visto como una relación de compatibilidad C entre Ω y Θ . Una relación de compatibilidad C entre Ω y Θ caracteriza la relación posibilística entre sus elementos. Un elemento ω de Ω es compatible con un elemento θ de Θ si es posible que ω sea una respuesta para Ω y θ sea una respuesta para Θ al mismo tiempo, y el granulo de ω es el conjunto de todos los elementos en Θ que son compatibles con ω .

$$G(\omega) = \{\theta | \theta \in \Theta, \omega C \theta\}$$

Dada una distribución de probabilidad del espacio Ω y una relación de compatibilidad entre Ω y Θ , se induce una asignación de probabilidad básica del espacio Θ , denotado por $m : 2^\Theta \rightarrow [0, 1]$, esta asignación de probabilidad básica es:

$$m(A) = \frac{\sum_{G(\omega_i)=A} p(\omega_i)}{1 - \sum_{G(\omega_i)=\emptyset} p(\omega_i)} \quad (5.5)$$

donde el subconjunto A es también llamado elemento focal. Si se asume que Γ no mapea ningún elemento del espacio Ω al conjunto vacío, el denominador (factor de normalización) en (5.5) viene a ser uno (Yen 1990).

La distribución probabilística del espacio Θ , a la que se le refiere como *el marco de discernimiento*, es delimitada por la asignación de probabilidad básica, pero en general, esto no es únicamente determinado por la asignación de probabilidad básica. La medida de creencia bel y la medida de plausibilidad pl de un conjunto B son, respectivamente, la probabilidad inferior y la probabilidad superior del conjunto sujeto a estas delimitaciones. Estas dos cantidades son obtenidas de la asignación de probabilidad básica como sigue:

$$bel(B) = \sum_{A_i \subset B} m(A_i)$$

$$pl(B) = \sum_{A_i \cap B \neq \emptyset} m(A_i)$$

De aquí que, el intervalo de creencia $[bel(B), pl(B)]$ es el intervalo de probabilidad de B , la longitud de este intervalo nos da la medida de imprecisión con respecto al valor de la incertidumbre.

Una importante ventaja de la teoría de D-S es su habilidad para expresar grados de ignorancia. En la teoría, el compromiso de creencia para con un subconjunto no obliga a la creencia remanente a estar comprometida con su complemento, esto es, $bel(B) + bel(\bar{B}) \leq 1$. La cantidad de creencia comprometida para B y para el complemento de B no implica el grado de ignorancia. Consecuentemente, la teoría provee un marco de trabajo dentro del cual la no creencia puede ser distinguida a partir de la falta de evidencias para la creencia. Las fusiones de creencia bel y de plausibilidad pl tienen las siguientes propiedades:

$$\begin{aligned} bel(B) &\leq pl(B) \\ pl(B) &= 1 - bel(\bar{B}) \end{aligned}$$

donde \bar{B} es la hipótesis complementaria de B : $B \cup \bar{B} = \Theta$ y $B \cap \bar{B} = \phi$.

En el caso de la teoría Bayesiana, la incertidumbre con respecto a un evento es medida por un valor único y la imprecisión con respecto a la medida de incertidumbre se asume que es nula. En el caso de la teoría de D-S (Hégarat-Masclé 1997), el valor de creencia de la hipótesis B puede ser interpretado como el valor de mínima incertidumbre (máxima certidumbre) con respecto a B , y su valor de plausibilidad, que también es el valor de la no creencia de la hipótesis complementaria \bar{B} , puede ser interpretado como el valor de máxima incertidumbre (mínima certidumbre) de B .

5.3.1. Aplicaciones de la teoría de Dempster-Shafer

Las aplicaciones más comunes para la teoría de D-S son en la clasificación y toma de decisiones (Yager 1995). Una situación que manifiesta las características de esta teoría sería la siguiente. Asumiendo que tenemos una rueda de la fortuna dividida en tres colores: rojo, amarillo y verde, al terminar el giro de la rueda el puntero parará en uno de los tres colores (figura 5.2), siendo el anterior un evento aleatorio, la probabilidad depende de la porción del círculo que corresponda a este color.

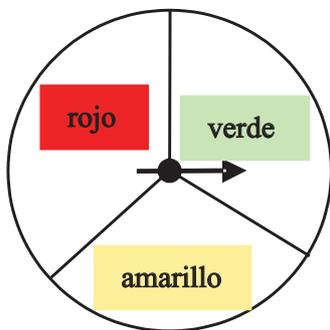


Figura 5.2: Rueda de la fortuna dividida en tres áreas

Asumiremos que tenemos un coche en el que a su chofer le es permitido ir a cierta velocidad, dependiendo del color en el que el puntero se detenga. Si el puntero se detiene en rojo, el chofer puede ir a cualquier velocidad entre 10 y 30 kph, si el puntero se detiene

en amarillo, el chofer podrá ir a cualquier velocidad entre 30 y 50 y si el puntero se detiene en verde, el chofer podrá ir a la velocidad que él desee (0 a 180). En este caso la variable m es la velocidad del coche. Los elementos focales son $A_1 = [10 - 30]$, $A_2 = [30 - 50]$, $A_3 = [0 - 180]$. Si las áreas de los colores fuesen iguales entre si, y la velocidad medida m oscilara ente 40 y 50 sería fácil concluir que el puntero cayo en amarillo, pero si las áreas de los colores son todas diferentes entre si y la velocidad medida m oscila entre 25 y 65, será más difícil apreciar cuál es la solución correcta.

Ejemplo 5.1 *Retomando la aplicación anterior sobre la rueda de la fortuna, se plantea un problema de fusión de información sobre dos fuentes evidenciales, usando la teoría de D-S. El marco de discernimiento Θ , es el conjunto hipotético de la velocidad entre 0 y 180 Km/Hr. La variable “m”, sobre la cual se asume compatibilidad con respecto a las diferentes fuentes evidenciales, es la velocidad. Consideraremos que las fuentes evidenciales de información A_i y B_j han sido probadas con anterioridad y son confiables. Los elementos focales de A_i están formados por diferentes intervalos de velocidad donde: $A_1 = [10 - 30]$, $A_2 = [30 - 50]$, $A_3 = [0 - 180]$. Estos eventos son aleatorios y la asignación de probabilidad básica para cada uno de ellos, dependerá del área que cada uno tenga en la rueda de la fortuna, así, $m(A_1) = 0,5$, $m(A_2) = 0,3$ y $m(A_3) = 0,2$. La otra fuente evidencial de información B_j corresponde a los intervalos de velocidad observados en el coche que dependen de los supuestos por A_i (por lo que deben ser en cierta medida compatibles) de esta manera se consideran los siguientes intervalos de velocidad observados en el coche $B_1 = [0 - 40]$, $B_2 = [20 - 60]$, $B_3 = [35 - 90]$ para los cuales la asignación básica de probabilidad depende de la observación frecuentista más que de las áreas en la rueda de la fortuna, siendo así que, $m(B_1) = 0,6$, $m(B_2) = 0,3$, $m(B_3) = 0,1$. Dado lo anterior podríamos preguntarnos:*

¿En que medida el intervalo de velocidades $C = [70 - 100]$ y $D = [0 - 70]$ pertenecen a la fuente evidencial B_j ?

Determinamos el intervalo de creencia para C

$$bel(C) = \sum_{B_i \subset C}^i m(B_i) = 0$$

$$pl(C) = \sum_{B_i \cap C \neq \emptyset}^i m(B_i) = m(B_3) = 0,1$$

el intervalo de creencia para C se define como el intervalo de probabilidad $Pr(C) = [0, 0,1]$, dado que la credibilidad es cero y que aún la plausibilidad es muy pequeña es poco probable que éste intervalo pertenezca a uno de los elementos focales de B_j . En el ámbito de evaluación y toma de decisiones podemos concluir que el intervalo C es totalmente incoherente con respecto a las fuentes evidenciales consideradas.

Determinamos el intervalo de creencia para D

$$bel(D) = \sum_{B_i \overset{i}{\subset} D} m(B_i) = m(B_1) + m(B_2) = 0,9$$

$$pl(D) = \sum_{B_i \overset{i}{\cap} D \neq \phi} m(B_i) = m(B_1) + m(B_2) + m(B_3) = 1$$

el intervalo de creencia para D , se define como $Pr(D) = [0,9, 1]$, según este resultado podemos concluir que el intervalo de velocidad D es coherente con los elementos focales de B_j .

5.3.2. Combinación de evidencia

Una característica importante en la teoría de D-S es el procedimiento para agregar múltiples estructuras de creencia en la misma variable. Esto puede ser visto como un problema de fusión de información (Yager 1995). Este procedimiento para la múltiple combinación de estructuras de creencia se basa en la regla de agregación de Dempster y es visto como una clase de conjunción (intersección) de las estructuras de creencia.

Asumiendo que m_1 y m_2 son dos estructuras de creencia independientes en el espacio Θ , entonces su conjunción es otra estructura de creencia m , que se denota como $m = m_1 \oplus m_2$. La estructura de creencia m se obtiene de la siguiente manera. Suponiendo que m_1 tiene elementos focales A_i , $i = 1, \dots, n_1$ y suponiendo que m_2 tiene elementos focales B_j , $j = 1, \dots, n_2$. Los elementos focales de m son todos los subconjuntos F_K de Θ donde $F_k = A_i \cap B_j$, y $F_K \neq \phi$. Los pesos asociados con cada F_K son definidos por:

$$m(F_K) = \frac{1}{1 - K} \sum_{A_i \cap B_j = F_K} (m_1(A_i) * m_2(B_j)) \quad (5.6)$$

donde

$$K = \sum_{A_i \cap B_j = \phi} (m_1(A_i) * m_2(B_j)) \quad (5.7)$$

Se puede ver que K ($K \in [0, 1]$) representa la masa que será asignada al conjunto vacío, después de la combinación en la ausencia de normalización (división por $(1 - K)$ en (5.6)), De manera que, K es en ocasiones interpretada como una medida de conflicto entre las diferentes fuentes y se introduce en (5.6) como el factor de normalización. Entre mayor sea K ($0 \leq K \leq 1$) significa que mayor número de fuentes están en conflicto entre si y sus combinaciones tienen menor sentido. Finalmente, la suma ortogonal no existe cuando K es igual a 1. En este caso, se dice que las fuentes son totalmente contradictorias, y no es posible hacer una combinación de ellas.

Ejemplo 5.2 *Se asume que nuestro universo de discurso es $\Theta = \{1, 2, 3, 4, 5, 6\}$*

$\underline{m_1}$		$\underline{m_2}$
$A_1 = \{1, 2, 3\}$	$m(A_1) = 0,5$	$B_1 = \{2, 5, 6\}$ $m(B_1) = 0,6$
$A_2 = \{2, 3, 6\}$	$m(A_2) = 0,3$	$B_2 = \{1, 4\}$ $m(B_2) = 0,4$
$A_3 = \{1, 2, 3, 4, 5, 6\}$	$m(A_3) = 0,2$	

tomando la conjunción se obtiene

$$\begin{aligned} F_1 &= A_1 \cap B_1 = \{2\} \\ F_2 &= A_1 \cap B_2 = \{1\} \\ F_3 &= A_2 \cap B_1 = \{2, 6\} \\ F_4 &= A_3 \cap B_1 = \{2, 5, 6\} \\ F_5 &= A_3 \cap B_2 = \{1, 4\} \end{aligned}$$

notamos también que $A_2 \cap B_2 = \phi$, debido a que sólo una intersección nos da un conjunto nulo (5.7), entonces

$$K = m_1(A_2) * m_2(B_2) = 0,12$$

de aquí que

$$1 - K = 0,88$$

usando esto, calculamos

$$\begin{aligned} m(F_1) &= \left(\frac{1}{0,88} \right) (0,5)(0,6) = 0,341 \\ m(F_2) &= \left(\frac{1}{0,88} \right) (0,5)(0,4) = 0,227 \\ m(F_3) &= \left(\frac{1}{0,88} \right) (0,3)(0,6) = 0,205 \\ m(F_4) &= \left(\frac{1}{0,88} \right) (0,2)(0,6) = 0,136 \\ m(F_5) &= \left(\frac{1}{0,88} \right) (0,2)(0,4) = 0,090 \end{aligned}$$

Aquí vemos que la estructura de creencia resultante de la combinación cumple con la norma $\sum_i m(A_i) = 1$ necesaria de toda estructura.

Ejemplo 5.3 *En este ejemplo de diagnóstico médico (Knight 1991), se muestra el uso de la combinación evidencial como una herramienta para la toma de decisiones. Inicialmente, $\Theta = \{Alergia, Gripe, Resfriado, Neumonía, \dots\dots\}$ donde $\{\Theta\} = 1$, $A:Alergia$, $G:Gripe$, $R:Resfriado$, $N:Neumonía$. Si se adquiere información como fiebre, tenemos una pieza de evidencia m_1 (cada pieza evidencial en este tipo de problemas sólo tiene dos elementos focales).*

$$\begin{aligned} \{G, R, N\} &= 0,6 \\ \Theta &= 0,4 \end{aligned}$$

El remanente de nuestra creencia aún reside en el gran conjunto Θ . Esto no quiere decir que este remanente debe ser asignado al complemento de $\{G, R, N\}$, sino que será asignado al gran conjunto Θ que cuenta con muchos desordenes físicos. El complemento de $\{G, R, N\}$, “podría” ser Alergia, pero en esta etapa del diagnóstico aún no lo sabemos.

Lo sabremos hasta que se complete el total de evidencias (en realidad el complemento de $\{G, R, N\}$ serán otras tres evidencias más, a parte de Alergia).

Supongamos que obtenemos otra fuente de evidencia m_2 al observar fluido en la nariz.

$$\begin{aligned}\{A, G, R\} &= 0,8 \\ \Theta &= 0,2\end{aligned}$$

Se obtiene una nueva estructura de creencia m_3 por medio de la combinación

	$\{A, G, R\} = 0,8$	$\Theta = 0,2$
$\{G, R, N\} = 0,6$	$\{G, R\} = 0,48$	$\{G, R, N\} = 0,12$
$\Theta = 0,4$	$\{A, G, R\} = 0,32$	$\Theta = 0,08$

Las siguientes son las operaciones que generan la tabla anterior:

$$\begin{aligned}\{G, R\} &= \{G, R, N\} \cap \{A, G, R\} = 0,6 * 0,8 = 0,48 \\ \{G, R, N\} &= \{G, R, N\} \cap \{\Theta\} = 0,6 * 0,2 = 0,12 \\ \{A, G, R\} &= \{\Theta\} \cap \{A, G, R\} = 0,4 * 0,8 = 0,32 \\ \{\Theta\} &= \{\Theta\} \cap \{\Theta\} = 0,4 * 0,2 = 0,08\end{aligned}$$

$K = 0$ debido a que no existe conjunto vacío ϕ . Como resultado de aplicar m_1 y m_2 , se produce m_3

$$\begin{aligned}\{G, R\} &= 0,48 \\ \{A, G, R\} &= 0,32 \\ \{G, R, N\} &= 0,12 \\ \{\Theta\} &= 0,08\end{aligned}$$

Ahora implementaremos la creencia m_4 , que se produce por la evidencia de que el problema desaparece cuando el paciente se va de viaje

$$\begin{aligned}\{A\} &= 0,9 \\ \{\Theta\} &= 0,1.\end{aligned}$$

Se puede aplicar el numerador a la regla de combinación para producir (donde ϕ denota el conjunto vacío):

	$\{A\} = 0,9$	$\Theta = 0,1$
$\{G, R\} = 0,48$	$\phi = 0,432$	$\{G, R\} = 0,048$
$\{A, G, R\} = 0,32$	$\{A\} = 0,288$	$\{A, G, R\} = 0,032$
$\{G, R, N\} = 0,12$	$\phi = 0,108$	$\{G, R, N\} = 0,012$
$\Theta = 0,08$	$\{A\} = 0,072$	$\Theta = 0,008$

Ahora una creencia de 0,54 esta asociada con el conjunto vacío ϕ , sólo 0,46 se encuentra asociada con los resultados que son posibles, así que se puede escalar los valores

asociados al conjunto vacío ϕ por el factor $1 - K = 1 - 0,54 = 0,46$, entonces

$$\begin{aligned}\{G, R\} &= \{G, R\}/1 - K = 0,048/0,46 = 0,1043 \\ \{A\} &= (\{A\} + \{A\})/1 - K = 0,36/0,46 = 0,7826 \\ \{A, G, R\} &= \{A, G, R\}/1 - K = 0,32/0,46 = 0,0696 \\ \{G, R, N\} &= \{G, R, N\}/1 - K = 0,012/0,46 = 0,0261 \\ \Theta &= \Theta/1 - K = 0,008/0,46 = 0,0174.\end{aligned}$$

Así

$$\begin{aligned}\{G, R\} &= 0,1043 \\ \{A\} &= 0,7826 \\ \{A, G, R\} &= 0,0696 \\ \{G, R, N\} &= 0,0261 \\ \{\Theta\} &= 0,0174.\end{aligned}$$

Dado estos resultados, podemos concluir casi con toda seguridad que el paciente es alérgico. La desventaja de la combinación evidencial, es que todas las fuentes evidenciales son mutuamente excluyentes, lo que en un principio hace que los resultados de esta técnica sean incoherentes. Sin embargo, esta desventaja desaparece a medida que las tendencias en las evidencias se van haciendo cada vez más claras, debido a que las hipótesis falsas tienden a cero, como se puede observar en el ejemplo anterior.

5.4. Generalización de la teoría de Dempster-Shafer a conjuntos borrosos

Los problemas que se pueden resolver con esta generalización, pueden ser ilustrados con el siguiente sistema de cómputo (Yen 1990). Este sistema, infiere la edad de una persona, basada en varias informaciones acerca de ella. Un sistema como tal, puede contener dos cuerpos de evidencia, uno referente al decoro con el que la persona viste, el otro referente a que si él o ella viste o no, a estilo punk rock. Las reglas para estos dos cuerpos de evidencia, son las siguientes:

SI la persona viste con decoro, ENTONCES es ANCIANA. SI la persona no viste con decoro, ENTONCES su edad es DESCONOCIDA.

SI la persona viste estilo punk rock, ENTONCES es JOVEN. SI la persona no viste estilo punk rock, ENTONCES su edad es DESCONOCIDA.

Donde ANCIANA y JOVEN son subconjuntos borrosos del intervalo $[15, 80]$. Supongamos que una persona llamada Juan exhibe el siguiente juicio probabilístico.

$$\begin{aligned}P(\text{decoro}) &= 0,8 & P(\text{no decoro}) &= 0,2 \\ P(\text{como punk}) &= 0,4 & P(\text{no punk}) &= 0,6\end{aligned}$$

Se le podría pedir al sistema que determine que tan probable es que Juan sea una persona de edad madura.

La característica importante de este problema es que la información que se considera aquí es probabilística y también de naturaleza vaga o borrosa (Anciana, madura, etc.). La teoría de D-S ha probado solucionar el caso especial de este tipo de problemas cuando A_i y B son conjuntos binarios (*crisp*) con su generalización podría incluir conjuntos borrosos.

5.4.1. Función de creencia generalizada

En los siguientes puntos, se muestra la generalización de la función de creencia de la teoría de D-S, a la teoría de conjuntos borrosos, según el trabajo de John Yen (Yen 1990) que hasta la actualidad es el más aceptado.

La optimización para obtener la función de creencia

$Pl(B)$ y $bel(B)$ son los límites de probabilidad inferior y superior del conjunto B bajo las restricciones impuestas por la asignación de probabilidad básica. De aquí que, la función de creencia bel se puede obtener resolviendo el siguiente problema de optimización:

$$LP1 \div \min \sum_{x_i \in B} \sum_j m(x_i : A_j)$$

sujeto al siguiente condicionamiento:

$$m(x_i : A_j) \geq 0, \quad i = 1, \dots, n; \quad j = 1, \dots, 1 \quad (5.8)$$

$$m(x_i : A_j) = 0, \quad \forall x_i \notin A_j \quad (5.9)$$

$$\sum_i m(x_i : A_j) = m(A_j) \quad j = 1, \dots, 1. \quad (5.10)$$

La variable $m(x_i : A_j)$ denota la masa de probabilidad localizada en x_i de la probabilidad básica del elemento focal A_j . La función objetiva simplemente computa la probabilidad total del conjunto B donde la suma interior proporciona la probabilidad de un elemento x_i . El condicionamiento de la desigualdad, especificada en (5.8) establece la no negatividad de las masas de probabilidad. La ecuación (5.9) prohíbe la asignación de probabilidad básica para cualquier elemento focal fuera del focal A_j . La ecuación (5.11) expresa que toda la masa de probabilidad asignada por un focal debe sumarse a su probabilidad básica. Podemos notar de (5.8) y (5.10), que el límite superior en $m(x_i : A_j)$ es $m(A_j)$.

Como la distribución de masa de los elementos focales no interactúa una con otra, puede ser optimizada individualmente, y luego sumar cada una de las distribuciones de masa hasta alcanzar una solución global óptima. La solución del problema lineal sobre la

plausibilidad del conjunto B difiere del anterior sólo en la dirección de optimización, y se formula como sigue:

$$LP2 \div \text{máx} \sum_{x_i \in B} \sum_j m(x_i : A_j) \text{ sujeto a (5.8)-(5.10)}$$

Las soluciones óptimas del subproblema de minimización $LP1$ y el subproblema de maximización $LP2$ se expresan como $m_*(B : A_j)$ y $m^*(B : A_j)$ respectivamente. Sumando las soluciones óptimas de los subproblemas, podemos obtener la medida de creencia y la medida de plausibilidad de B tal como se muestra a continuación.

$$bel(B) = \sum_{A_j \subseteq \Theta} m_*(B : A_j) \quad (5.11)$$

$$pl(B) = \sum_{A_j \subseteq \Theta} m^*(B : A_j) \quad (5.12)$$

Las soluciones óptimas de los subproblemas son las siguientes:

$$m_*(B : A_j) = \begin{cases} m(A_j) & \text{si } A_j \subset B \\ 0 & \text{de otra forma} \end{cases} \quad (5.13)$$

$$m^*(B : A_j) = \begin{cases} m(A_j) & \text{si } A_j \cap B \neq \phi \\ 0 & \text{de otra forma} \end{cases} \quad (5.14)$$

Generalizando funciones objetivas

Las funciones objetivas $LP1$ y $LP2$ computan la probabilidad de un conjunto binario (*crisp*) B . Si B es un subconjunto borroso del marco de discernimiento, su probabilidad se define como

$$P(B) = \sum_{x_i} P(x_i) \times \mu_B(x_i)$$

en la teoría de conjuntos borrosos. Por lo tanto podemos generalizar la función objetiva a

$$\sum_{x_i} \sum_j m(x_i : A_j) \times \mu_B(x_i).$$

Basados en esta generalización de las funciones objetivas, es posible obtener la función de creencia a partir de conjuntos borrosos, para una asignación de probabilidad básica no borrosa.

Suponiendo que A es un elemento focal no borroso. Las masas de probabilidad máxima y mínima que pueden ser localizadas en un conjunto borroso B a partir de A son

$$m_*(B : A) = m(A) \times \inf_{x \in A} \mu_B(x) \quad (5.15)$$

$$m^*(B : A) = m(A) \times \sup_{x \in A} \mu_B(x) \quad (5.16)$$

$m^*(B : A)$ es la solución óptima al siguiente problema de programación lineal:

$$\min \sum_x m(x : A) \times \mu_B(x)$$

sujeto a las restricciones impuestas por (5.8)-(5.10). Una solución óptima para este problema de programación lineal se puede obtener asignando todas las masas de A a un elemento de A que tiene el más bajo grado de pertenencia en B . Esto es, $m_*(B : A) = m(A) \times \inf_{x \in A} \mu_B(x)$. La ecuación (5.16) se puede probar de igual forma.

De las ecuaciones (5.11), (5.12), (5.15), y (5.16) se obtiene la siguiente formula para computar la función de creencia y plausibilidad de conjuntos borrosos a partir de la asignación de probabilidad básica *crisp*:

$$\begin{aligned} bel(B) &= \sum_{A_j \subseteq \Theta} m(A_j) \times \inf_{x \in A_j} \mu_B(x) \\ pl(B) &= \sum_{A_j \subseteq \Theta} m(A_j) \times \sup_{x \in A_j} \mu_B(x) \end{aligned}$$

Representación del condicionamiento probabilístico de los elementos focales borrosos a través de la descomposición

Para trabajar con elementos focales borrosos, primero se descomponen en elementos focales no borrosos cuyo condicionamiento fue descrito en la subsección “*La optimización para obtener la función de creencia*”. Un elemento focal borroso tiene dos componentes: un subconjunto borroso del marco de discernimiento y la asignación de probabilidad básica al subconjunto.

Un conjunto de nivel- α de A , un subconjunto de Θ , es un conjunto *crisp* denotado por A_α que contiene todos los elementos de Θ cuyo grado de pertenencia en A es más grande que o igual a α :

$$A_\alpha = \{x | \mu_A(x) \geq \alpha\}$$

Un conjunto borroso A puede descomponerse en sus conjuntos de niveles α a través de la resolución identidad:

$$A = \sum_{\alpha} \alpha A_\alpha$$

donde la sumatoria denota la operación de unión de conjuntos y denota un conjunto borroso con una función de pertenencia bi-valuada definida por

$$\begin{aligned} \mu_{\alpha A_\alpha}(x) &= \alpha \quad \text{para } x \in A_\alpha \\ \mu_{\alpha A_\alpha}(x) &= 0 \quad \text{de otra forma} \end{aligned}$$

Con el fin de descomponer un elemento focal borroso, es necesario descomponer la probabilidad básica del elemento focal y distribuirla entre los conjuntos de niveles α de los elementos focales. La descomposición tiene que satisfacer dos condiciones:

- a) Las probabilidades básicas descompuestas deben sumarse a la probabilidad básica asignada al elemento focal borroso.

$$\sum_{\alpha} m(A_{\alpha}) = m(A)$$

- b) Las probabilidades básicas descompuestas no deben ser negativas.

$$m(A_{\alpha}) \geq 0$$

Dubois y Prade (Dubois 1982) han demostrado que si la asignación de probabilidad básica es un conjunto de elementos focales anidados, $A_1 \supset A_2 \cdots \supset A_n$, que pueden ser relacionados con la distribución de posibilidad inducida, denotada como $Poss(x)$, como sigue:

$$m(A_i) = \pi_i - \pi_{i-1} \quad (5.17)$$

donde $\pi_i = \inf_{x \in A_i} poss(x)$, $\pi_0 = 0$, $\pi_n = 1$. Este resultado se puede aplicar directamente a la descomposición de elementos focales borrosos cuyo valor de probabilidad básica es uno ($m(A) = 1$) debido a los conjuntos de nivel $-\alpha$ de A derivados de un conjunto de elementos focales anidados. Desde que $\pi_i = \inf_{x \in A_i} poss(x) = \alpha_i$, el π_i en (5.17) se convierte en el valor alfa α_i de los conjuntos de nivel. De manera que se obtiene

$$m(A_{\alpha_i}) = \alpha_i - \alpha_{i-1} \quad (5.18)$$

Se extiende esta idea para descomponer elementos focales borrosos con masa de probabilidad arbitraria ($0 \leq m(A) \leq 1$) al multiplicar la masa del elemento focal con el lado derecho de la ecuación (5.18). La descomposición de un elemento focal borroso A se define como: una colección de subconjuntos no-borrosos tal que:

- a) Son los conjuntos de niveles de A que forman la resolución identidad
- b) Sus probabilidades básicas son

$$m(A_{\alpha_i}) = (\alpha_i - \alpha_{i-1}) \times m(A) \quad i = 1, 2, \dots, n$$

donde $\alpha_0 = 0$ y $\alpha_n = 1$. Cuando el elemento focal es un conjunto *crisp*, su descomposición es el elemento focal en si mismo debido a que la descomposición contiene un conjunto de un solo nivel, que corresponde al grado de la función de pertenencia “uno”.

El condicionamiento probabilístico de un focal borroso se define para que éste sea hallado a partir de su descomposición, el cual es un conjunto de focales no borrosos. Como ahora ya sabemos como trabajar con focales no borrosos, descomponiendo un focal borroso a un focal no-borroso podemos calcular la función de creencia que estará condicionada por los focales borrosos. Esto es la masa de probabilidad con la que un focal borroso A contribuye para formar la función de creencia (que incluye la plausibilidad) de un subconjunto borroso B , es la contribución de la descomposición total de A para la creencia (y plausibilidad) de B .

$$m^*(B : A) = \sum_{\alpha} m^*(B : A_{\alpha})$$

$$m_*(B : A) = \sum_{\alpha} m_*(B : A_{\alpha})$$

Uso de de la función de creencia

Basados en la generalización de la función objetiva y expresando el condicionamiento probabilístico de los elementos focales borrosos, a través de su descomposición, es posible derivar la siguiente formula para el computo de la función de creencia y la función de plausibilidad.

$$\begin{aligned} bel(B) &= \sum_A m(A) \sum_{\alpha_i} [\alpha_i - \alpha_{i-1}] \times \inf_{x \in A_{\alpha_i}} \mu_B(x) \\ pl(B) &= \sum_A m(A) \sum_{\alpha_i} [\alpha_i - \alpha_{i-1}] \times \sup_{x \in A_{\alpha_i}} \mu_B(x) \end{aligned} \quad (5.19)$$

Ejemplo 5.4 *Este ejemplo ilustra el uso de las formulas anteriores (Caro 1999). Supongamos que el marco de discernimiento es el conjunto de los números enteros entre 1 y 10. Una asignación de probabilidad básica borrosa que consiste de dos elementos focales “A” y “C”:*

$$\begin{aligned} A &= \{0,25/1, 0,5/2, 0,75/3, 1/4, 1/5, 0,75/6, 0,5/7, 0,25/8\} \\ C &= \{0,5/5, 1/6, 0,8/7, 0,4/8\}. \end{aligned}$$

Donde cada miembro de la lista esta en la forma de $\mu_A(x_i)/x_i$. Como se puede ver en la figura 5.3, “A” cuenta con cuatro niveles de corte α_i , (se describe uno por cada inciso de la figura). Se busca encontrar el grado de creencia y el grado de plausibilidad del subconjunto borroso “B”, también descrito en la figura 5.3:

$$B = \{0,5/2, 1/3, 1/4, 1/5, 0,9/6, 0,6/7, 0,3/8\}.$$

La descomposición del focal borroso “A” consiste de cuatro focales no-borrosos:

$$\begin{aligned} A_{0,25} &= \{1, 2, \dots, 8\} \text{ con masa de } 0,25 \times m(A) \\ A_{0,5} &= \{2, 3, \dots, 7\} \text{ con masa de } 0,25 \times m(A) \\ A_{0,75} &= \{3, 4, \dots, 6\} \text{ con masa de } 0,25 \times m(A) \\ A_1 &= \{4, 5\} \text{ con masa de } 0,25 \times m(A) \end{aligned}$$

la descomposición del focal borroso “C” también consiste de cuatro focales no-borrosos:

$$\begin{aligned} C_{0,4} &= \{5, 6, 7, 8\} \text{ con masa de } 0,4 \times m(C) \\ C_{0,5} &= \{5, 6, 7\} \text{ con masa de } 0,1 \times m(C) \\ C_{0,8} &= \{6, 7\} \text{ con masa de } 0,3 \times m(C) \\ C_1 &= \{6\} \text{ con masa de } 0,2 \times m(C) \end{aligned}$$

y la descomposición del focal borroso “B” consiste de cinco focales no borrosos:

$$\begin{aligned} B_{0,3} &= \{2, 3, \dots, 8\} \text{ con masa de } 0,3 \times m(B) \\ B_{0,5} &= \{2, 3, \dots, 7\} \text{ con masa de } 0,2 \times m(B) \\ B_{0,6} &= \{3, 4, \dots, 7\} \text{ con masa de } 0,1 \times m(B) \\ B_{0,9} &= \{3, 4, 5, 6\} \text{ con masa de } 0,3 \times m(B) \\ B_1 &= \{3, 4, 5\} \text{ con masa de } 0,1 \times m(B) \end{aligned}$$

La función de pertenencia $\mu_B(x)$ en (5.19) puede recobrase a partir de los niveles corte como sigue:

$$\mu(x) = \sup\{\alpha : x \in F_\alpha\} \quad (5.20)$$

Donde F_α es llamada “continuidad superior” (Dubois 2000a) y se define como:

$$F_\alpha = \lim_{n \rightarrow \infty} \bigcap_{i=1,n} F_{\alpha_i}$$

Primeramente recobramos $\mu_B(x)$ por el uso de la ecuación (5.20) donde la única condición según (5.19) es que $x \in A_{\alpha_i}$. En la figura 5.3 se ilustra como se obtienen límites superiores e inferiores de $\mu_B(x)$ para cada A_{α_i} necesarios para obtener la credibilidad y la plausibilidad (5.19). En inciso a) para A_{α_1} , $x = \{1, 2, \dots, 8\}$ donde aplicando (5.20) obtenemos la función de pertenencia borrosa $\mu_B(x)$ (línea discontinua en inciso a) de la figura 5.3) que produce $\inf_{x \in A_{\alpha_1}} \mu_B(x) = 0$, y $\sup_{x \in A_{\alpha_1}} \mu_B(x) = 1$, según se observa. En b) para A_{α_2} , $x = \{2, 3, \dots, 7\}$ de donde resulta $\inf_{x \in A_{\alpha_2}} \mu_B(x) = 0,5$ y $\sup_{x \in A_{\alpha_2}} \mu_B(x) = 1$. En c) para A_{α_3} , $x = \{3, 4, 5, 6\}$ que produce $\inf_{x \in A_{\alpha_3}} \mu_B(x) = 0,9$ y $\sup_{x \in A_{\alpha_3}} \mu_B(x) = 1$, y en d) para A_{α_4} , $x = \{4, 5\}$ que produce $\inf_{x \in A_{\alpha_4}} \mu_B(x) = 1$ y $\sup_{x \in A_{\alpha_4}} \mu_B(x) = 1$. Para obtener la credibilidad del B con respecto al A la expresión $\inf_{x \in A_{\alpha_i}} \mu_B(x)$ es denotada como $f_{B,A}(\alpha_i)$ de manera que se pueda tener

$$\begin{aligned} m_*(B : A) &= m(A) \times [0,25 \times f_{B,A}(0,25) + 0,25 \times f_{B,A}(0,5) + 0,25 \times f_{B,A}(0,75) + \\ &\quad 0,25 \times f_{B,A}(1)] \\ m_*(B : A) &= m(A) \times [0,25 \times 0 + 0,25 \times 0,5 + 0,25 \times 0,9 + 0,25 \times 1] \\ m_*(B : A) &= 0,6 \times m(A) \end{aligned}$$

La credibilidad de B con respecto a C se obtiene de la misma forma. Para analizar como se obtiene $f_{B,C}(\alpha_i)$ podemos auxiliarnos de la figura 5.4.

$$\begin{aligned} m_*(B : C) &= m(C) \times [0,4 \times f_{B,C}(0,4) + 0,1 \times f_{B,C}(0,5) + 0,3 \times f_{B,C}(0,8) + \\ &\quad 0,2 \times f_{B,C}(1)] \\ m_*(B : C) &= m(C) \times [0,4 \times 0,3 + 0,1 \times 0,6 + 0,3 \times 0,6 + 0,2 \times 0,9] \\ m_*(B : C) &= 0,54 \times m(C) \end{aligned}$$

de lo anterior tenemos que

$$bel(B) = 0,6m(A) + 0,54m(C)$$

Para obtener la plausibilidad de B se calculan $m^*(B : A)$ y $m^*(B : C)$. Los valores $\sup_{x \in A_{\alpha_i}} \mu_B(x)$ respectivos se obtienen según se vio al recobrar $\mu_B(x)$ (ver figura 5.3 y figura 5.4). De esta manera la plausibilidad de B sería

$$pl(B) = m(A) + 0,95m(C)$$

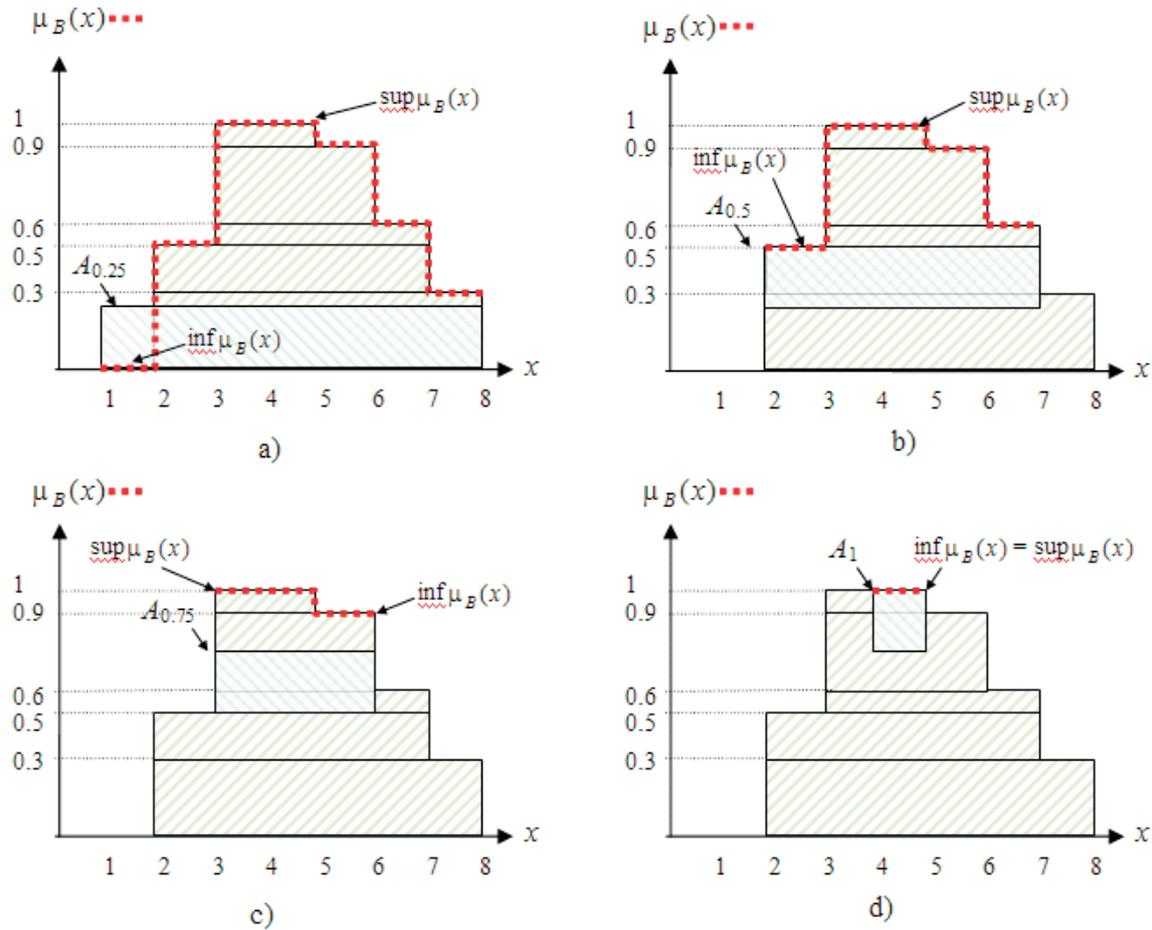


Figura 5.3: Descripción gráfica de cómo se obtiene $f_{B,A}(\alpha_i)$. Obtención de: a) $f_{B,A}(0,25)$, b) $f_{B,A}(0,5)$, c) $f_{B,A}(0,75)$, d) $f_{B,A}(1)$

5.4.2. Regla de combinación generalizada

La regla de D-S combina los efectos de dos fuentes evidenciales independientes, es posible la combinación de relaciones compatibles borrosas al asumir la no interacción entre las variables, lo cual es un concepto análogo a la independencia entre las variables aleatorias en la teoría de la probabilidad. El proceso de combinar relaciones compatibles borrosas produce elementos focales borrosos subnormales de aquí la importancia de describir como se normalizan los elementos focales borrosos subnormales (Yen 1990).

Normalización de los elementos focales borrosos

Una asunción importante en este trabajo es que todos los elementos focales son normales. Se evitan los elementos focales borrosos porque estos asignan masa de probabilidad al conjunto vacío. Sin embargo la intersección de dos conjuntos borrosos normales produce un conjunto borroso subnormal. De aquí que, la combinación de la asignación de probabilidad básica borrosa debe tratar con la normalización de elementos focales borrosos subnormales. Para lograr esto es necesario normalizar los dos componentes de un

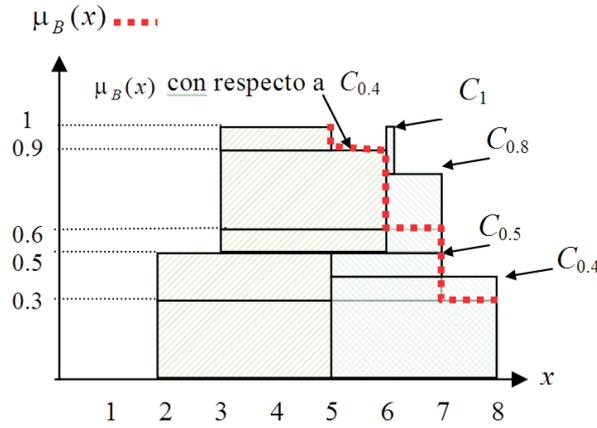


Figura 5.4: Descripción gráfica de los elementos focales de B y C

elemento focal borroso: El focal en si mismo, el cual es un conjunto borroso subnormal, y la masa de probabilidad asignada al focal.

Suponiendo que A es un conjunto borroso subnormal caracterizado por la función de pertenencia $\mu_A(x)$. El conjunto normalizado de A es denotado como \bar{A} , y se caracteriza por la siguiente función de pertenencia.

$$\mu_{\bar{A}}(x) = \frac{\mu_A(x)}{\max_x \mu_A(x)} = k \times \mu_A(x)$$

donde k es el factor de normalización

$$k = 1 / \max_x \mu_A(x)$$

El criterio para normalizar la masa de probabilidad de un focal subnormal es que el condicionamiento probabilístico impuesto por el focal subnormal debe preservarse después de la normalización. Desde que se usa la descomposición de un focal para representar su condicionamiento probabilístico, esto significa que la masa de probabilidad asignada a un focal descompuesto no debe ser cambiada por el proceso de normalización. Debido a que el nivel de corte α_i del focal subnormal se convierte en el nivel corte $k\alpha_i$ del focal normalizado, la masa de probabilidad asignada para ellos debe ser la misma:

$$m(A_{\alpha_i}) = m(\bar{A}_{k\alpha_i}) \quad (5.21)$$

A partir de esta condición, es posible derivar la relación entre $m(\bar{A})$ y $m(A)$ como sigue. El lado izquierdo de la ecuación (5.21) se puede escribir como

$$m(A_{\alpha_i}) = m(A)(\alpha_i - \alpha_{i-1}).$$

El lado derecho de (5.21) puede describirse como

$$m(\bar{A}_{k\alpha_i}) = m(\bar{A})(k\alpha_i - k\alpha_{i-1}) = k \times m(\bar{A})(\alpha_i - \alpha_{i-1}).$$

Podemos entender a partir de las tres ecuaciones anteriores que la masa del focal normalizado es reducida por un factor recíproco a la proporción por la cual su función de pertenencia es escalada.

$$m(\bar{A}) = m(A)/k$$

La masa remanente $(1 - 1/k)m(A)$ es la cantidad asignada al conjunto vacío por el focal borroso subnormal, y por lo tanto debe ser parte del factor de normalización en la regla de generalización de D-S.

Regla generalizada de combinación

La conmutatividad es un importante requerimiento para cualquier regla de combinación de evidencias, porque es muy recomendable tener los efectos de agregación de evidencias independientes del orden de combinación. La regla de Dempster es conmutativa. El paso de normalización discutido en la sección anterior no es conmutativo porque este modifica la función de pertenencia de la intersección subnormal del elemento focal. Para resolver este problema, primero se demuestra que el proceso de normalización en la regla de Dempster se puede posponer sin cambio en los resultados de la combinación. Después, se describe la regla de combinación generalizada propuesta en (Yen 1990) donde el proceso de normalización es pospuesto con el fin de lograr conmutatividad.

En la siguiente combinación, se usa el símbolo \otimes para denotar la regla de Dempster sin normalización ($K = 0$ en (5.6)), la letra “N” para denotar el proceso de normalización, y la letra m' con apóstrofe para denotar la asignación de probabilidad básica sin normalizar. Para demostrar que se obtiene los mismos resultados, se considera el marco de discernimiento de tres asignaciones básicas de probabilidad: m_1 , m_2 , y m_3 . Se busca demostrar que aplicando la normalización después que las tres asignaciones de probabilidad básicas son combinadas sin normalización producen el mismo resultado que al usar la regla de combinación de Dempster en la forma convencional,

$$(m_1 \oplus m_2) \oplus m_3 = N[(m_1 \otimes m_2) \otimes m_3] \quad (5.22)$$

Primero se expande los resultados de combinar las dos primeras asignaciones de probabilidad básica usando la regla de Dempster.

$$m_1 \oplus m_2(C) = \frac{m'_{12}(C)}{1 - k_{12}}$$

donde

$$m'_{12}(C) = \sum_{A \cap B = C} m_1(A)m_2(B) \quad (5.23)$$

y

$$k_{12} = \sum_{A \cap B = \phi} m_1(A)m_2(B) \quad (5.24)$$

El lado izquierdo de (5.22) se convierte en

$$(m_1 \oplus m_2) \oplus m_3(E) = \frac{\frac{1}{1-k_{12}} \sum_{C \cap D = E} m'_{12}(C)m_3(D)}{1 - \frac{1}{1-k_{12}} \sum_{C \cap D = \phi} m'_{12}(C)m_3(D)}$$

substituyendo $m'_{12}(C)$ con el lado derecho de (5.23), se obtiene

$$(m_1 \oplus m_2) \oplus m_3(E) = \frac{\frac{1}{1-k_{12}} \sum_{A \cap B \cap D = E} m_1(A)m_2(B)m_3(D)}{1 - \frac{1}{1-k_{12}} \sum_{A \cap B \cap D = \phi} m_1(A)m_2(B)m_3(D)}$$

multiplicando ambos el numerador y el denominador por $1 - k_{12}$, se tiene

$$(m_1 \oplus m_2) \oplus m_3(E) = \frac{\sum_{A \cap B \cap D = E} m_1(A)m_2(B)m_3(D)}{1 - k_{12} - \sum_{A \cap B \cap D = \phi, A \cap B \neq \phi} m_1(A)m_2(B)m_3(D)}$$

substituyendo k_{12} con el lado derecho de la ecuación (5.24), se obtiene

$$(m_1 \oplus m_2) \oplus m_3(E) = \frac{\sum_{A \cap B \cap D = E} m_1(A)m_2(B)m_3(D)}{1 - \left[\sum_{a \cap B \neq \phi} m_1(A)m_2(B) + \sum_{A \cap B \cap D = \phi, A \cap B \neq \phi} m_1(A)m_2(B)m_3(D) \right]}.$$

Debido a que $\sum_D m_3(D) = 1$, podemos reformular el factor de normalización:

$$(m_1 \oplus m_2) \oplus m_3(E) = \frac{\sum_{A \cap B \cap D = E} m_1(A)m_2(B)m_3(D)}{1 - \left[\sum_{a \cap B \neq \phi, A \cap B = \phi} m_1(A)m_2(B)m_3(D) + \sum_{A \cap B \cap D = \phi, A \cap B \neq \phi} m_1(A)m_2(B)m_3(D) \right]}$$

finalmente se obtiene

$$(m_1 \oplus m_2) \oplus m_3(E) = \frac{\sum_{A \cap B \cap D = E} m_1(A)m_2(B)m_3(D)}{1 - \sum_{A \cap B \cap D = \phi} m_1(A)m_2(B)m_3(D)}$$

$$(m_1 \oplus m_2) \oplus m_3(E) = N[(m_1 \otimes m_2) \otimes m_3(D)].$$

Por lo tanto, con esto se demuestra que el paso de normalización de la regla de Dempster se puede posponer sin cambiar el resultado de la combinación.

La regla de combinación generalizada presentada aquí consiste de dos operaciones: Una operación de producto cruz y un proceso de normalización. Primero es combinada la asignación de probabilidad básica borrosa por medio del siguiente producto cruz generalizado:

$$m'_{12}(C) = m_1 \otimes m_2(C) = \sum_{A \cap B = C} m_1(A)m_2(B) \quad (5.25)$$

Donde \cap denotadas por el operador de intersección borrosa y C es una intersección no normalizada de un elemento focal, el cual puede ser un subconjunto borroso subnormal del marco de discernimiento. El conjunto borroso es una clase especial de elemento focal subnormal. Para computar la asignación de probabilidad básica combinada y normalizada (para la obtención de su función de creencia) se aplica el siguiente proceso de normalización para la asignación de probabilidad básica combinada no normalizada:

$$N[m'](D) = \frac{\sum_{\bar{C} = D} \max_{x_i} \mu_C(x_i) m'(C)}{1 - \sum_{C \in \Theta} (1 - \max_{x_i} \mu_C(x_i)) m'(C)}.$$

Por ejemplo, si se necesita combinar tres asignaciones de probabilidad básica del marco de discernimiento Θ , el resultado de la combinación es calculado primero combinando las tres asignaciones de probabilidad básica sin la normalización usando (5.25), y después normalizando el resultado final:

$$m_1 \oplus m_2 \oplus m_3 = N[(m_1 \otimes m_2) \otimes m_3]$$

El producto cruz generalizado es conmutativo y al posponer el proceso de normalización podemos combinar la asignación de probabilidad básica borrosa independientemente del orden en que se presente. En el caso especial donde existen sólo dos asignaciones de probabilidad básica borrosa, la combinación usando la regla de Dempster generalizada sería:

$$m_1 \oplus m_2(C) = N[m_1 \otimes m_2](C) = \frac{\sum_{(A \cap B)=C} \max_{x_i} \mu_{A \cap B}(x_i) m_1(A) m_2(B)}{1 - \sum_{A, B} (1 - \max_{x_i} \mu_{A \cap B}(x_i)) m_1(A) m_2(B)} \quad (5.26)$$

Uso de la regla de combinación generalizada de D-S

Retomando el ejemplo de introducción de la sección 5.4 se muestra el uso de la regla de combinación generalizada de D-S y aprovechando lo completo del ejemplo se obtiene también la función de creencia.

Ejemplo 5.5 *Se tiene un sistema, que infiere la edad de una persona, basada en varias informaciones acerca de ella. Un sistema como tal, puede contener dos cuerpos de evidencia, uno referente al decoro con el que la persona viste, el otro referente a que si él o ella viste o no, a estilo punk rock. Las reglas para estos dos cuerpos de evidencia, son las siguientes:*

SI la persona viste con decoro, ENTONCES es ANCIANA. SI la persona no viste con decoro, ENTONCES su edad es DESCONOCIDA.

SI la persona viste estilo punk rock, ENTONCES es JOVEN. SI la persona no viste estilo punk rock, ENTONCES su edad es DESCONOCIDA.

Donde ANCIANA y JOVEN son subconjuntos borrosos del intervalo [15, 80]. La edad DESCONOCIDA se considera entre [15, 80] que es un conjunto referencial considerado elemento focal.

Ahora supongamos la existencia de dos fuentes evidenciales de información que describen a una persona X. Una de estas fuentes es probabilística y la otra borrosa. La probabilística es la asignación básica de probabilidad que alguien da a X descrita como:

$$M(JOVEN) = 0, \quad m(MADURA) = 0,2, \quad m(ANCIANA) = 0,8.$$

El conjunto borroso es la opinión de otra persona que estima que al menos la persona X tiene 45 años pero que esta completamente segura que tiene 55 o más. Los elementos focales que describen al conjunto borroso son los siguientes:

$$B = \{0/42,5, 0,2/45, 0,4/47,5, 0,6/50, 0,8/52,5, 1/55, 1/57,5, 1/60, 1/62,5, 1/65, 1/67,5, 1/70, 1/72,5, 1/75, 1/77,5, 1/80\}.$$

Lo interesante en este caso sería obtener la función de creencia de B y de esa manera saber que tan creíble o plausible resulta la opinión de B con respecto a la edad de X, y en base a esto poder decidir quien es X de acuerdo al conocimiento a priori que se

tenga. Los elementos focales que describen cuando una persona es JOVEN, MADURA, o ANCIANA se describen a continuación:

$$JOVEN = \{1/15, 1/17,5, 1/20, 1/22,5, 1/25, 0,9/27,5, 0,8/30, 0,7/32,5, 0,6/35, 0,5/37,5, 0,4/40, 0,3/42,5, 0,2/45, 0,1/47,5, 0/50\}$$

$$MADURA = \{0/25, 0,25/27,5, 0,5/30, 0,75/32,5, 1/35, 1/37,5, 1/40, 1/42,5, 1/45, 1/47,5, 1/50, 0,75/52,5, 0,5/55, 0,25/57,5, 0/60\}$$

$$ANCIANA = \{0/35, 0,1/37,5, 0,2/40, 0,3/42,5, 0,4/45, 0,5/47,5, 0,6/50, 0,7/52,5, 0,8/55, 0,9/57,5, 1/60, 1/62,5, 1/65, 1/67,5, 1/70, 1/72,5, 1/75, 1/77,5, 1/80\}$$

La figura 5.5 describe los subconjuntos borrosos que describen las edades usadas en este ejemplo.

La descomposición del elemento focal JOVEN consiste de diez elementos focales no borrosos

$$JOVEN_{0,1} = \{[15, 47,5]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

$$JOVEN_{0,2} = \{[15, 45]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

$$JOVEN_{0,3} = \{[15, 42,5]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

$$JOVEN_{0,4} = \{[15, 40]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

$$JOVEN_{0,5} = \{[15, 37,5]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

$$JOVEN_{0,6} = \{[15, 35]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

$$JOVEN_{0,7} = \{[15, 32,5]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

$$JOVEN_{0,8} = \{[15, 30]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

$$JOVEN_{0,9} = \{[15, 27,5]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

$$JOVEN_1 = \{[15, 25]\} \text{ con masa de } 0,1 \times m(JOVEN)$$

La descomposición del elemento focal MADURA consiste de cuatro elementos focales no borrosos

$$MADURA_{0,25} = \{[27,5, 57,5]\} \text{ con masa de } 0,25 \times m(MADURA)$$

$$MADURA_{0,5} = \{[30, 55]\} \text{ con masa de } 0,25 \times m(MADURA)$$

$$MADURA_{0,75} = \{[32,5, 52,5]\} \text{ con masa de } 0,25 \times m(MADURA)$$

$$MADURA_1 = \{[35, 50]\} \text{ con masa de } 0,25 \times m(MADURA)$$

La descomposición del elemento focal ANCIANA consiste de diez elementos focales no

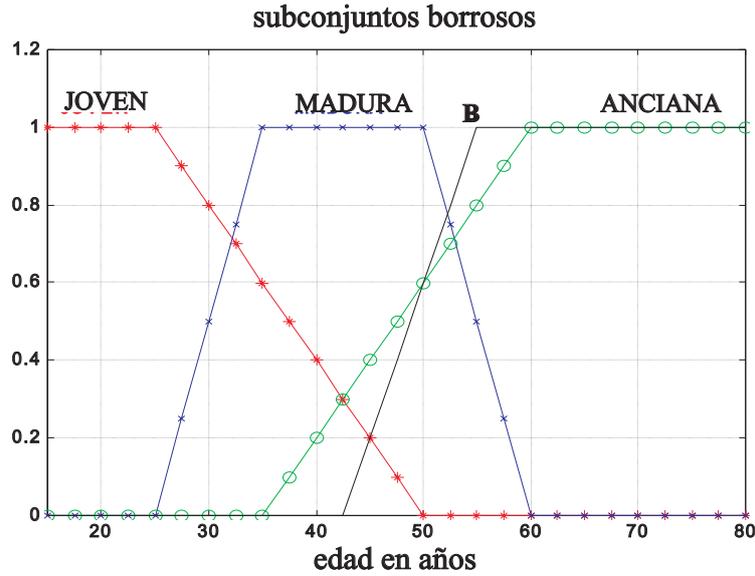


Figura 5.5: Descripción de los subconjuntos borrosos

borrosos

$ANCIANA_{0,1} = \{[37,5, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

$ANCIANA_{0,2} = \{[40, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

$ANCIANA_{0,3} = \{[42,5, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

$ANCIANA_{0,4} = \{[45, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

$ANCIANA_{0,5} = \{[47,5, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

$ANCIANA_{0,6} = \{[50, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

$ANCIANA_{0,7} = \{[52,5, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

$ANCIANA_{0,8} = \{[55, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

$ANCIANA_{0,9} = \{[57,5, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

$ANCIANA_1 = \{[60, 80]\}$ con masa de $0,1 \times m(ANCIANA)$

Credibilidad de B

Denotamos $\inf_{x \in JOVEN_{\alpha_i}} \mu_B(x)$ como $f_{B,JOVEN}(\alpha_i)$, se tiene entonces $m_*(B : JOVEN)$

$$m_*(B : JOVEN) = m(JOVEN) \times [0,1 \times f_{B,JOVEN}(0,1) + 0,1 \times f_{B,JOVEN}(0,2) + 0,1 \times f_{B,JOVEN}(0,3) + 0,1 \times f_{B,JOVEN}(0,4) + 0,1 \times f_{B,JOVEN}(0,5) + 0,1 \times f_{B,JOVEN}(0,6) + 0,1 \times f_{B,JOVEN}(0,7) + 0,1 \times f_{B,JOVEN}(0,8) + 0,1 \times f_{B,JOVEN}(0,9) + 0,1 \times f_{B,JOVEN}(1)]$$

Como los elementos focales de $JOVEN$ no son una inclusión de los elementos focales B entonces $f_{B,JOVEN}(\alpha_i) = 0$ para $i = 1, 2, \dots, n$, por lo tanto

$$m_*(B : JOVEN) = 0$$

Denotamos $\inf_{x \in MADURA_{\alpha_i}} \mu_B(x)$ como $f_{B,MADURA}(\alpha_i)$, se tiene entonces $m_*(B : MADURA)$

$$m_*(B : MADURA) = m(MADURA) \times [0,25 \times f_{B,MADURA}(0,25) + 0,25 \times f_{B,MADURA}(0,5) + 0,25 \times f_{B,MADURA}(0,75) + 0,25 \times f_{B,MADURA}(1)]$$

Como los elementos focales de *MADURA* no son una inclusión de los elementos focales *B* entonces $f_{B,MADURA}(\alpha_i) = 0$ para $i = 1, 2, \dots, n$, por lo tanto

$$m_*(B : MADURA) = 0$$

Denotamos $\inf_{x \in ANCIANA_{\alpha_i}} \mu_B(x)$ como $f_{B,ANCIANA}(\alpha_i)$, se tiene entonces $m_*(B : ANCIANA)$

$$m_*(B : ANCIANA) = m(ANCIANA) \times [0,1 \times f_{B,ANCIANA}(0,1) + 0,1 \times f_{B,ANCIANA}(0,2) + 0,1 \times f_{B,ANCIANA}(0,3) + 0,1 \times f_{B,ANCIANA}(0,4) + 0,1 \times f_{B,ANCIANA}(0,5) + 0,1 \times f_{B,ANCIANA}(0,6) + 0,1 \times f_{B,ANCIANA}(0,7) + 0,1 \times f_{B,ANCIANA}(0,8) + 0,1 \times f_{B,ANCIANA}(0,9) + 0,1 \times f_{B,ANCIANA}(1)]$$

$$m_*(B : ANCIANA) = m(ANCIANA) \times [0,1 \times 0 + 0,1 \times 0 + 0,1 \times 0 + 0,1 \times 0,2 + 0,1 \times 0,4 + 0,1 \times 0,6 + 0,1 \times 0,8 + 0,1 \times 1 + 0,1 \times 1 + 0,1 \times 1]$$

$$m_*(B : ANCIANA) = 0,5 \times m(ANCIANA)$$

$$m_*(B : ANCIANA) = 0,4$$

por lo tanto

$$bel(B) = 0,4$$

Plausibilidad de *B*

Denotamos $\sup_{x \in JOVEN_{\alpha_i}} \mu_B(x)$ como $f_{B,JOVEN}(\alpha_i)$, se tiene entonces $m^*(B : JOVEN)$

$$m^*(B : JOVEN) = m(JOVEN) \times [0,1 \times f_{B,JOVEN}(0,1) + 0,1 \times f_{B,JOVEN}(0,2) + 0,1 \times f_{B,JOVEN}(0,3) + 0,1 \times f_{B,JOVEN}(0,4) + 0,1 \times f_{B,JOVEN}(0,5) + 0,1 \times f_{B,JOVEN}(0,6) + 0,1 \times f_{B,JOVEN}(0,7) + 0,1 \times f_{B,JOVEN}(0,8) + 0,1 \times f_{B,JOVEN}(0,9) + 0,1 \times f_{B,JOVEN}(1)]$$

$$m^*(B : JOVEN) = m(JOVEN) \times [0,1 \times 0,4 + 0,1 \times 0,2 + 0,1 \times 0 + 0,1 \times 0]$$

$$m^*(B : JOVEN) = 0,06 \times m(JOVEN)$$

como $m(JOVEN) = 0$, entonces

$$m^*(B : JOVEN) = 0$$

Denotamos $\sup_{x \in MADURA_{\alpha_i}} \mu_B(x)$ como $f_{B,MADURA}(\alpha_i)$, se tiene entonces $m^*(B : MADURA)$

$$m^*(B : MADURA) = m(MADURA) \times [0,25 \times f_{B:MADURA}(0,25) + 0,25 \times f_{B:MADURA}(0,5) + 0,25 \times f_{B:MADURA}(0,75) + 0,25 \times f_{B:MADURA}(1)]$$

$$m^*(B : MADURA) = m(MADURA) \times [0,25 \times 1 + 0,25 \times 1 + 0,25 \times 0,8 + 0,25 \times 0,6]$$

$$m^*(B : MADURA) = 0,85 \times m(MADURA)$$

$$m^*(B : MADURA) = 0,17$$

Denotamos $\sup_{x \in ANCIANA_{\alpha_i}} \mu_B(x)$ como $f_{B,ANCIANA}(\alpha_i)$, se tiene entonces $m^*(B : ANCIANA)$

$$m^*(B : ANCIANA) = m(ANCIANA) \times [0,1 \times f_{B,ANCIANA}(0,1) + 0,1 \times f_{B,ANCIANA}(0,2) + 0,1 \times f_{B,ANCIANA}(0,3) + 0,1 \times f_{B,ANCIANA}(0,4) + 0,1 \times f_{B,ANCIANA}(0,5) + 0,1 \times f_{B,ANCIANA}(0,6) + 0,1 \times f_{B,ANCIANA}(0,7) + 0,1 \times f_{B,ANCIANA}(0,8) + 0,1 \times f_{B,ANCIANA}(0,9) + 0,1 \times f_{B,ANCIANA}(1)]$$

$$m^*(B : ANCIANA) = m(ANCIANA) \times [0,1 \times 1 + 0,1 \times 1]$$

$$m^*(B : ANCIANA) = 1 \times m(ANCIANA)$$

$$m^*(B : ANCIANA) = 0,8$$

por lo tanto

$$pl(B) = m^*(B : MADURA) + m^*(B : ANCIANA) = 0,17 + 0,8$$

$$pl(B) = 0,97$$

Función de creencia de B

$$\Pr(B) = [bel(B), pl(B)]$$

$$\Pr(B) = [0,4, 0,97]$$

Con los valores de la función de creencia de la opinión B (en ocasiones es útil usar su media) podríamos decidir que tan creíble es la opinión de B con respecto a la edad de

X y en base a este dato junto con otras apreciaciones podríamos inferir a que persona X se están refiriendo las fuentes evidenciales de información.

Ejemplo 5.6 Retomando el ejemplo 5.5 donde a través de unas reglas de inferencia se deduce la edad de una persona X , donde ANCIANA y JOVEN son subconjuntos borrosos del intervalo $[15, 80]$. La edad DESCONOCIDA se considera entre $[15, 80]$ que es un conjunto referencial considerado elemento focal. Supongamos la existencia de dos fuentes evidenciales probabilísticas que describen a una persona X . Estos dos cuerpos de evidencia suponen la opinión de dos personas (que por supuesto tienen puntos de vista diferentes) con respecto a la edad de cierta persona X .

El primer cuerpo de evidencia es la calificación de la persona como de edad $\{ANCIANA, DESCONOCIDA_1\}$ en los siguientes porcentajes:

$$m(ANCIANA) = 0,8, \quad m(DESCONOCIDA_1) = 0,2$$

El segundo cuerpo de evidencia califica a la persona como de edad $\{JOVEN, DESCONOCIDA_2\}$ en los siguientes porcentajes:

$$m(JOVEN) = 0,4, \quad m(DESCONOCIDA_2) = 0,6.$$

La combinación de las dos fuentes evidenciales anteriores nos da un nuevo marco evidencial que se obtiene usando la expresión (5.26) que se obtuvo en una sección anterior de este capítulo. Lo que se necesita obtener son las intersecciones entre los conjuntos descritos en el ejemplo 5.5 y que se ilustran en la figura 5.6.

$$\begin{aligned} \max_{x_i} \mu_{ANCIANA \cap JOVEN}(x_i) &= 0,3 && \text{cuando } x_i = 37,5, \dots, 47,5 \\ \max_{x_i} \mu_{ANCIANA \cap DESCONOCIDA_2}(x_i) &= 1 && \text{cuando } x_i = 37,5, \dots, 80 \\ \max_{x_i} \mu_{DESCONOCIDA_1 \cap JOVEN}(x_i) &= 1 && \text{cuando } x_i = 15, \dots, 47,5 \\ \max_{x_i} \mu_{DESCONOCIDA_1 \cap DESCONOCIDA_2}(x_i) &= 1 && \text{cuando } x_i = 15, \dots, 80 \end{aligned}$$

si hacemos

$$K = \sum_{A,B} \left(1 - \max_{x_i} \mu_{A \cap B}(x_i) \right) m_1(A) m_2(B)$$

entonces

$$\begin{aligned} K &= \left(1 - \max_{x_i} \mu_{ANCIANA \cap JOVEN}(x_i) \right) m(ANCIANA) m(JOVEN) + \dots \\ K &= (1 - 0,3) \times 0,8 \times 0,4 + \dots \\ K &= 0,2240 \end{aligned}$$

y así

$$1 - K = 1 - 0,2240 = 0,7760$$

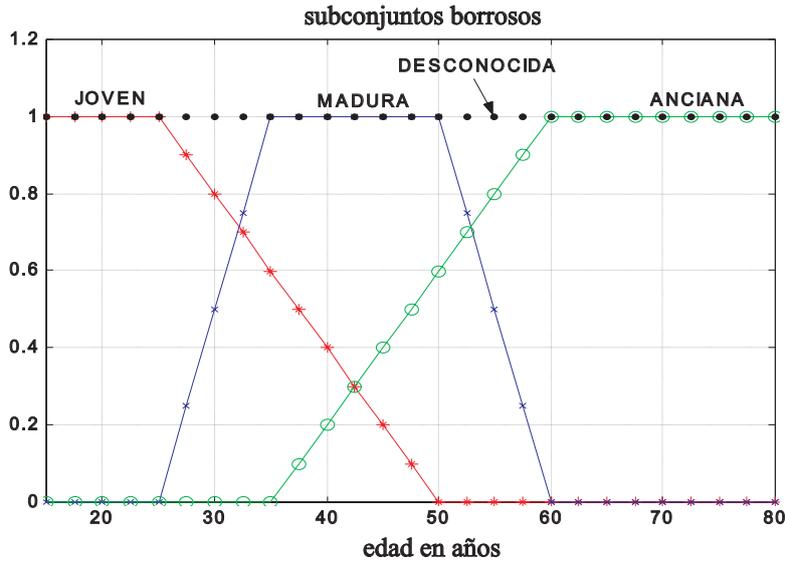


Figura 5.6: Subconjuntos usados en la combinación de dos fuentes evidenciales

$$m(ANCIANA \cap JOVEN) = \frac{\max_{x_i} \mu_{ANCIANA \cap JOVEN}(x_i) m(ANCIANA) m(JOVEN)}{1 - K}$$

$$m(ANCIANA \cap JOVEN) = \frac{0,3 \times 0,8 \times 0,4}{0,7760}$$

$$m(ANCIANA \cap DESCONOCIDA_2) = \frac{1 \times 0,8 \times 0,6}{0,7760}$$

$$m(DESCONOCIDA_1 \cap JOVEN) = \frac{1 \times 0,2 \times 0,4}{0,7760}$$

$$m(DESCONOCIDA_1 \cap DESCONOCIDA_2) = \frac{1 \times 0,2 \times 0,6}{0,7760}$$

$$m(ANCIANA \cap JOVEN) = 0,1237$$

$$m(ANCIANA \cap DESCONOCIDA_2) = 0,6186$$

$$m(DESCONOCIDA_1 \cap JOVEN) = 0,1031$$

$$m(DESCONOCIDA_1 \cap DESCONOCIDA_2) = 0,1546$$

Como todo marco evidencial la suma de todas las asignaciones de probabilidad básica de las intersecciones que resultan de la combinación suman uno. De los resultados deducimos que es muy probable que X sea una persona anciana ($m(ANCIANA) = 0,6186$) y que lo menos probable es que sea una persona joven ($m(JOVEN) = 0,1031$).

Conclusiones del capítulo

Se ha dado una introducción a la teoría probabilista Bayesiana, se ha explicado en que consiste la teoría de evidencia de Dempster-Shafer, ilustrando su aplicación con unos ejemplos. Como parte principal de este capítulo se explicó en que consiste la generalización de la teoría de evidencia de Dempster-Shafer a conjuntos borrosos. Se ilustró con algunos ejemplos las aplicaciones que puede tener el intervalo de creencia, donde también incluimos la combinación evidencial. Para ejemplos no complejos como los estudiados en este capítulo, los resultados resuelven en cierta medida el problema de la incertidumbre. En problemas de mayor complejidad, la toma de decisiones por el análisis de intervalos que resultan de la teoría de evidencia, se puede tornar en algo muy difícil a medida que nuestras variables de decisión se incrementan.

Capítulo 6

Diagnóstico de fallos mediante redes Bayesianas

6.1. Introducción

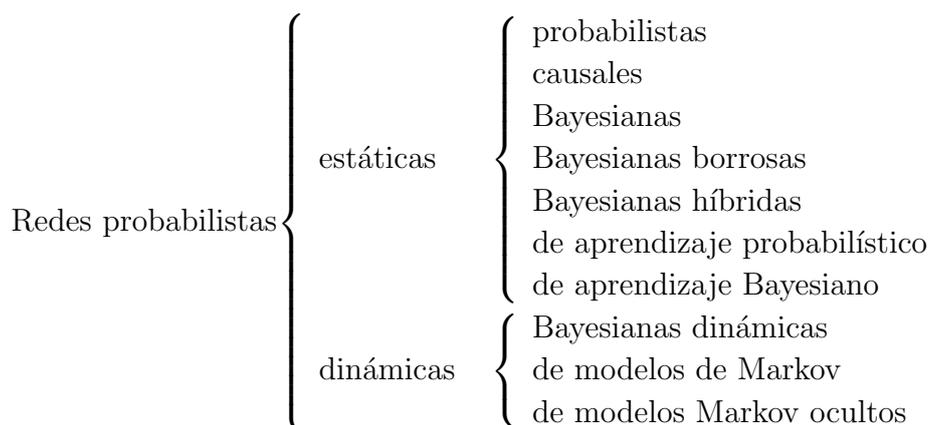
Existen principalmente tres teorías que incorporan el manejo de incertidumbre en sistemas de diagnóstico de fallos: la lógica borrosa maneja vaguedad, la teoría de evidencia de Dempster-Shafer trata con ignorancia o ambigüedad, y la teoría probabilística Bayesiana trata aleatoriedad (Boston 2000). En la década de los 90 se dio un cambio tanto en contenido como en metodología en las investigaciones de Inteligencia Artificial (Russell 2003).

La aparición de Razonamiento probabilístico y sistemas inteligentes, libro escrito por Judea Pearl en 1988 (Pearl 1988), señaló la bienvenida al uso de la probabilidad y de la teoría de decisiones en IA. La invención del formalismo de Redes Bayesianas (o de Creencia) obedece a la necesidad de poder razonar eficientemente ante una combinación de evidencias inciertas. Este enfoque supera con creces a los sistemas de razonamiento probabilístico de los años 60 y 70, y actualmente tiene bastante aceptación en la investigación que se hace en IA sobre razonamiento incierto y sistemas expertos. Es por esas razones que en esta investigación también se incluye el estudio de Redes Bayesianas. Junto con la lógica borrosa y la teoría de evidencia de D-S, tratados en capítulos anteriores, completan el estudio de la incertidumbre en sistemas de diagnóstico de fallos.

Las redes Bayesianas derivan de las redes probabilistas, y en el cuadro sinóptico 6.1 se clasifican con respecto a sus características estáticas y dinámicas. Las de tipo estático, son las que no contemplan cambios de estado o que no varían los valores de probabilidad de sus variables con respecto al tiempo, no así las de tipo dinámico. Las redes Bayesianas tienen una estrecha relación con todos los tipos de redes que derivan de redes probabilistas o que están fundadas en la teoría de probabilidad.

En la literatura a las redes Bayesianas se les menciona como redes de creencia, redes Bayesianas de creencia y algunos autores las llaman redes probabilistas causales. En este

Cuadro 6.1: Clasificación de redes probabilistas



trabajo las llamaremos redes Bayesianas (RB). Existen muchos campos de aplicación para RB, como: minería de datos, diagnóstico médico, análisis de software y otros más. En esta investigación enfocaremos su aplicación a diagnóstico de fallos.

6.2. Aplicaciones de redes Bayesianas

Los avances en los últimos años en el campo de redes Bayesianas aplicadas a diagnóstico de fallos, plantean sistemas de diagnóstico como los de (Sheppard 2003). Estos caracterizan la incertidumbre de los sensores a través de técnicas tomadas de la ciencia de mediciones, que son técnicas probabilistas. Finalmente en la etapa de evaluación de fallos utilizan RB, con lo que le dan un tratamiento probabilístico al manejo de la incertidumbre de principio a fin.

Algunos trabajos como en (Liu 2002), validan la información de los sensores por multivariantes de estado en lo referente a detección de fallos y luego aplican RB para diagnóstico. Acercamientos como los de (Mehrambod 2003), caracterizan los sensores como nodos de una RB, donde cada sensor-singular esta formado por tres nodos de entrada: bias, ruido, y valor de la variable, que son padres de un nodo de salida hijo, que contiene las probabilidades condicionales *a priori* de sus antecesores. Existen también algunos trabajos de aplicación dedicados a diagnóstico con RB, como los de (Chen-Fu 2002) en redes de distribución de energía eléctrica. Otros como los de (Weidl 2005), usan RB modeladas en programación orientada a objetos para sistemas de diagnóstico en procesos de producción de papel. Los de (Dey 2005) en multiprocesos de maquinado con múltiples sensores, para el monitorizado completo del proceso. Por último citamos los de (Mast 1999; Sahin 2007) en diagnóstico de turbinas de aeroplano.

La construcción de Redes Bayesianas es una combinación de arte y ciencia. Algunos autores como (Przytula 2000), creen que sin un significativo progreso en las técnicas de creación de modelos Bayesianos, será muy difícil que esta tecnología llegue a ser muy usada en sistemas de diagnóstico prácticos. Para solventar el problema, propone

la simplificación en la creación de RB dividiéndola en subsistemas. En (Shao-Zhong 2002) proponen, la creación automática de RB, separándolas en dos tareas: aprendizaje de estructura y aprendizaje de parámetros. La primera, consiste en crear la estructura de una RB a partir de una colección de datos. La segunda, consiste en calcular los parámetros numéricos de una estructura dada. Otras contribuciones como las de (Peng 2003), proponen algoritmos de búsqueda para refinar la generación automática de grafos de RB.

En el dominio de las “redes Bayesianas dinámicas” (RBD), además de los “modelos ocultos de Markov” (MOM), acercamientos como los de (Gong 2003), han introducido los “dinámicamente multi-enlazados modelos ocultos de Markov” (DME-MOM). Estos sustituyen los esquemas alternativos como los “modelos ocultos de Markov paralelos” (MOMPa). Estos dos últimos son esquemas parecidos, pero que observan variaciones con respecto a conexiones causales entre los múltiples procesos temporales. Existen aplicaciones recientes de RBD en reconocimiento de escenarios de video (Ayers 2000) y diagnóstico de fallos (Sterritt 2000).

En “redes Bayesianas híbridas” (RBH), podemos mencionar contribuciones como las de (Whei-Min 2004), que proponen “redes probabilistas neuronales”, que usan información primaria para crear conjuntos de entrenamiento. Cuando la topología de red cambia, sus capacidades adaptativas como red neuronal entran en función. Baldwin (Baldwin 2003) propone “redes Bayesianas borrosas”, donde conjuntos borrosos son usados para rescribir información contenida en la base de datos. Lo anterior con el fin de reducir la complejidad del aprendizaje automático de datos en RB.

6.3. Redes Bayesianas

La principal ventaja del razonamiento probabilístico en comparación con otro, como el lógico, reside en que a través del razonamiento probabilístico, se pueden tomar decisiones racionales. Esto es, podemos tomar una decisión aun sin disponer de suficiente información para probar categóricamente que una acción dada funcionará. Sin embargo, aunque no pruebe cien por ciento una hipótesis dada, los resultados esperados son muy aceptables.

En temas anteriores se explicó el empleo de la regla de Bayes. Las relaciones de independencia condicional que existen entre las variables, permiten simplificar el cálculo de los resultados de una consulta. También permite reducir considerablemente la cantidad de probabilidades condicionales que es necesario especificar.

6.3.1. Representación del conocimiento en un dominio incierto

Para representar la dependencia que existe entre determinadas variables se utiliza una estructura de datos conocida como red Bayesiana, y también para especificar de manera

concreta la distribución de probabilidad conjunta. La RB es una gráfica en la que se cumple lo siguiente:

- Los nodos de la red están formados por un conjunto de variables aleatorias.
- Cada par de nodos se conecta entre sí mediante un conjunto de enlaces o flechas. El significado implícito de una flecha que vaya del nodo X al nodo Y es el de que X ejerce una influencia directa sobre Y .
- Por cada nodo hay una tabla de probabilidad condicional que sirve para cuantificar los efectos de los padres sobre el nodo. Los padres de un nodo son aquellos nodos cuyas flechas apuntan hacia éste.
- La gráfica no tiene ciclos dirigidos (por lo tanto, es una gráfica dirigida acíclica, o GDA).

Por lo general, para el experto de un dominio resulta sencillo saber qué relaciones de dependencia condicional son válidas: mucho más fácil que la especificación de las probabilidades. Una vez definida la topología de la RB, sólo hay que especificar las probabilidades condicionales de los nodos que intervienen en las dependencias directas y utilizarlos en el cálculo de cualquier otro valor de probabilidad. En la inferencia probabilista las conectivas \wedge , \vee , \neg , pueden interpretarse respectivamente, como la intersección, unión y complementación de probabilidades.

Consideremos la construcción de una Red Bayesiana a través de un ejemplo de apilamiento de bloques por medio de un brazo de robot (ver figura 6.1).

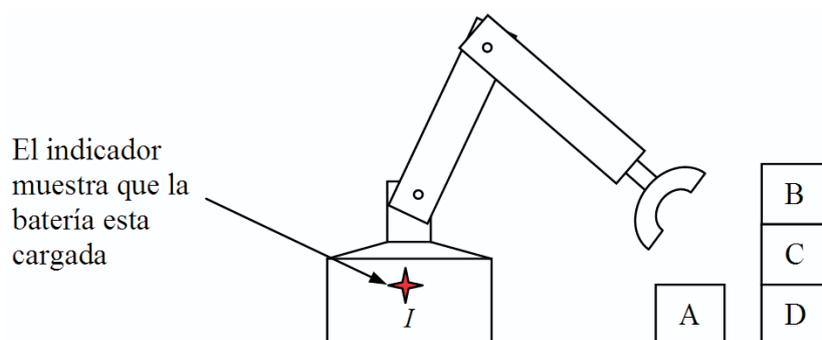


Figura 6.1: Apilamiento de bloques por un brazo de robot

Comenzaremos con lo que imaginamos son las “primeras causas” en este dominio, a saber, las variables que corresponden a las proposiciones “la batería está cargada” (B) y “el bloque se puede levantar” (L). B y L son causas de S (“el brazo se mueve”), y B es una causa de I (“el indicador muestra que la batería está cargada”). Así, dibujaríamos la RB de este problema como la que se muestra en la figura 6.2. Fijémonos en que, sobre otras cosas, la red explicita que $p(S|I, B, L) = p(S|B, L)$. Si hubiese un nodo en la red, llamémoslo E (cuyo significado fuese que el bloque está elevado), no podría

sucedir que $p(S|I, B, L, E) = p(S|B, L)$, porque E sería un descendiente de S (que el bloque esté elevado influye a la probabilidad de que el brazo se mueva. ¿De qué otra forma podría elevarse el bloque?). La expresión de la probabilidad conjunta de todos los nodos de la red se muestra en la figura 6.2.

Vemos que para poder calcular el valor de las probabilidades conjuntas dadas por la RB, necesitamos conocer las funciones de probabilidad condicional de cada nodo de la red que esté influido por sus padres, tal como se puede observar en la figura 6.2. Para los nodos sin padres, las probabilidades no están condicionadas por otros nodos, por lo que se les denomina probabilidades *a priori* de estas variables (nodos). Por tanto una especificación completa de las probabilidades de un conjunto de variables aleatorias significa tener una RB de estas variables, junto a las tablas de probabilidades condicionales (TPC) de cada nodo de la red. La fórmula de la función de probabilidad

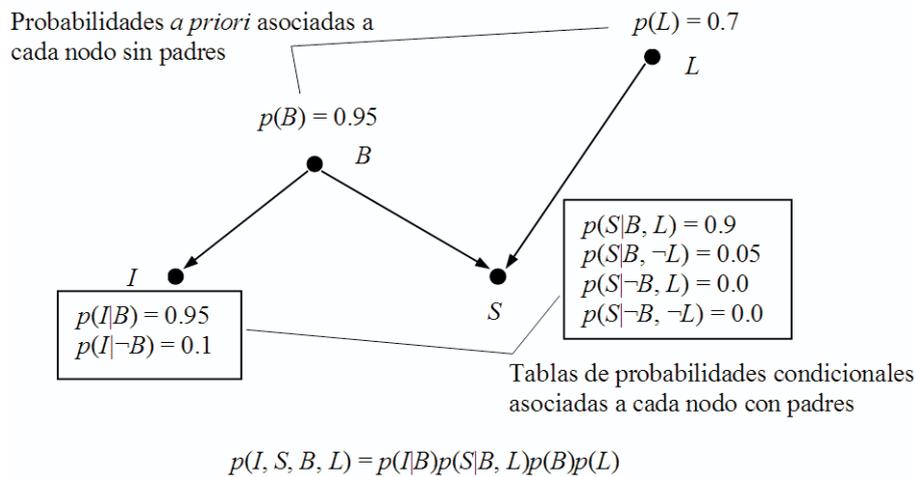


Figura 6.2: Ejemplo de una red Bayesiana

conjunta de la RB de nuestro ejemplo del apilamiento de bloques, se podría comparar con una parecida (sin independencias condicionales), que se obtendría mediante la regla encadenada:

$$p(I, S, B, L) = p(I \wedge S \wedge B \wedge L) = p(I|S, B, L)p(S|B, L)p(B|L)p(L)$$

Notemos que la fórmula de la RB es mucho más sencilla (ver figura 6.2). Sin las independencias condicionales especificadas en la RB, la especificación de la probabilidad conjunta para las cuatro variables del ejemplo requiere la especificación de 16 probabilidades conjuntas distintas. Tal como se evidencia en la figura 6.2, las asunciones hechas en la RB nos permiten tener que calcular sólo ocho probabilidades. Cuando hay muchas independencias condicionales entre las variables del dominio, la expresión de la probabilidad conjunta que se calcula a partir de la RB necesita la especificación de muchas menos probabilidades que las que necesitaría sin dichas independencias. Esta reducción convierte en tratables muchos problemas que, de otro modo, no lo serían.

6.3.2. Representación de la distribución de probabilidad conjunta

La RB permite obtener una descripción completa del dominio. Es posible calcular cada una de las entradas de la distribución de probabilidad conjunta desde la información de la red. Una entrada genérica en la probabilidad conjunta es la probabilidad de que se dé la conjunción de determinadas asignaciones para cada una de las variables, como sería $p(X_1 = x_1 \wedge \dots \wedge x_n = x_n)$. Usamos la notación $p(x_1, \dots, x_n)$ como una abreviatura para esto. El valor de esta entrada está dado por la siguiente expresión:

$$p(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i | \text{padres}(X_i)) \quad (6.1)$$

Por lo tanto, las entradas de la probabilidad conjunta se representan mediante el producto de los respectivos elementos de las tablas de probabilidad condicional (TPC) de la RB. Es decir, las TPC constituyen una representación analítica de la probabilidad conjunta.

Como ejemplo, podemos calcular la probabilidad del evento de que el brazo no se pueda mover, dadas las condiciones de que el indicador este encendido, la batería esta cargada y el bloque se pueda levantar.

$$\begin{aligned} p(\neg S \wedge I \wedge B \wedge L) &= p(\neg S | I, B, L) p(I | B, L) p(B | L) p(L) \\ &= p(\neg S | B, L) p(I | B) p(B) p(L) \\ &= 0,1 \times 0,95 \times 0,95 \times 0,7 \\ &= 0,0631 \end{aligned}$$

Si la RB es una representación de la probabilidad conjunta, también puede servir para responder cualquier tipo de consulta.

6.3.3. Un método para construir redes Bayesianas

La ecuación (6.1) define el significado de una determinada RB. Sin embargo, no explica cómo construir una RB de manera que la distribución conjunta resultante sea una buena representación de un determinado dominio. Se mostrará a continuación que la ecuación (6.1) implica ciertas relaciones de independencia condicional que sirven para construir la topología de la RB. Primero, se reescribirá la conjunción en función de una probabilidad condicional; para ello se utilizará la definición:

$$p(x_1, \dots, x_n) = p(x_n | x_{n-1}, \dots, x_1) p(x_{n-1}, \dots, x_1)$$

Se repite este procedimiento, reduciendo cada una de las probabilidades conjuntas a una probabilidad condicional y a una conjunción más pequeña. Esto también puede

ser visto como una probabilidad conjunta sobre la base de una *cadena* de probabilidades condicionales. Por último, se obtiene el siguiente producto general que también es conocido como *regla encadenada*:

$$p(x_1, \dots, x_n) = p(x_n|x_{n-1}, \dots, x_1)p(x_{n-1}|x_{n-2}, \dots, x_1) \cdots p(x_2|x_1)p(x_1)$$

$$p(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i|x_{i-1}, \dots, x_1)$$

Si se compara esta expresión con la ecuación (6.1), veremos que la especificación de la conjunción equivale a la aseveración general:

$$p(X_i|X_{i-1}, \dots, X_1) = p(X_i|Padres(X_i)) \quad (6.2)$$

suponiendo que $Padres(X_i) \subseteq \{x_{i-1}, \dots, x_1\}$. Esta última condición se puede satisfacer sin mayor problema identificando los nodos de acuerdo con un orden congruente con el orden parcial implícito en la estructura de la gráfica.

La ecuación (6.2) significa que la RB será la representación correcta del dominio sólo si cada uno de los nodos tiene la independencia condicional respecto de sus predecesores en la secuencia de nodos, tomando como referencia a sus padres. Por lo tanto, si se desea construir una RB cuya estructura sea la adecuada para el dominio, hay que escoger los padres de cada nodo de manera que se satisfaga la propiedad anterior.

El procedimiento general para la construcción de una RB es el siguiente:

- 1. Escoja el conjunto de variables X_i que sirvan para describir el dominio.
- 2. Defina la manera como se van a ordenar las variables.
- 3. Siempre que haya variables:
 - a). Por cada variable X_i que escoja, añada a la red un nodo.
 - b). Asigne $Padres(X_i)$ a un conjunto mínimo de nodos que esté presente en la red, para de esta manera satisfacer la propiedad de independencia condicional (6.2).
 - c). Elabore la tabla de probabilidad condicional correspondiente a X_i .

Puesto que en esta construcción un nodo está conectado solamente con nodos anteriores, este método garantiza la obtención de una red acíclica. Otra característica importante de las RB es que en ellas no hay valores de probabilidad redundante.

6.3.4. Representación de las tablas de probabilidad condicional

En una RB cuando la relación que existe entre padres e hijos es totalmente arbitraria, para la elaboración de las tablas de probabilidad condicional se necesitan muchos números. Dada la cantidad de posibles casos de condicionamiento, para llenar las tablas se

necesita bastante tiempo y mucha experiencia. En realidad lo anterior sucede en el peor de los casos, por lo general, la relación padre-hijo corresponde a una de varias categorías que tienen distribuciones canónicas; es decir, corresponden a un patrón estándar. En casos como el anterior, para especificar la tabla basta con indicar un patrón específico y proporcionar algunos parámetros.

Nodos deterministas

El ejemplo más sencillo de lo anterior lo constituyen los nodos deterministas. El valor de un nodo determinista está definido exactamente por los valores de sus padres, sin lugar a incertidumbre. La relación puede ser de tipo lógico: por ejemplo, la relación que existe entre los nodos padre *canadiense*, *estadounidense*, *mexicano* y el nodo hijo *norteamericano* consiste en que el hijo es la disyunción de los padres. La relación puede ser también de tipo numérico. Por ejemplo, si los nodos padres son los diversos precios de un coche en distintas distribuidoras, y el nodo hijo es el precio final pagado por un cazador de ofertas, entonces el nodo hijo será el mínimo de los valores de los padres.

Relaciones lógicas ruidosas

A menudo, las relaciones inciertas se pueden caracterizar por lo que se denomina relaciones lógicas ruidosas. El ejemplo típico es la llamada relación *O-ruidosa*, que viene a ser una generalización de la relación *O lógica*. En la lógica propositiva se puede afirmar que fiebre es verdadera si y sólo si “*gripe, catarro o malaria*” son también verdaderas. El modelo *O-ruidoso* añade cierta incertidumbre a este enfoque estrictamente lógico.

Este modelo hace tres suposiciones. Primero, se supone que cada causa tiene una posibilidad independiente de producir el efecto. Segundo, se supone que todas las causas posibles están listadas. En tercer lugar se supone que lo que impida, por ejemplo, a “*resfriado*” producir “*fiebre*” es independiente de lo que impida a “*gripe*” producir “*fiebre*”. Estos inhibidores no se presentan como nodos, sino que se agrupan como “parámetros de ruido”. Si $p(\text{fiebre}|\text{resfriado}) = 0,4$, $p(\text{fiebre}|\text{gripe}) = 0,8$ y $p(\text{fiebre}|\text{malaria}) = 0,9$, entonces los parámetros de ruido son 0.6, 0.2 y 0.1, respectivamente. Si ninguno de los nodos padres es verdadero, la certeza de que el nodo de salida sea falso es de 100%. Si sólo uno de los padres es verdadero, la probabilidad de que la salida sea falsa es igual a la del parámetro de ruido correspondiente a ese nodo. En general, la probabilidad de que el nodo de salida sea falso es precisamente el producto de los parámetros de ruido de todos los nodos de entrada que son verdaderos. Una ilustración de lo anterior se muestra en el cuadro 6.2. Las relaciones lógicas ruidosas en las que una variable depende de k padres se describirá utilizando $O(k)$ parámetros en vez de $O(2^k)$, obteniendo toda la tabla de probabilidad condicional. De esta forma la evaluación y el aprendizaje se simplifican bastante.

Cuadro 6.2: Ejemplo de parámetros de ruido

<i>resfriado</i>	<i>gripe</i>	<i>malaria</i>	$p(\textit{fiebre})$	$p(\neg\textit{fiebre})$
F	F	F	0,0	1,0
F	F	T	0,9	0,1
F	T	F	0,8	0,2
F	T	T	0,98	$0,02 = 0,2 \times 0,1$
T	F	F	0,4	0,6
T	F	T	0,94	$0,06 = 0,6 \times 0,1$
T	T	F	0,88	$0,12 = 0,6 \times 0,2$
T	T	T	0,988	$0,012 = 0,6 \times 0,2 \times 0,1$

6.4. Inferencia en redes Bayesianas

El principal objetivo de un sistema de inferencia probabilista es el cálculo de la distribución de probabilidad posterior de un conjunto de variables de consulta, con base en determinadas variables de evidencia. Es decir, el sistema calcula $p(\textit{consulta}|\textit{evidencia})$. Las RB tienen una flexibilidad que permiten a un nodo funcionar como variable de consulta o como variable de evidencia.

6.4.1. Patrones de inferencia en redes Bayesianas

Existen tres patrones importantes de inferencia en las RB. Para explicarlos se hace referencia al ejemplo del brazo de robot de la subsección 6.3.1.

Inferencia causal (o descendente)

Supongamos que deseamos calcular $p(S|L)$, la probabilidad de que el brazo se mueva dado que el objeto se puede elevar. Como el hecho de que el objeto sea elevable es una de las causas de que el brazo pueda moverse, decimos que este cálculo es un ejemplo de *razonamiento causal*. A L se le denomina *evidencia* utilizada en la inferencia, y a S el nodo de *consulta*. Así es cómo realizamos la inferencia: primero, expandimos $p(S|L)$ (probabilidad marginal) en la sumatoria de dos probabilidades conjuntas (porque debemos tener en cuenta el otro padre de S , B):

$$p(S|L) = p(S, B|L) + p(S, \neg B|L).$$

Luego, queremos que S esté condicionado, tanto a este otro padre, como a L , así que utilizamos la expresión de la regla encadenada para escribir:

$$p(S|L) = p(S|B, L)p(B|L) + p(S|\neg B, L)p(\neg B|L).$$

Pero $p(B|L) = p(B)$ (se obtiene de la propia estructura de la red; fijémonos en que B no tiene padres). E igualmente, $p(\neg B|L) = p(\neg B)$. Por tanto, $p(S|L) = p(S|B, L)p(B) + p(S|\neg B, L)p(\neg B)$. Como todas estas cantidades se dan en la red, podemos calcular y obtener:

$$p(S|L) = (0,9 \times 0,95) + (0 \times 0,05)$$

$$p(S|L) = 0,855$$

Las operaciones que hemos realizado en este ejemplo merecen anotarse, porque pueden generalizarse a versiones más complejas de razonamiento causal. Las principales operaciones son:

- a) Escribir las probabilidades condicionales del nodo de consulta X , dada la evidencia, en términos de la probabilidad conjunta de X y todos sus padres (los cuales no son la evidencia), dada la evidencia.
- b) Expresar esta probabilidad conjunta como la probabilidad de X , condicionada por todos sus padres.

Inferencia de diagnóstico o ascendente

Ahora se calculará $p(\neg L|\neg S)$, la probabilidad de que el bloque no se pueda elevar dado que el brazo no se mueve. Aquí, los papeles de la consulta y de la evidencia son a la inversa del último ejemplo. Como estamos utilizando un efecto o síntoma para inferir una causa, a este tipo de razonamiento lo denominaremos *razonamiento de diagnóstico*.

$$p(\neg L|\neg S) = \frac{p(\neg S|\neg L)p(\neg L)}{p(\neg S)} \quad (\text{teorema de Bayes})$$

Ahora calculamos $p(\neg S|\neg L) = 0,9525$ (mediante razonamiento causal), y luego sustituimos en

$$P(\neg L|\neg S) = \frac{0,9525 \times 0,3}{p(\neg S)} = \frac{0,2857}{p(\neg S)}$$

de igual modo,

$$p(L|\neg S) = \frac{p(\neg S|L)p(L)}{p(\neg S)} = \frac{0,145 \times 0,7}{p(\neg S)} = \frac{0,1015}{p(\neg S)}$$

y como ambas expresiones deben sumar 1, tenemos $p(\neg L|\neg S) = 0,7378$. Los cálculos realizados en este sencillo ejemplo de razonamiento diagnóstico también se pueden generalizar. El paso principal es la utilización del teorema de Bayes para convertir el problema a uno de razonamiento causal.

Inferencia intercausal o justificación

Si nuestra única evidencia es que $\neg S$ esto es, el brazo no se mueve, podemos calcular la probabilidad de que el bloque no se puede elevar, $\neg L$ tal como hicimos. Pero si también

tenemos $\neg B$ esto es, la batería no está cargada, entonces $\neg L$ debería ser menos cierto (tener menos peso). En este caso, decimos que $\neg B$ *justifica* $\neg S$, haciendo a $\neg L$ menos cierto. Este tipo de inferencia utiliza un proceso de razonamiento causal o descendente incrustado en un proceso de diagnóstico o ascendente.

$$p(\neg L|\neg B, \neg S) = \frac{p(\neg S, \neg B|\neg L)p(\neg L)}{p(\neg B, \neg S)} \text{ teorema de Bayes}$$

$$p(\neg L|\neg B, \neg S) = \frac{p(\neg S|\neg B, \neg L)p(\neg B|\neg L)p(\neg L)}{p(\neg B, \neg S)} \text{ (def. de probabilidad condicional)}$$

$$p(\neg L|\neg B, \neg S) = \frac{p(\neg S|\neg B, \neg L)p(\neg B)p(\neg L)}{p(\neg B, \neg S)} \text{ (estructura de la RB)}$$

De esta expresión, utilizando las probabilidades de la red, y resolviendo $p(\neg B, \neg S)$ de la forma habitual, obtenemos $p(\neg L|\neg B, \neg S) = 0,30$, que es tal como se esperaba mucho menor que $p(\neg L|\neg S)$ (calculado anteriormente). Otra vez, nos debemos fijar en el uso del teorema de Bayes, que es un paso bastante importante de la justificación.

6.4.2. Separación dependiente de la dirección o separación-d

Ocurre que una RB tiene más independencias condicionales que las que sólo involucran a los padres de un nodo. Por ejemplo, en la figura 6.2. $p(S|I, B) = p(S|B)$, es decir, S es condicionalmente independiente de I dado B (aunque no nos dan ambos padres de S). De forma intuitiva, el conocimiento del efecto I puede influir sobre el conocimiento de la causa B , el cual influye sobre el conocimiento del efecto S . Pero si nos dan la causa B , no hay nada que nos pueda decir I acerca de S . En este caso, decimos que B *separa-d*, esta es la separación dependiente de la dirección a S de I .

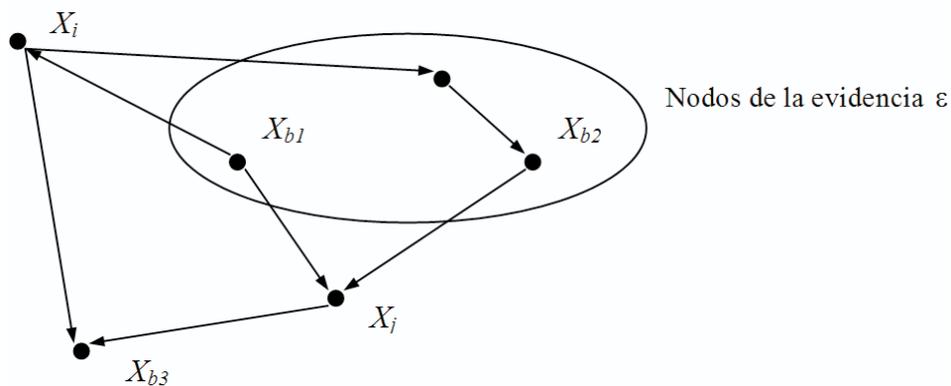


Figura 6.3: Ejemplo de RB que observa *separación-d*

Dos nodos X_i y X_j son independientes condicionalmente dado un conjunto de nodos ϵ (es decir, $\mathbf{I}(X_i, X_j)|\epsilon$) si por cada camino no dirigido entre X_i y X_j hay algún nodo X_b , que cumple algunas de las siguientes tres propiedades (ver figura 6.3):

- 1. X_b pertenece a ϵ , y ambos arcos salen de X_b .
- 2. X_b pertenece a ϵ , y un arco va hacia X_b y otro arco sale de él.
- 3. Ni X_b ni ningún descendiente suyo pertenece a ϵ , y ambos arcos van hacia X_b .

En la figura 6.3, dado el conjunto de nodos de evidencia, X_i es independiente de X_j debido a que los tres caminos se bloquean entre sí. Los nodos que bloquean son:

- a) X_{b1} es un nodo evidencia, y ambos arcos salen de X_{b1} .
- b) X_{b2} es un nodo evidencia, y un arco va hacia él y otro sale de él.
- c) X_{b3} no es un nodo de evidencia, ni ninguno de sus descendientes, y los dos arcos van a él.

Cuando cualquiera de estas tres condiciones se cumple, decimos que el nodo X_b bloquea el camino dado ϵ . Tengamos en cuenta que los caminos (rutas) de los que hablamos son caminos indirectos, es decir, caminos que ignoran la dirección de los arcos. Si todos los caminos X_i y X_j están bloqueados, entonces decimos que ϵ separa-d X_i de X_j (separación dependiente de la dirección), y concluimos que X_i y X_j son independientes condicionalmente dado ϵ . Ejemplos en la figura 6.2 de independencia condicional (**I**) por causa de separación-d son:

- **I**($I, L|B$); por la regla 1, B bloquea el único camino entre I y L dado B ; por la regla 3, S también bloquea este camino dado B , porque S no es miembro del conjunto de evidencia.
- **I**(I, L); por la regla 3, S bloquea el único camino entre I y L , entre B y L , dado el conjunto de evidencia vacío (S no es miembro del conjunto de evidencia).

Sin embargo, nos debemos fijar en que B y L no son condicionalmente independientes dado S ($p(B, L|S)$), ya que, aunque S está en el camino entre B y L , los dos arcos de este camino van a S , y S pertenece al conjunto de evidencia; por tanto, S no bloquea el camino entre B y L .

El concepto de separación-d también se puede aplicar a conjuntos de nodos. Dos conjuntos X_i y X_j son condicionalmente independientes dado ϵ , si están separados-d por ϵ ; es decir, si cada camino no dirigido entre los nodos de X_i y todos los nodos de X_j están bloqueados dado ϵ .

Aun con la utilización de la separación-d, la inferencia probabilística en RB es por lo general complejo, (*NP-hard*). Sin embargo, se pueden hacer algunas simplificaciones para una clase muy importante de redes, denominada poliárboles. Un poliárbol es un GDA en el que, para cada par de nodos, sólo hay un camino a lo largo de los arcos en cualquier dirección (red de conexión única). Por ejemplo, la red de la figura 6.2 es un poliárbol. En la siguiente sección se muestra la inferencia probabilista en poliárboles con un ejemplo más desarrollado, el método que se muestra está basado en un algoritmo propuesto por (Russell 1996).

6.5. Inferencia probabilista en poliárboles

La RB que se muestra en la figura 6.4 es el típico ejemplo de poliárbol. Lo que queremos es calcular la probabilidad de Q a partir de los otros nodos en la RB. Fijémonos en que algunos nodos están conectados a Q sólo mediante los padres del mismo; entonces diremos que estos nodos están por encima de Q . Los otros nodos están conectados sólo mediante los sucesores inmediatos de Q (sus hijos); diremos que éstos están por debajo de Q . También nos debemos fijar en que no hay ningún otro camino (ruta) excepto los que unen un nodo por encima de Q a un nodo por debajo de Q (si no la RB ya no sería un poliárbol). Estas definiciones y propiedades de conectividad se aplican a todos los nodos del poliárbol; tendremos tres tipos de evidencias:

- 1. Todos los nodos de la evidencia están por encima de Q (apoyo causal). Como ejemplo típico de este tipo de evidencia, calcularemos $p(Q|P5, P4)$.
- 2. Todos los nodos de la evidencia están por debajo de Q (apoyo evidencial). Como ejemplo típico de este tipo de evidencia, calcularemos $p(Q|P12, P13, P14, P11)$.
- 3. Hay nodos de la evidencia por encima y por debajo de Q .

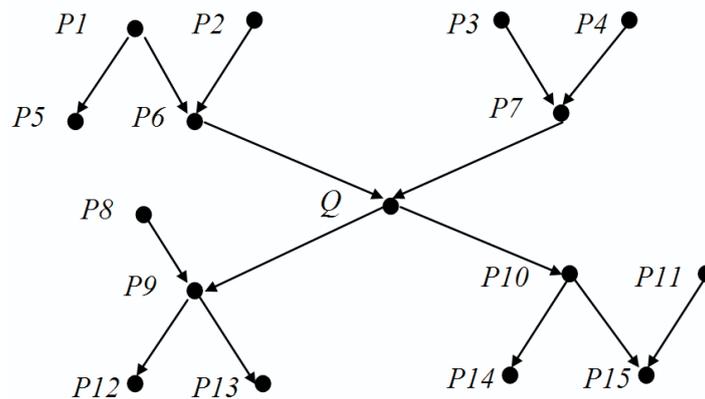


Figura 6.4: Un poliárbol típico

6.5.1. Apoyo causal

Se calculará $p(Q|P5, P4)$, donde todos los nodos de la evidencia están por encima de Q . El cálculo es un trazo de la ejecución de un algoritmo recursivo “ascendente” que calcula la probabilidad de cada antecesor de Q , dada la evidencia, hasta que se alcanza la evidencia o hasta que la evidencia está por debajo del antecesor. El algoritmo se comporta de la siguiente manera:

Primero involucramos a los padres de Q

$$p(Q|P5, P4) = \sum_{P6, P7} p(Q, P6, P7|P5, P4)$$

(La notación especial que se utiliza para indexar esta sumatoria, es para indicar que añadimos cuatro versiones de $p(Q, P6, P7|P5, P4)$, la original, una con $\neg P6$ en vez de $P6$, otra con $\neg P7$ en vez de $P7$, y la última con $\neg P6$ y $\neg P7$). Luego introducimos a los padres de Q como parte de la evidencia usando la definición de la independencia condicional:

$$p(Q, P6, P7|P5, P4) = p(Q|P6, P7, P5, P4)p(P6, P7|P5, P4)$$

Con la sustitución obtenemos:

$$p(Q|P5, P4) = \sum_{P6, P7} p(Q|P6, P7, P5, P4)p(P6, P7|P5, P4)$$

ahora, como un nodo es condicionalmente independiente de sus antecesoros, dados sus padres:

$$p(Q|P5, P4) = \sum_{P6, P7} p(Q|P6, P7)p(P6, P7|P5, P4)$$

entonces, la separación-d nos permite fragmentar a los padres:

$$p(Q|P5, P4) = \sum_{P6, P7} p(Q|P6, P7)p(P6|P5, P4)p(P7|P5, P4)$$

y finalmente, la separación-d nos permite ignorar la evidencia por encima de uno de los padres al calcular la probabilidad del otro padre:

$$p(Q|P5, P4) = \sum_{P6, P7} p(Q|P6, P7)p(P6|P5)p(P7|P4)$$

Es muy importante fijarnos en que los términos que se suman son *a*) las probabilidades del nodo de consulta, dados los valores de sus padres (cuyas probabilidades se dan en la RB), y *b*) las probabilidades de cada uno de los padres se dan a partir de la evidencia por encima de ellos mediante una llamada recursiva al algoritmo que estamos describiendo. Estos resultados se siguen de forma directa por el hecho de que estamos trabajando con un poliárbol.

Este mismo procedimiento se aplica recursivamente hasta que finalmente se alcanzan los nodos que tienen un nodo de la evidencia como padre o los que no tienen padres (nodos que ellos mismos no son evidencia). El primero de los dos casos lo tenemos en el cálculo de $p(P7|P4)$, donde el nodo de la evidencia $P4$ es uno de los padres del nodo de consulta $P7$. En este caso, el proceso de involucrar a los padres es más sencillo, ya que uno de los padres de $P7$ ya ha sido involucrado. Como $p(P7, P3|P4) = p(P7|P3, P4)p(P3|P4)$, podemos escribir:

$$p(P7|P4) = \sum_{P3} p(P7|P3, P4)p(P3|P4) = \sum_{P3} p(P7|P3, P4)p(P3)$$

El último paso se sigue de $\mathbf{I}(P3, P4)$. Todos los términos de la sumatoria se obtienen de la RB, entonces, el procedimiento para el ejemplo finaliza en esta ramificación. Al calcular $p(P6|P5)$, obtenemos:

$$p(P6|P5) = \sum_{P1, P2} p(P6|P1, P2)p(P1|P5)p(P2)$$

Aquí debemos calcular $p(P1|P5)$, y nos fijamos en que el nodo de la evidencia no se encuentra por encima del nodo de consulta, sino por debajo, por lo que no podemos proseguir con la llamada recursiva, pero debemos utilizar el procedimiento de “apoyo evidencial” (que aún no hemos descrito). En este ejemplo simplemente, utilizamos el teorema de Bayes para obtener

$$p(P1|P5) = \frac{p(P5|P1)p(P1)}{p(P5)}.$$

Después, todas las cantidades que se necesitan para calcular $p(P6|P5)$ se dan en la RB. Podemos juntar todos estos resultados (realizando todas las sumatorias) para obtener la respuesta final de $p(Q|P5, P4)$.

6.5.2. Apoyo evidencial

A continuación se calcula $p(Q|P12, P13, P14, P11)$, donde todos los nodos evidencia están por debajo de Q . Nuestros cálculos, van a ser traceados de la ejecución de un algoritmo recursivo. Su funcionamiento es que en el nivel principal utilizamos la regla de Bayes para escribir:

$$p(Q|P12, P13, P14, P11) = \frac{p(P12, P13, P14, P11|Q)p(Q)}{p(P12, P13, P14, P11)}$$

$$p(Q|P12, P13, P14, P11) = kp(P12, P13, P14, P11|Q)p(Q)$$

donde $k = \frac{1}{p(P12, P13, P14, P11)}$ es un factor de normalización que se calculará más adelante, de la misma forma como lo hicimos en los ejemplos iniciales. Por separación-d $\mathbf{I}(\{P12, P13\}, \{P14, P11\}|Q)$, obtenemos

$$p(Q|P12, P13, P14, P11) = kp(P12, P13|Q)p(P14, P11|Q)p(Q).$$

Nos tenemos que fijar en que hemos separado el conjunto $\{P12, P13, P14, P11\}$ en dos subconjuntos correspondientes a los de los dos hijos de Q . Cada uno de los términos $p(P12, P13|Q)$ y $p(P14, P11|Q)$ son un caso del cálculo de la probabilidad de un conjunto de nodos de consulta, dado un único nodo de evidencia que está por encima de ellos. Entonces, podemos utilizar un algoritmo parecido al anterior. Como sólo hay un único nodo de la evidencia, es conveniente utilizar el algoritmo recursivo descendente en vez del ascendente.

Vamos a ilustrar el funcionamiento de la versión descendente en el cálculo de $p(P12, P13|Q)$. El paso clave es involucrar al único hijo de Q , $P9$, que esta por encima del conjunto de nodos de consulta $\{P12, P13\}$. Primero, fijémonos en que $p(P12, P13, P9|Q) = p(P12, P13|P9, Q)p(P9|Q)$, por la definición de la independencia condicional. Entonces:

$$p(P12, P13|Q) = \sum_{P9} p(P12, P13|P9, Q)p(P9|Q)$$

ahora, por separación-d de $\mathbf{I}(\{P12, P13\}, Q|P9)$, tenemos

$$p(P12, P13|Q) = \sum_{P9} p(P12, P13|P9)p(P9|Q)$$

de los términos de la sumatoria, $p(P9|Q)$ se calcula involucrando a todos los padres de $P9$:

$$p(P9|Q) = \sum_{P8} p(P9|P8, Q)p(P8)$$

$p(P9|P8, Q)$ se da en la red. El otro término, $p(P12, P13|P9)$ es una llamada recursiva al procedimiento ascendente que calcula la probabilidad del conjunto de nodos de consulta dado un único nodo de la evidencia, que está por encima de ellos. En este caso, la llamada recursiva finaliza después de un paso, ya que los hijos de $P9$ son los nodos de la evidencia. Como $P12$ y $P13$ son independientes dado $P9$, tenemos que $p(P12, P13|P9) = p(P12|P9)p(P13|P9)$. Ambas probabilidades se dan en la red. Aplicando el procedimiento descendente sobre $p(P14, P11|Q)$, obtenemos:

$$p(P14, P11|Q) = \sum_{P10} p(P14, P11|P10)p(P10|Q)$$

y entonces:

$$p(P14, P11|Q) = \sum_{P10} p(P14|P10)p(P11|P10)p(P10|Q)$$

porque $\mathbf{I}(P14, P11|P10)$. Sólo el término central de la multiplicación no se da en la red; entonces utilizamos el procedimiento descendente para calcularlo:

$$p(P11|P10) = \sum_{P15} p(P11|P15, P10)p(P15|P10)$$

aquí;

$$p(P15, P10) = \sum_{P11} p(P15|P10, P11)p(P11)$$

Pero en $p(P11|P15, P10)$, el nodo de consulta $P11$ está por encima de los nodos de la evidencia, así que aplicamos otra vez el nivel principal del procedimiento (mediante el teorema de Bayes):

$$p(P11|P15, P10) = \frac{p(P15, P10|P11)}{p(P15, P10)} = k_1 p(P15, P10|P11)p(P11)$$

donde $k_1 = \frac{1}{p(P15, P10)}$ y $p(P11)$ se da directamente en la red. El algoritmo finaliza con:

$$p(P15, P10|P11) = p(P15|P10, P11)p(P10|P11) = p(P15|P10, P11)p(P10)$$

porque $P10$ y $P11$ son independientes.

Ahora se pueden agrupar todos los resultados, y las sumatorias, y k y k_1 se pueden calcular, para obtener la respuesta final de $p(Q|P12, P13, P14, P11)$. La complejidad de los algoritmos, el de apoyo causal y el de apoyo evidencial, es lineal en relación al número de nodos en la red (sólo en poliárboles).

6.5.3. Apoyo causal y evidencial

Si disponemos de la evidencia por encima y por debajo de Q , como en:

$$p(Q|\{P5, P4\}, \{P12, P13, P14, P11\})$$

separamos la evidencia en el subconjunto de apoyo causal ϵ^+ , y en el de apoyo evidencial ϵ^- , y utilizamos el teorema de Bayes para escribir:

$$p(Q|\epsilon^+, \epsilon^-) = \frac{p(\epsilon^-|Q, \epsilon^+)p(Q|\epsilon^+)}{p(\epsilon^-, \epsilon^+)}.$$

Como es habitual, tratamos $\frac{1}{p(\epsilon^-|\epsilon^+)} = k_2$ como un factor de normalización, y escribimos:

$$p(Q|\epsilon^+, \epsilon^-) = k_2 p(\epsilon^-|Q, \epsilon^+) p(Q|\epsilon^+)$$

fijémonos en que Q separa-d a ϵ^- de ϵ^+ , de esta manera:

$$p(Q|\epsilon^+, \epsilon^-) = k_2 p(\epsilon^-|Q) p(Q|\epsilon^+).$$

Ya calculamos la primera probabilidad como parte del procedimiento descendente para calcular $p(Q|\epsilon^-)$; y la segunda probabilidad se calculó directamente con el procedimiento ascendente.

6.5.4. Un ejemplo numérico

Se demuestra numéricamente a continuación la utilidad de estos métodos con un ejemplo, a partir del poliárbol abstracto que se muestra en la figura 6.5. Lo que se quiere es calcular $p(Q|U)$. Como es habitual en el razonamiento de diagnóstico, primero utilizamos el teorema de Bayes para obtener:

$$p(Q|U) = k p(U|Q) p(Q), \text{ donde } k = \frac{1}{p(U)}.$$

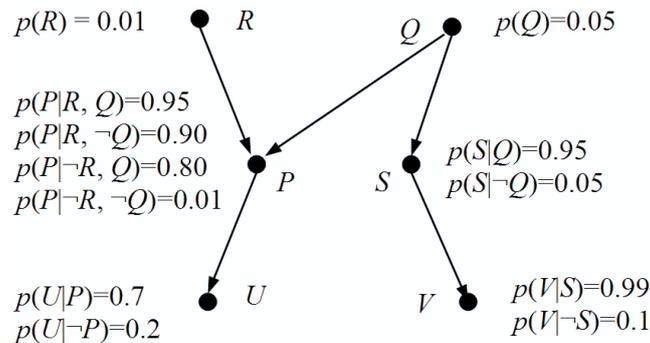


Figura 6.5: Ejemplo de poliárbol

El algoritmo descendente calcula de forma recursiva:

$$\begin{aligned}
 p(U|Q) &= \sum_P p(U|P)p(P|Q) \\
 p(P|Q) &= \sum_R p(P|R, Q)p(R) \\
 p(P|Q) &= p(P|R, Q)p(R) + p(P|\neg R, Q)p(\neg R) \\
 p(P|Q) &= 0,95 \times 0,01 + 0,8 \times 0,99 = 0,80 \\
 p(\neg P|Q) &= 0,20
 \end{aligned}$$

$$\begin{aligned}
 p(U|Q) &= p(U|P) \times 0,8 + p(U|\neg P) \times 0,2 \\
 p(U|Q) &= 0,7 \times 0,8 + 0,2 \times 0,2 = 0,60
 \end{aligned}$$

$$\begin{aligned}
 p(Q|U) &= k \times 0,6 \times 0,05 = k \times 0,03 \\
 p(\neg Q|U) &= kp(U|\neg Q)p(\neg Q) \\
 p(U|\neg Q) &= \sum_P p(U|P)p(P|\neg Q) \\
 p(P|\neg Q) &= \sum_R p(P|R, \neg Q)p(R)
 \end{aligned}$$

$$\begin{aligned}
 p(P|\neg Q) &= p(P|R, \neg Q)p(R) + p(P|\neg R, \neg Q)p(\neg R) \\
 p(P|\neg Q) &= 0,90 \times 0,01 + 0,01 \times 0,99 = 0,019 \\
 p(\neg P|\neg Q) &= 0,981
 \end{aligned}$$

$$\begin{aligned}
 p(U|\neg Q) &= p(U|P) \times 0,019 + p(U|\neg P) \times 0,981 \\
 p(U|\neg Q) &= 0,7 \times 0,019 + 0,2 \times 0,981 = 0,2095 \\
 p(\neg Q|U) &= k \times 0,2095 \times 0,95 = k \times 0,1990
 \end{aligned}$$

De esta manera, $k= 4.3668$, y finalmente,

$$p(Q|U) = 4,3668 \times 0,03 = 0,1310$$

Cuando la red no es un poliárbol, los procedimientos recursivos que acabamos de describir no finalizan, debido a los múltiples caminos que hay entre los nodos.

Conclusiones del capítulo

Se presento el estado del arte en RB, especificando la clasificación de las mismas con respecto a sus características estáticas y dinámicas. Se estudiaron los principios fundamentales en la construcción de una RB, haciendo mención: de la distribución de probabilidad conjunta, un método para construir RB, y la representación de las TPC. Se estudió ampliamente la inferencia en RB, especificando los patrones de inferencia más usados. Se estudió la separación dependiente de la dirección. Por último estudiamos el algoritmo de inferencia en poliárboles.

Podemos concluir, que las redes Bayesianas usadas para el diagnóstico de fallos, son una buena alternativa en la inferencia probabilista de fallos. Sin embargo, lo anterior esta condicionado a que el conocimiento que se tenga del sistema a diagnosticar, sea del tipo estadístico. Si el conocimiento que se tiene es del tipo vago e impreciso, se tiene que adecuar este conocimiento para que la propagación de la información a través de la red sea efectivo, tal como se propone en la sección 7.4.

Parte II

Aportaciones

Capítulo 7

Aportaciones en técnicas de Inteligencia Artificial para el diagnóstico

7.1. Introducción

En este capítulo se presentan todas las aportaciones hechas en la etapa investigativa inicial en el desarrollo de esta tesis. En el intervalo de varios años de trabajo se ha asistido a varios congresos donde se han presentado todos los avances que en lo referente a diagnóstico de fallos hemos tenido.

Estos avances giran al rededor de diferentes técnicas de Inteligencia Artificial como son sistemas expertos borrosos, diccionario de fallos, combinación de lógica borrosa con probabilidad por la generalización de la teoría de Dempster-Shafer a conjuntos borrosos, y las redes Bayesianas. En todas estas aportaciones se le ha dado un enfoque especial al trato con la incertidumbre inherente a los sistemas de diagnóstico. Este enfoque es la representación de las medidas en forma intervalar, esto es con el propósito de introducir mayor completitud en el conocimiento necesario para la toma de decisiones en los sistemas de diagnóstico.

En lo referente a información primaria para el diagnóstico tanto del tipo cuantitativo como del cualitativo, es más conveniente representar la información en forma de intervalos los cuales nos muestran la distribución de información y no tan sólo la media muestral. Además propagar la información a través del sistema en forma de intervalo, resulta en menor consumo de poder de computo, que propagar el conjunto completo de muestras de información (medidas).

En la primera sección de este capítulo, se presenta un “Sistemas expertos borrosos para el diagnóstico” (Ramírez 2002), aportación presentada en el marco del XXIV congreso internacional de electrónica ELECTRO 2002 , organizado por Instituto Tecnológico de

Chihuahua Méx. En la segunda sección se estudiará lo que es el “Diagnóstico de sistemas mediante combinación de lógica borrosa y probabilidad” (Ramírez 2004), aportación presentada en el marco de XXVI congreso internacional de electrónica ELECTRO 2004. Por último se estudiará el artículo “*Learning Bayesian Networks for Systems Diagnosis*” (Ramírez 2006), presentado en el congreso internacional CERMA 2006, organizado por la IEEE *Computer Society* en la ciudad de Cuernavaca Méx.

7.2. Diccionario de fallos borroso para el diagnóstico

En este artículo se presentó el trato de la incertidumbre inherente a los sistemas de diagnóstico representando las medidas de los sensores que es nuestra información primaria, en forma de intervalo. Resultó muy adecuado para el trato de los intervalos el uso de un diccionario de fallos que es nuestra base de conocimiento en nuestro sistema experto, que usa en su inferencia reglas del tipo *if-then* pero con proposiciones borrosas, para de esta forma introducir la gradualidad en el diagnóstico. En este sistema, se toman en cuenta las variables de contexto para evitar los fallos falsos positivos. También se introduce el diagnóstico predictivo por el uso de la tendencia de las variables. Todo lo anterior se simula a través de un programa codificado en MATLAB, con el cual pudimos hacer un análisis de resultados.

7.2.1. Introducción a un sistema de diagnóstico basado en un diccionario de fallos

Los sistemas expertos se pueden basar ya sea en conocimiento probabilístico o en conocimiento posibilístico. El conocimiento probabilístico resulta de hechos probados estadísticamente y los sistemas expertos basados en este conocimiento se dice que son usados para resolver problemas a través del razonamiento probabilístico.

Existe otra clase de conocimiento que se evalúa a través de un juicio subjetivo, que incluye información incompleta, imprecisa, o ambigua, el manejo de la incertidumbre resultante de este conocimiento conduce a un razonamiento inexacto o aproximado. Los sistemas expertos basados en este razonamiento aproximado, ponen énfasis en el punto de vista borroso o posibilístico que es una propiedad intrínseca del lenguaje natural.

Es obtenido el conocimiento, estimado por un experto tal que permite la interpretación y diagnóstico. Esta base de conocimiento es organizada en reglas de inferencia borrosas, que relacionan las medidas o resultado de pruebas, con las conclusiones a cerca de las condiciones del proceso o posibles fallos.

7.2.2. Diagnóstico lógico

Se definen las funciones de pertenencia a partir de los conjuntos borrosos, por ejemplo del conjunto borroso T1 (ver figura 7.1) tendremos tres funciones de pertenencia T1 baja, T1 normal y T1 alta

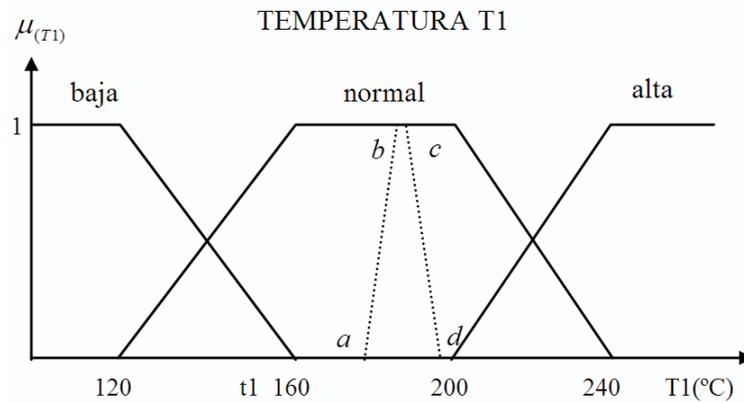


Figura 7.1: Funciones de pertenencia

las reglas de inferencia borrosas a usar son del tipo de enunciados *if-then* pero que usan proposiciones borrosas (Driankov 1996; Jang 1997). Como por ejemplo:

SI la “temperatura 2” (T2) es alta en un 75% y la “temperatura 1” (T1) y la “presión 1” (P1) son normales, *ENTONCES* se produce el fallo 4 con una gradualidad del 75%

Estas reglas de inferencia se organizan en un diccionario de fallos construido en base a las funciones de pertenencia de cada una de las medidas que se monitorean en el sistema. En este diccionario a cada función de pertenencia se le asignan límites mínimos y máximos, como se ve en los cuadros 7.1 y 7.2, donde se muestran sólo algunos de los posibles fallos únicos que se podrían presentar en determinado sistema.

Las reglas de inferencia que después se siguen, son la comparación de los valores que las medidas producen en cada función de pertenencia ($\mu_{T1_b}(t1)$), con los límites que cada fallo tiene para que este ocurra, en cada una de las funciones de pertenencia. Si el *Error* entre los valores que las medidas producen y los valores del fallo es cero, entonces el fallo ocurre con una severidad del 100 %, esto es, coinciden las medidas con los síntomas del fallo en su totalidad. Las reglas son de la siguiente forma:

Si $\mu_{T1_b}(t1)$ es menor o igual que el límite máximo de $\mu_{F.T1_b}(t1)$ y mayor o igual que el límite mínimo de $\mu_{F.T1_b}(t1)$, *entonces* el *Error* es cero.

SI $\mu_{T1_b}(t1)$ no se encuentra entre los límites máximos y mínimos de $\mu_{F.T1_b}(t1)$, *ENTONCES* el *Error* es igual a la diferencia entre el límite más próximo de $\mu_{F.T1_b}(t1)$ a $\mu_{T1_b}(t1)$.

Estas reglas se aplican a cada función de pertenencia de cada medida. En este caso cada medida tiene 3 funciones de pertenencia (Figura 7.1) por lo que hacen un total de 9 funciones de pertenencia (según se aprecia en los cuadros 7.1 y 7.2) con lo cual las reglas anteriores producirían 9 errores. Para obtener la severidad del fallo se obtiene el *Error* máximo de los 9 errores producidos por las reglas y se aplica la siguiente fórmula:

$$\text{Severidad de fallo} = 1 - \text{Error máximo}$$

Cuadro 7.1: Diccionario de fallos (límites mínimos)

LIMITES MÍNIMO									
	$\mu_{F_{T1}}(t1)$			$\mu_{F_{T2}}(t2)$			$\mu_{F_{P1}}(p1)$		
fallo	baja	normal	alta	baja	normal	alta	baja	normal	alta
1	0	0	0	1	0	0	1	0	0
2	0	0	1	0	0	0	0	0	1
3	0	1	0	1	0	0	0	1	0
4	0	1	0	0	0	1	0	1	0

Cuadro 7.2: Diccionario de fallos (límites máximos)

LIMITES MÁXIMOS									
	$\mu_{F_{T1}}(t1)$			$\mu_{F_{T2}}(t2)$			$\mu_{F_{P1}}(p1)$		
fallo	baja	normal	alta	baja	normal	alta	baja	normal	alta
1	1	1	1	1	0	0	1	0	0
2	0	0	1	1	1	1	0	0	1
3	0	1	0	1	0	0	0	1	0
4	0	1	0	0	0	1	0	1	0

Resumiendo el concepto del diagnóstico lógico presentado: consiste en un diccionario de fallos compuesto por dos tablas, que contienen la información de los fallos en base a las funciones de pertenencia de cada medida. En una tabla se describen los límites mínimos y en otra los límites máximos, en los que se considera que un fallo existe para cada medida, si una medida cae en ese intervalo, entonces diremos que el fallo se ha presentado con una severidad del 100 %, de lo contrario, dependerá de lo lejos que la medida se encuentre, del límite más próximo del intervalo que la regla estipula en nuestro diccionario de fallos.

7.2.3. Imprecisión en las medidas

Los sensores usados en los procesos industriales generalmente son fuentes de incertidumbre, dependiendo de la calidad de estas las medidas proporcionadas por ellos tendrán

mayor o menor grado de imprecisión. En este programa de diagnóstico las medidas se representan por un conjunto borroso trapezoidal definido por los puntos paramétricos a, b, c, d , como se muestra en la figura 7.1, donde en el peor caso la medida se podría encontrar entre los puntos a y d (soporte de la función de pertenencia) y en el mejor de los casos entre b y c (núcleo) con los cuales se proporcionan dos intervalos porcentuales en los posibles fallos al operador del sistema, que son: los fallos cuando se considera la incertidumbre en las medidas en el mejor de los casos (entre b y c) y los fallos cuando se considera la incertidumbre en las medidas en el peor de los casos (entre a y d).

Como generalmente cada sensor arrojará un intervalo en la medida, cada medida tendrá su valor máximo y su valor mínimo por lo que, para cada medida tendremos dos vectores de medidas el de valores máximos y el de valores mínimos, estos valores se acercarán entre si según la incertidumbre introducida por los sensores disminuya. Las reglas de inferencia aplicadas tanto a los valores máximos como a los valores mínimos de los vectores de medidas, son las mismas enunciadas en la sección anterior, con esto el diagnóstico de un fallo refleja la incertidumbre introducida por los sensores, por lo que un diagnóstico podría ser, que el fallo 4 ocurre con una severidad entre el 60 y 70 %.

7.2.4. Variables de contexto

Con el fin de tomar en cuenta las variaciones introducidas a los parámetros de las funciones de pertenencia, por los cambios en las variables de contexto como las perturbaciones, que en este caso son la temperatura y la presión ambiental, se introduce la matriz de sensibilidad (Slamani 1992) definida de la siguiente manera:

$$S_x^T = \frac{\Delta T}{\Delta x} | \Delta x \rightarrow 0$$

donde: ΔT es la variación del parámetro T , Δx es la variación de valor en el elemento de contexto y S_x^T es la matriz de sensibilidad, en otros acercamientos las variaciones de los parámetros se consideran relativas a los valores nominales de T y x .

En este programa, si la temperatura ambiente cambia con respecto a la nominal, tendremos Δx_{Ta} lo cual supone un ligero cambio en los parámetros de las medidas, sin embargo, esto no implica que un fallo este ocurriendo. Para evitar el diagnóstico de un fallo no existente (fallo falso positivo), se ajustan los valores de los parámetros de las funciones de pertenencia, en base a la matriz de sensibilidad. En la simulación del proceso descrito en la sección 7.2.6 se compone de 8 parámetros por medida (ver figura 7.1) y el cálculo de los valores (Liu 1994) se rige por la siguiente ecuación.

$$\begin{bmatrix} \Delta T1 \\ \Delta T2 \\ \Delta P1 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \\ S_{31} & S_{32} \end{bmatrix} \begin{bmatrix} \Delta Ta \\ \Delta Pa \end{bmatrix}$$

donde S_{ij} es un vector columna de 8 filas (número de parámetros por medida) ΔTa y ΔPa son los cambios ambientales en temperatura y presión respectivamente. Estas

matrices S_{ij} pueden ser obtenidas heurísticamente por el experto de diagnóstico, superponiendo los resultados obtenidos con la variación de la temperatura ambiente, con los resultados de la variación de la presión ambiental para determinar la variación total de los parámetros de las funciones de pertenencia.

7.2.5. Diagnóstico predictivo a través de la tendencia de las variables

Podemos aprovechar la capacidad computacional de los ordenadores, para determinar con cierta precisión la tendencia de las variables en el diagnóstico de un proceso industrial.

La derivada de una variable con respecto al tiempo en realidad es la pendiente de la variable medida, por lo que, podemos basarnos en la pendiente para entender la deriva de una señal o tendencia de una variable con respecto al tiempo. Si tomásemos la última pendiente producida por la señal, y dijésemos, esta será la tendencia de la variable en el futuro próximo, tal vez podríamos equivocarnos si la señal tuviese ciertas oscilaciones como es común en las medidas obtenidas en un sistema real

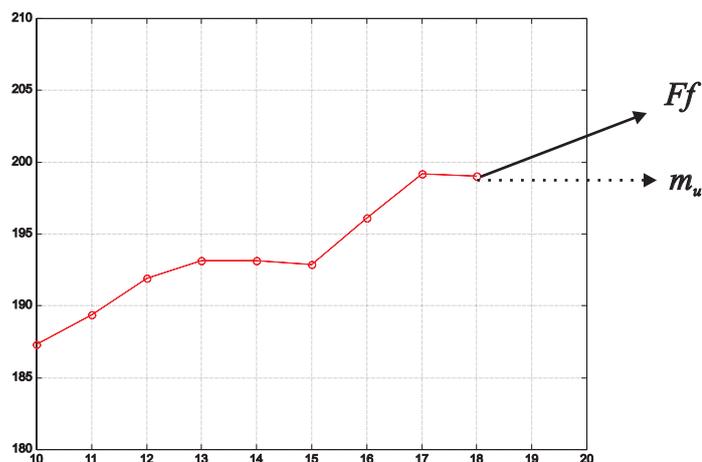


Figura 7.2: Pendientes m_u y Ff en la tendencia de una variable

por lo que, no se lograría el objetivo de predecir la tendencia de una variable, tal y como se demuestra en la figura 7.2.

La manera propuesta en este programa de diagnóstico es el uso de un filtro aplicado a las pendientes de la siguiente forma:

$$Ff = c_u m_u + c_{u-1} m_{u-1} + c_{u-2} m_{u-2} + \dots + c_{u-n} m_{u-n}$$

donde Ff es el filtro final o pendiente final, c_u es la ponderación para cada pendiente, y m_u es la última pendiente. Lo que se hace con este filtro es obtener una pendiente final

que es una extrapolación de todas las pendientes incluidas en el filtro, de esta manera la pendiente obtenida proporciona una dirección más cierta, porque se basa no sólo en la última pendiente si no en la historia de las pendientes, que pueden ser la segunda y tercera inmediata anterior (o alguna otra más).

Para esta simulación se usan las últimas tres pendientes obteniéndose buenos resultados para el tipo de señales propuestas. Aplicando esta técnica al ejemplo anterior, mientras la última pendiente es negativa la pendiente obtenida por el filtro marca una tendencia predictiva más adecuada según se puede ver en la figura 7.2.

Los valores de $c_u, c_{u-1}, \dots, c_{u-n}$ se proponen según el tipo de señal que produzca la variable, donde el valor máximo se le da a c_u puesto que es la última pendiente y es la de mayor importancia. La ponderación de las constantes debe seguir esta regla:

$$c_u + c_{u-1} + c_{u-2} + \dots + c_{u-n} = 1$$

Como la naturaleza del ordenador es discreta las muestras se toman por periodos de muestreo. Aquí la predicción se hace por periodos de muestreo en adelante. Una vez obtenida la pendiente final Ff para el punto de muestreo actual, se hace la predicción según esta pendiente incluida en la siguiente expresión

$$x(k+n) = n \times \Delta k \times Ff + x(k)$$

donde n es el número de periodos de muestreo de predicción, Δk es el valor del periodo de muestreo, y $x(k)$ es el valor en el punto actual de muestreo. Con lo anterior se puede predecir la tendencia de las variables con el número de periodos de muestreo en adelante deseados.

7.2.6. Simulación de un proceso

Se simula el diagnóstico de una planta industrial como la mostrada en la figura 7.3, donde la salida de tres sensores servirán como medida paramétricas. Del tanque de almacenamiento de vapor tendremos $T1$ y $P1$ (temperatura y presión) y de la caldera tendremos $T2$. Se consideran las siguientes condiciones de operación del sistemas : La presión en la caldera es mayor que la presión $P1$ de la cámara de vapor y $T2$ es mayor que $T1$, $R1$ regula el gasto de vapor , $R2$ regula $T1$ y $P1$, $R3$ y $R4$ regulan $T2$. El gasto de vapor es más o menos constante y las perturbaciones N pueden ser la temperatura ambiental (Ta) y la presión atmosférica (Pa). El programa se encarga sólo del diagnóstico y no del control de los parámetros.

Para cada medida se generan 3 funciones de pertenencia como las mostradas en la figura 7.1. El diccionario de fallos será el de los cuadros 7.1 y 7.2. Suponiendo que la $T2$ se presenta con una tendencia alcista esto supondría el fallo 4 que resulta de la avería de la válvula reguladora $R4$. Teniendo en cuenta que en el mejor de los casos los sensores presentan una variación de ± 2 unidades de medida y que los parámetros de las funciones de pertenencia se ajustan por pequeñas variaciones en la temperatura y presión ambiental, se obtienen los siguientes vectores de medidas

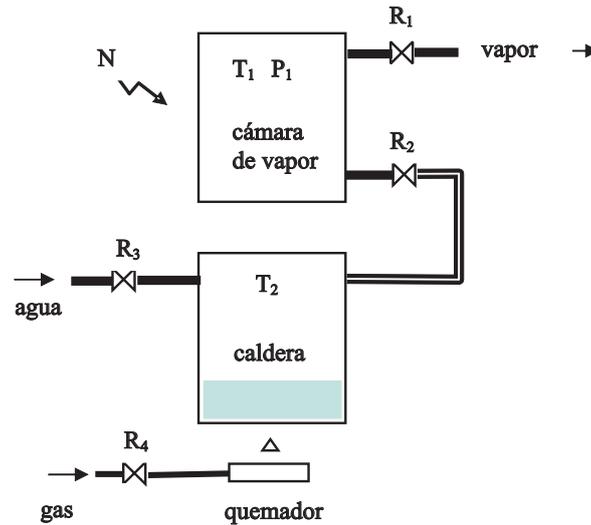


Figura 7.3: Proceso industrial a simular

μ_{T1}			μ_{T2}			μ_{P1}		
baja	normal	alta	baja	normal	alta	baja	normal	alta
0	1	0	0	0.83	0.26	0	1	0
0	1	0	0	0.73	0.16	0	1	0

la primera fila corresponde a las medidas máximas y la segunda fila al de medidas mínimas, se aplican las reglas de inferencia de la sección 2 con respecto a la cuadros 7.1 y 7.2 resultando los siguientes vectores de error donde la primera fila corresponde a los errores máximos de cada función de pertenencia y el de la segunda fila a los errores mínimos

$Error_{T1}$			$Error_{T2}$			$Error_{P1}$		
baja	normal	alta	baja	normal	alta	baja	normal	alta
0	0	0	0	0.83	0.84	0	0	0
0	0	0	0	0.73	0.74	0	0	0

por lo tanto obteniendo los valores máximos de cada uno de los vectores de error anteriores se calcula el intervalo del fallo 4 que en este caso ocurre en el periodo 14 de muestreo y esta entre el 16 y 26 % como se muestra en la figura 4.

$$\text{Máxima severidad de fallo} = 1 - Error_{2,6} \text{ cuando } k=14$$

$$\text{Mínima severidad de fallo} = 1 - Error_{1,6} \text{ cuando } k=14$$

Ahora con respecto a la tendencia de las variables si consideramos las 3 pendientes anteriores al periodo de muestreo 14, en el caso de las medidas máximas tendríamos:

$$Ff = c_u m_u + c_{u-1} m_{u-1} + c_{u-2} m_{u-2}$$

$$Ff = (0,5 \times 0,51) + (0,3 \times 0,24) + (0,2 \times 0,71)$$

$$Ff = 0,4690$$

de continuar esta tendencia en los 5 periodos de muestreo próximo ($n = 5$) dado que la temperatura máxima es de 291 °C para $k = 14$, la temperatura incrementaría hasta 293 °C para cuando k fuese igual a 19 periodos de muestreo, según lo muestra el siguiente cálculo.

$$\begin{aligned}
 x(k + n) &= n \times \Delta k \times Ff + x(k) \\
 x(k + 5) &= 5 \times 1 \times 0,4690 + 291 \\
 x(k + 5) &= 293,34
 \end{aligned}$$

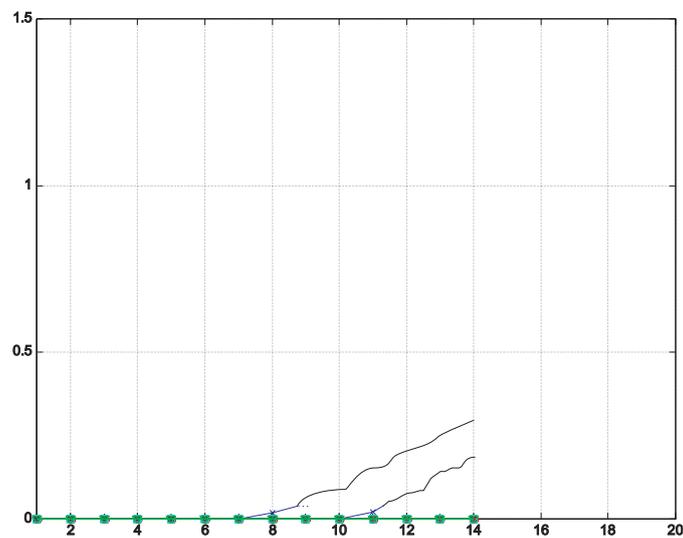


Figura 7.4: Fallo 4 cuando $k = 14$

7.2.7. Sumario sobre el sistema experto borroso

Se ha planteado una manera de resolver la incertidumbre en las medidas, como hacer el diagnóstico a través de reglas de inferencia borrosas organizadas en un diccionario de fallos, como evitar el diagnóstico de fallos falsos positivos por el uso de la matriz de sensibilidad, y como estimar la tendencia de las variables a través de un filtro sobre las pendientes. En la siguiente aportación publicada, se usará la técnica de diagnóstico lógico vista en esta sección, pero añadiremos la combinación de lógica borrosa con probabilidad, con lo que se busca incorporar más flexibilidad al sistema de diagnóstico.

7.3. Diagnóstico de sistemas mediante combinación de lógica borrosa y probabilidad

Para ejemplificar esta técnica se presenta el diagnóstico de una plantas industriales, que caracterizan la información provista por sus sensores en forma probabilística, pero que también dependen de otras variables de diagnóstico, que por su naturaleza, no son cuantitativas sino cualitativas y que por lo tanto, se representan en forma borrosa, de allí la necesidad de combinar lógica borrosa con probabilidad. En este trabajo, para demostrar la combinación de la lógica borrosa con la probabilidad, que se efectúa a través de la generalización de la teoría de evidencia, se simula un sistema de diagnóstico para dos depósitos de combustible interconectados.

7.3.1. Aplicaciones industriales de la teoría de evidencia

La teoría de evidencia o teoría de Dempster-Shafer se originó en los años 60 (Dempster 1967; Shafer 1976). En este trabajo en el capítulo 5 se ofrece una explicación de la misma. Si se quiere ahondar en esta teoría se puede consultar (Caro 1999).

En la actualidad aplicaciones que combinan lógica borrosa con probabilidad existen pocas. En el caso del diagnóstico de plantas industriales tenemos en (Lardon 2004) un trabajo que trata sobre el diagnóstico de un proceso de digestión anaeróbica, que se usa para la producción de biogás. Algunas de las variables de diagnóstico en este tipo de procesos son borrosas y dependen de la opinión del operador humano, como son la textura, color, olor y concentración de la biomasa, otras dependen de sensores y pueden ser probabilísticas, como en este caso, son el caudal de entrada de la biomasa, el caudal de salida del biogás, el pH y la concentración de hidrógeno por mencionar algunas.

Lo que se hizo en este trabajo de digestión anaeróbica, es borrosificar todas las variables de diagnóstico e interpretar todas estas medidas borrosas, como las asignaciones básicas de probabilidad utilizadas en la combinación evidencial, que se describe en el apartado 5.3.2. En (Yong 2003) puede consultarse un esquema similar usando redes neuronales para determinar las asignaciones básicas de probabilidad.

7.3.2. Sistema que combina lógica borrosa y probabilidad

En la prueba de la generalización de D-S a conjuntos borrosos, se simula a través de MATLAB un sistema con tablas de decisión borrosas en el diagnóstico de fallos, de un sistema de depósitos de combustible interconectados como el de la figura 7.5.

Modelo matemático de planta

La idea de este modelo de planta para diagnóstico fue tomado de (Sala 2000). Este modelo consiste en dos depósitos de combustible interconectados entre si, como se muestran

en la figura 7.5. Las ecuaciones principales en el sistema de depósitos son las siguientes:

$$\begin{aligned} q_1 &= 2 \frac{(h_1 - h_2)}{R_1}, \quad q_2 = 2 \frac{h_2}{R_2}, \quad qf_1 = 2 \frac{h_1}{Rf_1}, \quad qf_2 = 2 \frac{h_2}{Rf_2} \\ dh_1 &= \frac{1}{C_1}(q - q_1 - qf_1)dt, \quad dh_2 = \frac{1}{C_2}(q_1 - q_2 - qf_2)dt \end{aligned} \quad (7.1)$$

Aquí R es la resistencia para el líquido a través de un tubo y C es capacitancia de un depósito, q y h son respectivamente una desviación pequeña en el caudal y altura de su valor nominal, qf es considerado el caudal de fuga en los depósitos

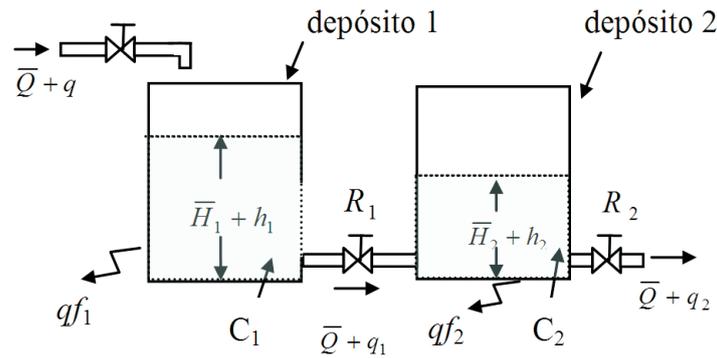


Figura 7.5: Sistema de dos depósitos interconectados

Diagnóstico lógico

El diagnóstico lógico se hace a través de un sistema de tablas de decisión borrosas (Ramírez 2002), que obtiene su base de conocimiento de un diccionario de fallos determinado experimentalmente (cuadros 7.3 y 7.4), por el uso del modelo matemático (7.1).

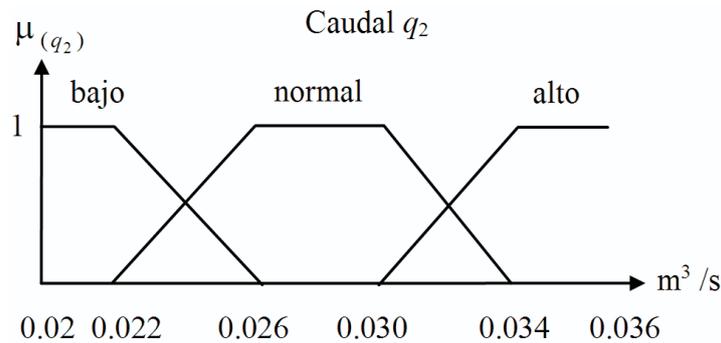


Figura 7.6: Funciones de pertenencia para q_2

Partiendo de los valores de punto de equilibrio nominal de las variables del sistema de depósitos $\bar{Q} = 0,02 \text{ m}^3/\text{s}$, $\bar{H}_1 = 2\text{m}$, $\bar{H}_2 = 1\text{m}$, se definen las funciones de pertenencia usadas en la inferencia lógica del sistema de diagnóstico borroso, tal como se hace para

el caudal de salida q_2 en la figura 7.6. Los cuadros 7.3 y 7.4 muestran parte de un diccionario de fallos, donde se indican los límites mínimos y máximos de las funciones de pertenencia propuestas para cada variable de diagnóstico a evaluar, en el diagnóstico de fallos.

Cuadro 7.3: Límites mínimos en el diccionario para diagnóstico del sistema de depósitos

LIMITES MÍNIMO									
	$\mu(q_2)$			$\mu(h_1)$			$\mu(h_2)$		
fallo	baja	normal	alta	baja	normal	alta	baja	normal	alta
2	0	0	0.8	0	0	0.9	0	0	1
3	0	1	0	0	0	1	0	1	0

Cuadro 7.4: Límites máximos en el diccionario para diagnóstico del sistema de depósitos

LIMITES MÁXIMOS									
	$\mu(q_2)$			$\mu(h_1)$			$\mu(h_2)$		
fallo	baja	normal	alta	baja	normal	alta	baja	normal	alta
2	0	0	1	0	0	1	0	0	1
3	0	1	0	0	0	1	0	1	0

En este diccionario de fallos que constituye los límites de las reglas de inferencia, las funciones de pertenencia para cada conjunto borroso son: bajo, normal y alto según lo indica en los cuadros 7.3 y 7.4. En este sistema de diagnóstico se consideran las medidas provenientes de los sensores o de observadores analíticos como intervalos de creencia $[bel, pl]$, también se considera que los límites que observan las reglas son intervalares, como se explica a través de los cuadros 7.3 y 7.4.

Los límites mínimos y máximos de las reglas de inferencia que encontramos en este diccionario, son útiles, pues nos ayudan afinando la delimitación de fallos. Así para el fallo 2 si $\mu(q_2)_{alta} = 0,8$, $\mu(h_1)_{alta} = 0,9$, $\mu(h_2)_{alta} = 1$, entonces el fallo se presentará en un 100%, esto se debe a que los límites de $\mu(q_2)_{alta}$ en el diccionario de fallos están entre $[0,8, 1]$ y los de $\mu(h_1)_{alta}$ entre $[0,9, 1]$ $\mu(h_2)_{alta}$ no tiene intervalo. Una interpretación lingüística del diccionario de fallos sería:

SI fallo 2 *ENTONCES* q_2 , h_1 , y h_2 son altos.

SI fallo 3 *ENTONCES* q_2 , y h_2 son normal y h_1 es alto.

El *Error* o diferencia entre el intervalo de creencia y el intervalo que la regla estipula, nos dirá que tan lejos o que tan cerca estamos de diagnosticar un fallo, de hecho es obvio que si no hay una diferencia o *Error* entre las medidas de creencia y las condiciones necesarias para diagnosticar un fallo que la regla indica, diremos coherentemente que estamos ante un fallo. En la ilustración de la figura 7.7, se muestra como se conciben los errores en este sistema de creencias y reglas intervalares.

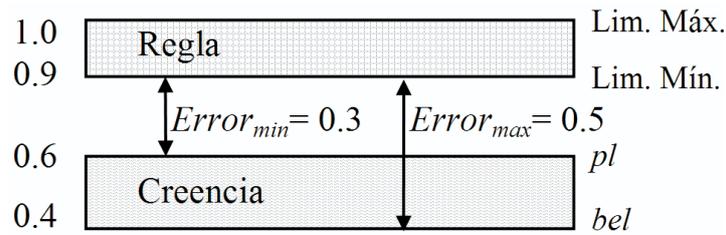


Figura 7.7: Errores entre la creencia y la regla

En la figura 7.7 se observa que la distancia mínima ($Error_{min} = 0,3$) y la distancia máxima ($Error_{max} = 0,5$) entre la creencia y el límite inferior de la regla (que para este caso es el más cercano a la creencia) nos define la severidad del fallo que estaría entre un 50 y un 70 %. Esto se obtiene restándole un 1 a los errores.

Conversión de conjuntos borrosos a probabilidad

Una de las características del ser humano es de valerse de diferentes fuentes de información, cuando no esta seguro de poder tomar una decisión, así la teoría de evidencia de D-S, se vale de la fusión de información para disminuir la incertidumbre. Debido a la inflamabilidad los sensores de nivel en los depósitos de combustible (Thangaraj 1994), son del tipo no invasivo y presentan alguna clase de incertidumbre. En este trabajo se propone fusionar dos sensores, para medir el nivel de combustible y fusionar otros dos, para medir el caudal de salida.

La fusión se llevará acabo a través de la generalización de la teoría de D-S a conjuntos borrosos (Yen 1990). Así h_1 -ultrasónico y h_2 -ultrasónico por ser de tipo más preciso, serán asignaciones básicas de probabilidad $m(A)$ (5.19). Estas medidas se borrosifican por las funciones de pertenencia de nuestro sistema de diagnóstico borroso (ver figura 7.6). Luego h_1 -visión y h_2 -visión, que son sensores de visión por ordenador, debido a su naturaleza vaga o difícil de interpretar, se caracterizarán en forma trapezoidal borrosa (ver figura 7.8).

Detalles sobre cómo caracterizar este tipo de sensores se encuentran en (Chakravarthy 2002; Li 2002). También en los sensores de caudal son usados los de tipo ultrasónico (Bohs 1994), por lo tanto nuestras últimas variables de diagnóstico serán: q_2 -ultrasónico y q_2 -presión, que se caracterizan cómo asignación básica de probabilidad y trapezoidal borrosa respectivamente.

En la conversión de conjuntos borrosos a probabilidad, según las expresiones (5.19) vistos en la seccion 5.4, se encuentran involucradas las funciones de pertenencia A y B que se convierten a no borrosas, por el uso de la resolución dada por los niveles de corte α_i , según se puede apreciar en la figura 7.8. Desde el momento que dividimos las funciones de pertenencia borrosas por los niveles de corte α_i , se convierten en focales no borrosos, por lo que cada focal, no será considerado como medida borrosa sino probabilística.

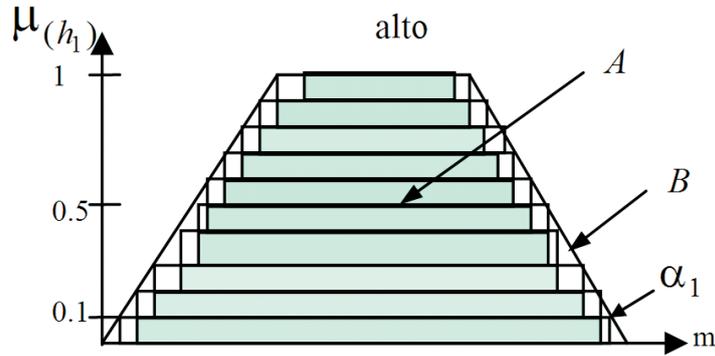


Figura 7.8: Funciones de pertenencia no borrosas A y B

En la figura 7.8 el focal no borroso A (color oscuro), corresponde a la opinión del sistema de diagnóstico que define al nivel h_1 como alto, el focal no borroso B , corresponde a h_1 -visión.

Para obtener la credibilidad $bel(B)$ (5.19), el resultado de cada $[\alpha_i - \alpha_{i-1}]$ se multiplica por cada ínf $\mu_B(x)$ correspondiente, el resultado de esta sumatoria se multiplica por la asignación básica de probabilidad $m(A)$ que para nuestro caso sería $\mu(h_1\text{-ultrasónico})$. La condición para obtener la cantidad borrosa ínf $\mu_B(x)$ es que $x \in A_{\alpha_i}$ (Caro 1999; Yen 1990) en este caso $x = m$. Obteniendo $bel(B)$ correspondiente a la figura 7.8 donde por simplicidad $\text{ínf}_{x \in A_{\alpha_i}} \mu_B(x)$ es denotada como $f_{B,A}(\alpha_i)$ y $m(A) = 1$ tendríamos:

$$\begin{aligned}
 bel(B) &= m(A) \times [0,1 \times f_{B,A}(0,1) + 0,1 \times f_{A,B}(0,2) + \dots] \\
 bel(B) &= 1 \times [0,1 \times 0,2 + 0,1 \times 0,3 + 0,1 \times 0,4 + 0,1 \times 0,5 \\
 &\quad + 0,1 \times 0,6 + 0,1 \times 0,7 + 0,1 \times 0,8 + 0,1 \times 0,9 + 0,1 \times 1] \\
 bel(B) &= 0,64
 \end{aligned}$$

Para obtener la plausibilidad $pl(B)$ (5.19), el resultado de cada $[\alpha_i - \alpha_{i-1}]$ se multiplica por cada sup $\mu_B(x)$ correspondiente, el resultado de esta sumatoria se multiplica por la asignación básica de probabilidad $m(A)$ que para nuestro caso sería $\mu(h_1\text{-ultrasónico})$.

En la figura 7.9 se muestra la aparición de varios fallos que ocurren en los depósitos de combustible. El fallo 2 y 3 que ahí se presentan, se describen en los cuadros 7.3 y 7.4. En la simulación de estos fallos partimos de condiciones normales, luego aumentamos gradualmente (forma exponencial) el caudal de entrada q de nuestro modelo de depósitos con esto todas las variables de diagnóstico suben gradualmente lo que ocasiona que se produzca el fallo 2, que en su forma incipiente también se parece al fallo 3 por lo que también aparece inicialmente junto con el fallo 2 como podemos ver en la gráfica de la figura 7.9. Con el fin de hacer más demostrativa esta técnica, entre el segundo 1400 y 2100 tan sólo estamos midiendo las asignaciones básicas de probabilidad (h_1 -ultrasónico, h_2 -ultrasónico, q_2 -ultrasónico) con periodo $T=10$ seg. esto significa que el periodo de muestreo para h_1 -visión, h_2 -visión, q_2 -presión es de 700 seg.

Debido a que éstas variables tienen un componente de ruido aleatorio, se observan perturbaciones, donde en este intervalo de tiempo, sólo observamos la plausibilidad de

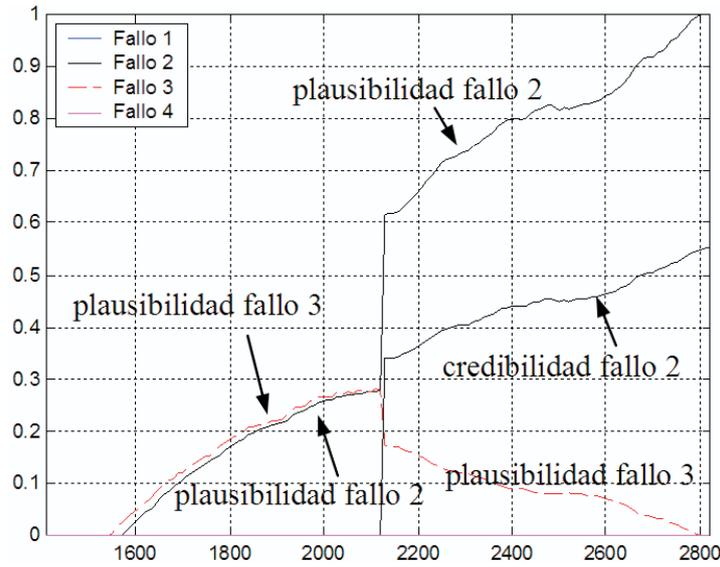


Figura 7.9: Gráfica de credibilidad y plausibilidad

ambos fallos, siendo sus respectivas credibilidades iguales a cero. Para el segundo 2100 ya se ha confirmado la tendencia de las variables, es entonces cuando volvemos a leer las medidas trapezoidales borrosas (h_1 -visión, h_2 -visión, q_2 -presión), permitiendo su fusión. Para este punto ya se han confirmado las tendencias de las variables. La evidencia que proporcionan estos sensores (focales no borrosos B), es que todas las variables se encuentran altas. Vemos en el segundo 2100 como al fusionar las nuevas evidencias trapezoidales borrosas, la tendencia del fallo 2 se confirma, y podemos observar su credibilidad y su plausibilidad. La plausibilidad del fallo 3 empieza a desaparecer, confirmando que el fallo que en realidad está ocurriendo es el fallo 2. Si fusionásemos más fuentes evidenciales a las ya existentes, la incertidumbre con respecto a la tendencia disminuiría aún más.

El efecto que logramos observar en esta simulación, es que en el intervalo entre el segundo 1400 y 2100, aunque existiese ausencia, fallo o intermitencia en un sensor, con todo el sistema sería capaz de detectar y diagnósticos los fallos. Una ventaja de este sistema, es que al fusionar sensores del tipo trapezoidal borroso, no agregamos al comportamiento del fallo más ruido aleatorio. Otra de las ventajas que la teoría de evidencia ofrece, son dos puntos de vista complementarios para la toma de decisiones: la medida de creencia $[bel, pl]$ tratada en esta sección, que son medidas probabilísticas que reflejan los valores reales de nuestras variables de diagnóstico, y la combinación evidencial que será tratada en el próximo apartado.

7.3.3. Análisis de resultados por combinación evidencial

La combinación evidencial es una medida normalizada con respecto al marco de discernimiento que los fallos presentes ofrecen, la cuál distribuye de forma proporcional, las probabilidades correspondientes a cada uno de ellos, de manera que por comparación

nos demos una idea más clara, de cuál es el fallo que en realidad esta ocurriendo. Esta técnica se estudio en la sección 5.3.2.

Introducimos la combinación evidencial de Dempster-Shafer (5.6) y no la combinación dada en la generalización de conjuntos borrosos a D-S (5.26), esto es debido a que para este punto, todas las medidas fueron convertidas en asignaciones básicas de probabilidad $bel(B)$ y $pl(B)$ que reflejan las medidas proporcionadas por los sensores del sistema.

Cada fuente evidencial estará compuesta por dos elementos focales, esto es, los fallos detectados serán un elemento focal y el complemento será el otro elemento focal. Cada fallo diferente (nueva combinación) será una nueva fuente evidencial conformada por dos elementos focales, los fallos detectados y su complemento. El resultado obtenido de la combinación evidencial es desplegado de la siguiente manera por el sistema:

Tipo de fallo	Fallos	Plausibilidad
único	{2}	0.9983
único	{3}	0.0001
único	{5}	0.0012
múltiple	{1,2,6,7,8}	0.0003

El cuadro anterior también corresponde al resultado de los fallos presentes en la gráfica de la figura 7.9. Aquí sólo se presenta la plausibilidad pero también puede presentarse la credibilidad. El cuadro anterior fue desplegado al segundo 2800, sí nos damos cuenta, los resultados son equivalentes a los de la figura 7.9 donde el fallo confirmado es el 2 con una plausibilidad del 99.83%.

La razón por lo que esta técnica funciona, es que a medida que las tendencias de los conjuntos de fallos se clarifican, los conjuntos de fallos que son menos probables de ocurrir, tienden hacia una plausibilidad igual a cero como podemos observar en el cuadro. Los resultados obtenidos por la combinación evidencial, nos confirma coherentemente que los obtenidos por la función de creencia $[bel, pl]$ (figura 7.9), son correctos.

7.3.4. Sumario de la combinación de lógica borrosa y probabilidad

En este trabajo hacemos una conversión estricta entre lógica borrosa y probabilidad, que nos da más exactitud debido a la mayor fusión de evidencias, mientras otros enfoques (Lardon 2004; Yong 2003) tan sólo usan la combinación evidencial, por lo cuál, sólo tienen un punto de vista, mientras nosotros tenemos dos equivalentes y complementarios: la función de creencia y la combinación evidencial. Probamos por simulación, la eficacia de la generalización de la teoría de D-S a conjuntos borrosos, en la disminución de la incertidumbre en la evaluación de fallos. Representamos la asignación básica de probabilidad como una cantidad borrosa, usada para obtener nuestra función de creencia y combinación evidencial. También adaptamos la función de creencia a la lógica intervalar

borrosa ya existente. En la siguiente sección, también hacemos uso de la combinación de lógica borrosa con probabilidad, esto con el fin de transportar nuestra inferencia de diagnóstico a un entorno totalmente probabilístico, como son las redes Bayesianas.

7.4. Redes Bayesianas en el diagnóstico de sistemas con aprendizaje de parámetros

En esta aportación se presenta la construcción de una red Bayesiana para el diagnóstico de fallos en un sistema industrial. La construcción se basa en el modelo matemático de planta e incluye aprendizaje de parámetros y estructura de red Bayesiana a través de las distribuciones Beta Dirichlet (Haider 2003). Experimentamos la anterior metodología por medio de un caso de estudio que incluye el diagnóstico de fallos con los cuales fue entrenada la red. En esta aportación tratamos a la incertidumbre existente en las variables de diagnóstico desde el enfoque probabilístico y borroso.

Existen muchos campos de aplicación para redes Bayesianas, como minería de datos, diagnóstico médico, análisis de software y otros más, en esta investigación enfocaremos su aplicación a diagnóstico de fallos.

Los avances en los últimos años en el campo de redes Bayesianas aplicadas a diagnóstico de fallos, plantean sistemas de diagnóstico (Sheppard 2003), que caracterizan la incertidumbre de los sensores a través de técnicas tomadas de la ciencia de mediciones, que son técnicas probabilistas. Finalmente en la etapa de evaluación de fallos utilizan redes Bayesianas, con lo que le dan un tratamiento probabilista al manejo de la incertidumbre de principio a fin.

Algunos trabajos como en (Liu 2002), validan la información de los sensores por técnicas de estimación residual. En caso de ocurrir un fallo, entonces en el análisis de residuos del componente que fallo, se le asigna probabilidad a dicho componente, y por último aplican redes Bayesianas para el diagnóstico. Acercamientos como los de (Mehrambod 2003), caracterizan los sensores como nodos de una red, donde cada sensor-singular esta formado por tres nodos de entrada: bias, ruido, y valor de la variable, que son padres de un nodo de salida hijo, que contiene las probabilidades condicionales *a priori* de sus antecesores. Existen también algunos trabajos de aplicación dedicados a diagnóstico con redes Bayesianas, como los de (Chen-Fu 2002) en redes de distribución de energía eléctrica, o los de (Mast 1999) en turbinas de aeroplano.

7.4.1. Caso de estudio

Se simula a través de MATLAB un sistema de depósitos interconectados como el de la figura 7.5. Las ecuaciones principales que rigen el comportamiento del sistema de depósitos son (7.1) vistas en la sección 7.3.

En las ecuaciones que describen el comportamiento dinámico de los depósitos, R es la resistencia para el líquido a través de un tubo y C es capacitancia de un depósito, q y h son respectivamente una desviación pequeña en el caudal y altura de su valor nominal. Al variar el caudal q de cada válvula se simularán los diferentes fallos, donde las causas serán fallos en las válvulas Va , Vi , Vs , y las evidencias que conducen a los fallos serán los niveles en q_2 , h_1 y h_2 .

7.4.2. Descripción breve del programa

El programa de simulación para este caso de estudio está hecho bajo el paradigma de programación estructurada y las partes principales que lo conforman son: 1) el muestreo de datos a partir del modelo matemático (7.1), 2) aprendizaje supervisado de estructura de red, 3) aprendizaje paramétrico y 4) el diagnóstico de fallos a partir de los parámetros aprendidos.

Muestreo de datos

Este proyecto es visto desde el punto de vista binomial, por lo que el muestreo se lleva a cabo por niveles de umbral, donde se determina cuál variable observa o no observa fallo. Las variables que se observarán en el muestreo al variar la q de cada válvula serán: q_2 , h_1 , h_2 , R , R_1 , y R_2 . Al variar la resistencia de las válvulas más allá de un punto de referencia (*set point*) dado, significará que están atascadas o averiadas, de manera que permiten un caudal mayor o menor que lo normal.

Se simulan todos los fallos posibles que el sistema vaya a diagnosticar (por simplicidad se simularán 7 pero pudiesen ser más); Estos pueden ser fallos únicos, como lo son: válvula Vs atascada, válvula Vi atascada, válvula Va dañada. También pueden ser diferentes combinaciones de fallos múltiples, como son: válvulas Vi y Vs están atascadas, válvula Va dañada al momento que Vs está atascada, válvula Va dañada al mismo tiempo que Vi está atascada, válvula Va dañada cuando Vi y Vs están atascadas. Al momento de la simulación de estos fallos también se toman muestras de las variables de diagnóstico del sistema, las cuales se almacenan en una matriz de 6 columnas por 20 filas de datos de cada fallo (esto significa 140 filas para un total de 7 fallos), donde cada columna corresponde a una variable de muestreo o de diagnóstico.

Aprendizaje supervisado de estructura

En esta matriz de datos de muestreo tenemos la materia prima para dar inicio al aprendizaje supervisado de estructura de red, se dice que es supervisado porque el experto debe proveer las posibles estructuras de red, de las cuales una resultará ser la óptima. Como se muestra en la figura 7.10, para este caso de estudio proponemos dos posibles estructuras de red, la que obtenga la mayor probabilidad según la expresión (7.4) esa

será la elegida. Se aclara que aunque las redes sólo incluyen 3 fallos únicos (Va , Vi , Vs) es posible hacer inferencias para fallos múltiples (por ejemplo, $p(Va, Vi|q_2, h_1, h_2)$). Las expresiones del (7.2) al (7.4) son usadas también en (Haider 2003).

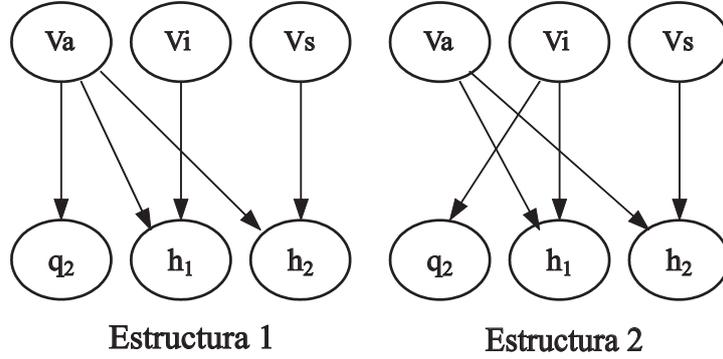


Figura 7.10: Estructuras de red propuestas en el aprendizaje supervisado

$$p[\theta_1] = \frac{\alpha_1 + N_1}{\alpha_1 + \alpha_2 + N}, \quad p[\theta_2] = \frac{\alpha_2 + N_2}{\alpha_1 + \alpha_2 + N} \quad (7.2)$$

$$\int \theta^{n_1} (1 - \theta)^{n_2} d\theta = \frac{n_1! n_2!}{(n_1 + n_2 + 1)!} = \tau_i \quad (7.3)$$

$$p(D|S, \theta_S) = \prod_i \tau_i \quad (7.4)$$

En el aprendizaje de estructura la expresión (7.3) es para calcular la probabilidad *a priori* de cada nodo ya sea de evidencia o de consulta. Las probabilidades *a priori* que nos interesan en las redes propuestas son: $p(Va)$, $p(Vi)$ y $p(Vs)$. En la expresión (7.3) n_1 significa el número de unos en los datos de muestreo del nodo o variable de la cuál se obtiene la probabilidad, n_2 significa el número de ceros en los datos de muestreo de esta misma variable.

Para calcular la probabilidad *a posteriori* (probabilidad condicional) que se usará en el aprendizaje de estructura, utilizaremos también la expresión (7.3), salvo que se tomaran en cuenta las probabilidades condicionales para definir n_1 y n_2 . Por ejemplo en $p(q_2|Va)$, n_1 es el número de unos en q_2 en cuyas filas en la matriz de datos de muestreo, también coincide con unos en Va y n_2 es el número de ceros en q_2 en cuyas filas en la matriz de datos de muestreo, coincide con unos en Va .

Las probabilidades *a posteriori* a calcular para la estructura 1 serán: $p(q_2|Va)$, $p(h_1|Va, Vi)$, $p(h_2|Va, Vs)$, para la estructura 2 serán: $p(q_2|Vi)$, $p(h_1|Va, Vi)$, $p(h_2|Va, Vs)$. Una vez obtenidas las probabilidades *a priori* y *a posteriori* para el aprendizaje de estructura, es posible calcular cuál estructura es la de mayor probabilidad, por el uso de la expresión (7.4) que establece la multiplicación de probabilidades *a priori* y *a posteriori* de cada estructura.

Aprendizaje paramétrico

Una vez concluido el aprendizaje de estructura pasamos al aprendizaje de parámetros, dado que consideramos el caso binomial asumiremos una distribución Beta Dirichlet. Para esto usaremos la expresión (7.2) donde $\alpha_1 = 1$ y $\alpha_2 = 1$ son los parámetros que se quieren aprender. Como en la expresión (7.3), también la expresión (7.2) nos puede servir para calcular tanto la probabilidad *a priori* como la *a posteriori* en el aprendizaje paramétrico.

No obstante que el comportamiento de las variables en este programa, incluyen cierto grado de ruido aleatorio, la estructura que generalmente es elegida por el aprendizaje de estructura es la estructura 1. Las probabilidades *a priori* que necesitamos aprender para los parámetros de la red elegida son: $p(Va)$, $p(Vi)$ y $p(Vs)$. En la expresión (7.2) N_1 es el número de unos, N_2 es el número de ceros y N es el total de muestras de la variable o nodo del que se quiere obtener el parámetro. En la expresión (7.2) $p(\theta_1)$ equivale a la probabilidad positiva de la variable (ej., $p(Va)$) y $p(\theta_2)$ equivale a la probabilidad negativa de la variable (ej., $p(\neg Va)$).

Para calcular las probabilidades *a posteriori* en el aprendizaje paramétrico, también se usa la expresión (7.2), pero considerando la forma condicional de las variables. Por ejemplo en $p(q_2|Va)$, N_1 es la cantidad de unos en q_2 cuya correspondiente muestra en Va también es uno, N es el número total de unos en Va . Para $p(q_2|Va)$, N_2 es la cantidad de unos en q_2 cuya correspondiente muestra en Va es cero, N es el número total de ceros en Va . Las probabilidades *a posteriori* a calcular para completar los parámetros de la estructura 1 serán: $p(q_2|Va)$, $p(h_1|Va, Vi)$, $p(h_2|Va, Vs)$.

7.4.3. Diagnóstico de fallos usando una red Bayesiana

Una vez obtenidos los parámetros de la red, simulamos el diagnóstico de fallos a través de la red Bayesiana. En este punto, el programa nos pregunta que fallo queremos simular de un listado de fallos que nos presenta, al elegir uno damos inicio a la simulación del fallo.

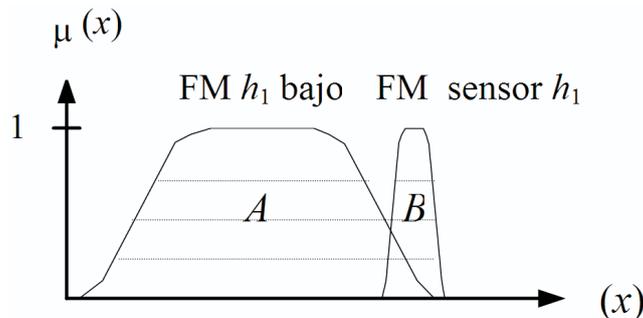


Figura 7.11: Funciones de pertenencia que caracterizan al nivel de h_1 bajo y al sensor h_1

La señal muestreada de los sensores q_2 , h_1 , h_2 , es caracterizada a través de la función de pertenencia campana generalizada (Jang 1997), que es muy parecida a la función de pertenencia trapezoidal borrosa (ver figura 7.11), muy usada en la lógica borrosa. Luego esta interpretación borrosa de las variables de diagnóstico, es convertida a un intervalo de probabilidad a través de la teoría de Dempster-Shafer (Shafer 1976; Yen 1990), entonces en vez de tener representadas 20 muestras de una variable de diagnóstico (B) en función de pertenencia campana, obtendremos de esta función la probabilidad inferior bel (credibilidad) y su probabilidad superior pl (plausibilidad).

Las expresiones en (5.19) son usadas para calcular credibilidad (bel) y plausibilidad (pl) respectivamente donde $m(A)$, es la asignación básica de probabilidad de B con respecto al focal no borroso A (representación borrosa de la media de B con respecto a A $\mu_{\bar{x}(B)}(A)$), los focales son subdivisiones horizontales no borrosas de las funciones de pertenencia A y B (ver figura 7.11), los niveles de corte α_i forman la resolución de los focales no borrosos que componen A y B , μ_B es la medida borrosa inferior o superior representativa del focal no borroso B con respecto a los focales de A . De esta manera la credibilidad y plausibilidad de que el nivel h_1 (B) sea bajo dependerá de la relación entre las expresiones en (5.19) de la FP del sensor h_1 (B) con la FP del nivel h_1 bajo (A). Una explicación más detallada en la conversión de conjuntos borrosos a intervalo de probabilidad se encuentra en (Yen 1990).

Propagar un intervalo de probabilidad [bel, pl] proporciona más completitud en la información, que sólo propagar la media muestral \bar{x} a través de la red Bayesiana de diagnóstico. La inferencia probabilista de diagnóstico para este caso de estudio, propaga sus probabilidades a través de la red usando el algoritmo para poliárboles (Nilsson 2001; Russell 2003). Un ejemplo en la inferencia de un fallo usando este algoritmo en la estructura 1 de la figura 7.10 sería la expresión (7.5), que calcula la probabilidad de que Vi esta obstruida dado que los sensores q_2 , h_1 , y h_2 , están altos.

$$p(Vi|Sq_2, Sh_1, Sh_2) = \frac{p(Sq_2, Sh_1, Sh_2|Vi)p(Vi)}{p(Sq_2, Sh_1, Sh_2|Vi)p(Vi) + p(Sq_2, Sh_1, Sh_2|\neg Vi)p(\neg Vi)} \quad (7.5)$$

Recordemos que las probabilidades *a priori* como $p(Vi)$ y otras probabilidades *a posteriori* que se usan para llegar a la expresión (7.5), fueron obtenidas en el aprendizaje paramétrico. En la figura 7.12 se muestra las gráficas de plausibilidad y credibilidad que resultan de la simulación del fallo de Vi .

Para obtener las inferencias probabilísticas para las otras dos válvulas se calculan expresiones similares a (7.5). Donde las variables Sq_2 , Sh_1 , y Sh_2 representan el intervalo probabilístico para la actualización de valores de los sensores que aportan las nuevas evidencias a la red de las variables q_2 , h_1 , y h_2 respectivamente.

En esta gráfica k es el meta periodo de muestreo que es de 200 seg. El periodo de muestreo base es de 10 seg. tiempo suficientemente rápido para el monitoizado de variables para un sistema de depósitos como el de la figura 7.5 donde los valores nominales de $q_2 = 0,02 \text{ m}^3/\text{s}$, $h_1 = 2 \text{ m}$, $h_2 = 1 \text{ m}$.

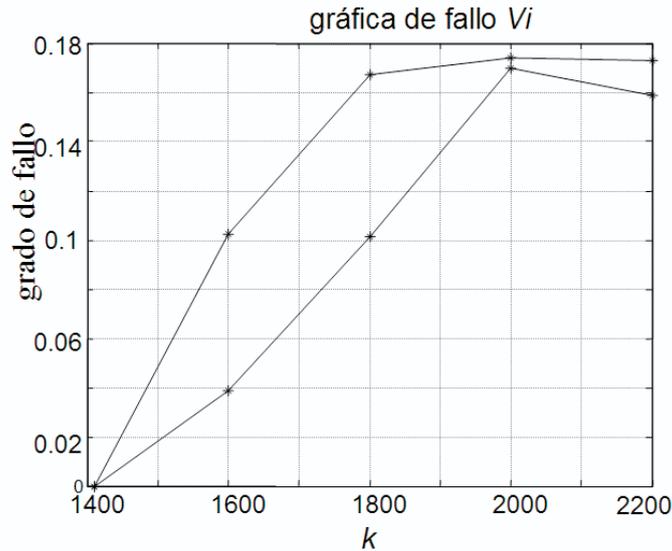


Figura 7.12: Gráfica de plausibilidad y credibilidad del fallo de V_i

Con el tiempo base de muestreo de 10 seg. se logran tomar 20 muestras cada meta periodo de muestreo k , con esas muestras, se construye la función de pertenencia campana generalizada para los sensores de q_2 , h_1 , y h_2 . Las funciones de pertenencia de los sensores se usarán para obtener la credibilidad y la plausibilidad a través de (5.19). Cabe aclarar que las funciones de pertenencia que definen a los niveles de las variables (h_1 , h_2 , q_2) como bajo, normal y alto son definidos desde un inicio en el programa y no cambian a menos que sucedan excepciones.

El intervalo de probabilidad $[bel, pl]$ propagado por la inferencia de diagnóstico (7.5) en la gráfica del fallo de V_i (figura 7.12), muestra dos gráficas la plausibilidad que es la gráfica superior y la credibilidad que es la gráfica inferior. Para simular este fallo se uso una perturbación aleatoria máxima lo que es evidente en la gráfica pues podemos observar como el intervalo crece y decrece que es precisamente lo que deseamos saber para dar una adecuada interpretación a la incertidumbre presente en las variables de diagnóstico.

7.4.4. Análisis de resultados en el diagnóstico por redes Bayesianas

Dado que este caso de estudio es una simulación, la misma proporción de muestras en la simulación de cada fallo fueron colectadas para el aprendizaje tanto de parámetros como de estructura de red, lo anterior produce que en el diagnóstico de fallos, la diferencia del intervalo de probabilidad entre un fallo y otro sea muy pequeña. En un sistema real, los resultados dependerán de los datos estadísticos que se tengan de las variables de diagnóstico. Como es probable que algunos fallos sean más frecuentes que otros, es posible que la diferencia entre fallos se incremente.

El hecho de que el intervalo de probabilidad sea grande en los fallos, indica que la incertidumbre en las señales provistas por los sensores es grande, lo cual puede deberse a la dinámica inherente al sistema, a la presencia de ruido o fallos en los sensores.

En este enfoque, tratamos la incertidumbre inherente a las variables de diagnóstico, desde el punto de vista aleatorio o probabilista, y también desde el punto de vista borroso, al convertir las medidas borrosas provistas por los sensores, a un intervalo de probabilidad a través de la teoría de Dempster-Shafer, que nos proporcionó mayor completitud en la información necesaria para la toma de decisiones cuando un fallo se presenta.

Conclusiones del capítulo

Hemos estudiado todas las aportaciones realizadas en el periodo de investigación inicial, donde pudimos evaluar diferentes técnicas de Inteligencia Artificial. El énfasis en el trato de la incertidumbre inherente a los sistemas de diagnóstico a sido puesto en la representación de la información primaria que alimenta a nuestro sistema en forma intervalar.

Como se puede ver se han logrado algunos avances en el trato de la incertidumbre. Sin embargo, aún nuestra salida de diagnóstico puede resultar compleja para el experto en diagnóstico, pues resulta difícil mantener en mente intervalos de fallos y cuando la superposición de esos intervalos ocurre, en el caso de fallos múltiples, se torna aún más compleja la toma de decisiones. Aunado a esto, el entorno complicado de diferentes conceptos como: intervalo borroso, intervalo probabilístico o intervalo posibilístico, intervalo de creencia, inferencia borrosa, inferencia probabilística o Bayesiana, inferencia posibilística.

Toda esta gama de terminología que se ha involucrado, hace confusa la diagnosis final para el técnico de diagnóstico, que generalmente no esta familiarizado con estas teorías, si no que se especializan en campo de aplicación concreto, como puede ser: el diagnóstico de procesos industriales (calderas, hornos, procesos biológicos etc.) o el diagnóstico de equipos, como son: las máquinas de combustión interna. En el próximo capítulo se propone una metodología de diagnóstico más directa

Capítulo 8

Diagnóstico posibilístico borroso como un problema de optimización lineal

8.1. Introducción

En este capítulo tratamos sobre la aportación principal de esta tesis. Se propone una técnica de diagnóstico borroso posibilístico muy sencilla de aplicar. Que ofrece como salida de diagnóstico un óptimo único para cada fallo o desorden. Una metodología de diagnóstico, la cuál el técnico en diagnóstico puede aprender sin necesidad de profundizar mucho en la fundamentación teórica de la misma. Además es fácil de mecanizar su aplicación, a cualquier sistema de diagnóstico que obedezca el principio causa-efecto. Esta metodología se explicará en los subsecuentes capítulos. Primero, estableciendo su fundamentación teórica, que se basa en las técnicas relacionales de conjuntos borrosos y en la optimización lineal posibilística. Después, ejemplificando su uso en un sistema de diagnóstico para motores de combustión interna, por el análisis de aceites usados.

El diagnóstico es el proceso de inferir desórdenes a partir de medidas las cuales denotan una desviación de un comportamiento “normal” de un sistema en particular. Los resultados del proceso de diagnóstico son la estimación del grado de presencia de cada desorden, el cual puede ser expresado por un valor “si/no” o por resultados graduados continuamente (denotando “severidades” de desórdenes desde cero a un 100 %). Diferentes enfoques aparecen en la literatura abordando este problema (Chiang 2001; Blanke 2003): base de datos, sistemas basados en el conocimiento, o ecuaciones diferenciales. El problema de diagnosis completo, debe incluir análisis temporal (Wainer 1997) y configuración probabilística (Bolgelt 2000), en caso de contar con un marco de trabajo como las redes Bayesianas recurrentes híbridas (Gharamani 2001; Russell 2003), o con criterios teóricos sobre toma de decisiones. Un problema como este es bastante complejo, normalmente se declaran asunciones simplificadas y aproximadas, como en el caso que aquí se trata.

En muchos casos prácticos, el conocimiento relacionado al diagnóstico es proporcionado por los expertos, quienes lo expresan en declaraciones como “*una ingestión de alcohol alta causa una presión sanguínea más alta de lo normal*”: El bloque de construcción básico en muchos enfoques basados en lógica para el diagnóstico de fallos son enunciados lógicos que relacionan *desórdenes* y *manifestaciones medidas*. La graduación en intensidad de manifestaciones y desórdenes ha inspirado el uso de la lógica borrosa (Angeli 1999; Carrasco 2004). Otros marcos de trabajo basados en el conocimiento, han extendido los esquemas de razonamiento lógico básico para tratar con la incertidumbre en las bases de conocimiento vía el razonamiento aproximado (Kruse 1991; Shafer 1990), el razonamiento posibilístico y abductivo (Mouzon 2000; Mouzon 2001; Dubois 2001a; Yamada 2002), redes Bayesianas (Castillo 1997; Russell 2003), etc.

Esta aportación ha sido inspirada en los trabajos (Dubois 1995; Cayrac 1995; Mouzon 2000). Sin embargo, en vez de seguir al razonamiento aproximado y a la argumentación lógica, el objetivo de esta aportación es formular una interpretación del diagnóstico borroso posibilístico en términos de una optimización con restricciones lineales (Dubois 1996), donde las restricciones y el índice de coste representa el conocimiento de diagnóstico codificado. La posibilidad de un desorden es definida como la exponencial del índice de coste mínimo factible, cuando el desorden se encuentra presente. La idea de transformar la inferencia lógica en restricciones algebraicas y optimización fue usada en (Sala 1998; Sala 2001b; Sala 2006a): Las implicaciones fueron transformadas en desigualdades. La aportación clave de este trabajo es demostrar que una relación compleja entre los desórdenes y los valores observados de las manifestaciones, pueden declararse en una manera simple, si se introducen entidades intermedias adicionales (como las variables de decisión): que pueden ser llamadas *variables de asociación*, *mediciones*, pero que preferentemente serán llamadas *variables artificiales*.

En la primera parte de este capítulo se presenta una introducción a la optimización lineal. Después describimos una metodología para codificar el conocimiento básico. Se asume que el conocimiento consiste en una descripción de “asociaciones” entre un conjunto de *manifestaciones* m_j medibles, y un *desorden* único particular, d_i . Las manifestaciones serán divididas en tres grupos: Aquellas que ciertamente son causadas por d_i , aquellas que se encuentran ausentes cuando únicamente d_i se encuentra presente, y aquellas cuya presencia o ausencia es incierta. En el caso binario, esta aportación coincide con (Dubois 1995). Sin embargo, en vez de codificar el conocimiento como una lista, como se hace en (Dubois 1995), el conocimiento será equivalentemente codificado como restricciones de igualdades y desigualdades. Las observaciones inciertas también serán definidas con el mismo tipo de restricciones. Entonces, el diagnóstico consistente (Dubois 1995; Peng 1990) será entendido como la solución factible, del conjunto de restricciones mencionadas anteriormente que representan el conocimiento y los hechos. En la segunda parte de este capítulo, se discute la situación donde un conjunto de observaciones puede tener muchas explicaciones alternativas. En este enfoque, las preferencias abductivas con respecto al conocimiento incierto (Dubois 1995; Mouzon 2001; Cayrac 1995; Peng 1990; Bolgelt 2000) se clasificarán implícitamente por la descripción de una distribución posibilística a través de un índice de costes, similar al enfoque presentado en (Dubois 1996). Entonces, maximizando la posibilidad, la abducción será definida

como la optimización funcional lineal o cuadrática sujeta a restricciones lineales, el cuál es un problema que se ha estudiado profundamente en el pasado (Sierksma 2001; Schrage 1998).

Existen expresiones equivalentes de las restricciones en términos de los enunciados lógicos. Sin embargo, el énfasis de esta aportación, como en (Sala 1998), es puesto en dirección opuesta: mostrando que el álgebra computacional y el software de optimización que han sido ampliamente divulgados, pueden ser usados para inferencia lógica (en el esquema particular propuesto).

Las principales ventajas de la técnica propuesta son: a) la habilidad para hacer uso de técnicas de optimización muy conocidas disponibles en software comercial muy rápidos capaces de tratar con miles de variables de decisión, b) la incorporación de *ambos* “borrrosidad” (entendida como la gradación en síntomas y desórdenes) y la “incertidumbre gradual” (entendida como restricciones de desigualdades relajadas); c) la habilidad para tratar naturalmente con múltiples fallos (incluyendo mediciones espurias).

8.2. Optimización lineal

Una de las características de los modelos de optimización es la existencia de un único decisor frente a otras disciplinas donde puede existir más de un decisor. Debido a la naturaleza del problema a resolver en esta aportación, el modelo de optimización a tratar será un modelo determinista estático con un único objetivo que es la minimización con restricciones lineales.

El planteamiento general de un problema de optimización es de la forma siguiente:

$$\begin{aligned} \text{Optimizar} \quad & f(x_1, x_2, \dots, x_n) \\ \text{sujeto a:} \quad & g_1(x_1, x_2, \dots, x_n) \leq b_1 \\ & g_2(x_1, x_2, \dots, x_n) \leq b_2 \\ & \dots \\ & g_m(x_1, x_2, \dots, x_n) \leq b_n \end{aligned}$$

La función f denominada *función objetivo*, es una función definida de un dominio de \mathfrak{R} , y representa una descripción matemática y cuantificada del objetivo que se pretende alcanzar con el problema planteado.

El vector x es el vector de *variables de decisión*, de entre cuyos valores posibles se trata de elegir aquél o aquellos que proporcionen el valor óptimo de la función f .

Conjunto de oportunidades, denominaremos así al conjunto de puntos $x \in \mathfrak{R}$ que verifica todas y cada una de las restricciones y al mismo tiempo pertenecen al dominio de definición de la función. El conjunto de oportunidades es representado por S .

La *forma estándar* en la que se formula un problema de programación lineal, nos será de mucha utilidad para nuestra propuesta. Es aquella donde se plantean todas las restricciones en forma de igualdad. Para convertir las desigualdades en igualdades podemos recurrir a las denominadas variables de holgura s_i (variables artificiales). Estas variables han de ser no negativas, por tanto para las restricciones de la forma \leq se introducen sumando al primer miembro mientras que para las desigualdades de la forma \geq se introducirán restando al primer miembro. Es decir:

Consideraremos la i -ésima restricción:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leq b_i$$

definimos la variable de holgura s_i como:

$$s_i = b_i - (a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) \geq 0$$

de modo que la restricción se puede escribir como:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + s_i = b_i$$

Si la restricción es mayor-igual (para minimización):

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \geq b_i$$

definimos la variable de holgura s_i como:

$$s_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - b_i \geq 0$$

por lo que la restricción se puede escribir como:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - s_i = b_i$$

Las variables de holgura o variables artificiales además de servirnos para convertir desigualdades en igualdades, tiene una utilidad adicional, pues, nos informa del comportamiento de las restricciones. Si una variable artificial es nula, ello significa que la restricción correspondiente se comporta como una igualdad, es decir, se trata de una restricción activa. En el caso de que la variable artificial sea positiva, la restricción correspondiente es una restricción inactiva. Este concepto de holgura en las variables artificiales, nos servirá en el análisis de señales de entrada para nuestro sistema de diagnóstico, que nos dirán cuando un dato de entrada puede ser considerado como un dato erróneo, o un caso no previsto en nuestra base de hechos.

En los problemas de programación lineal la función objetivo es lineal, cumple las propiedades de convexidad y concavidad de la función. El conjunto de oportunidades es siempre convexo, por este motivo los óptimos de los problemas lineales son siempre óptimos globales. Para la solución de problemas de programación lineal simples existe

el método algebraico, sin embargo, para resolver problemas de mediana y gran complejidad existen algoritmos como el simplex introducido por Dantzing en 1947. Este último aunque es muy usado por la sencillez de su planteamiento puede presentar errores de ciclado en la solución, por lo que se han implementado nuevos algoritmos más eficaces (Karmarkar, 1984) (Molochí 1999) que son usados en los software comerciales. En esta aportación los software que estarán siendo utilizados para la resolución de problemas de minimización serán MATLAB y MATHEMATICA.

8.3. Definiciones básicas

En esta sección, se presentarán las definiciones básicas y la notación que se usará en este capítulo. Los elementos claves en este sistema de diagnóstico serán: desórdenes, manifestaciones (causadas por los desórdenes) y observaciones (de algunas de las manifestaciones). La relación entre desórdenes y manifestaciones constituirán la base de conocimiento; las relaciones entre manifestaciones y observaciones serán incorporadas en la que llamaremos base de hechos.

Desórdenes. Consideremos un conjunto q de desórdenes únicos básicos $\{d_1, d_2, \dots, d_q\}$ que pudiesen presentarse en un sistema de diagnóstico (con la posibilidad de la presencia simultánea de algunos otros desórdenes). Cada desorden podría ser concebido como una estructura de dos campos: una etiqueta lingüística (tal como $d_1.etiqueta = \text{“fiebre”}$) y un coeficiente de severidad, tal como $d_i.f$, que denota la presencia ($d_i.f=1$) de algún desorden o la ausencia ($d_i.f=0$), o un grado parcial de presencia (Kitowski 1987) con una severidad borrosa ($0 \leq d_i.f \leq 1$). Abusando de la notación, dependiendo del contexto d_i será usado para denotar ambos, la severidad del desorden y la etiqueta (desde este punto en adelante los campos f y $etiqueta$ no serán utilizados).

El estado de los desórdenes en un sistema será considerado como un vector, D , cuyos componentes (coordenadas de desórdenes) son la severidad $D = (d_1, \dots, d_q)$. El vector $D = (0, 0, \dots, 0)$ será denotado, por definición, como situación *normal* (ausencia de desórdenes). Los vectores con más de una severidad no-cero serán denotados como *desórdenes múltiples*.

Manifestaciones. Las posibles manifestaciones (síntomas) de los desórdenes descritos anteriormente serán $\{m_1, \dots, m_n\}$ donde cada m_i , tiene una etiqueta lingüística asociada (tal como $m_1.etiqueta = \text{“temperatura anormal”}$, etc.) y una “intensidad de presencia” asociada, es decir, $m_i.s$, denotando su ausencia ($m_i.s = 0$), total presencia ($m_i.s = 1$), o un grado parcial de presencia $0 \leq m_i.s \leq 1$. Como en el caso de los desórdenes, dependiendo del contexto, m_i será usado para denotar, el nombre de la manifestación o su intensidad. Si es conveniente, las intensidades de las manifestaciones estarán colocadas en un vector, denotado por M .

Observaciones. Para cada caso de diagnóstico, algunas de las manifestaciones podrán ser observadas (medidas). Las mediciones m_j serán definidas como la obtención de una

aserción tal como “las manifestaciones m_j toman los valores de σ_j ”. Idealmente, la aserción sería equivalente a $m_j = \sigma_j$; Sin embargo la distinción entre m_j y σ_j será usada después para poder integrar fácilmente una clase especial de “desórdenes de sensores” (por ejemplo, medidas incorrectas y espurias). El conjunto de valores medidos será denotado como Σ .

Ejemplo 1. Considere una aplicación médica, tal como el ejemplo académico en (Yamada 2002). Algunos desórdenes pueden ser definidos, tales como “ d_1 : fiebre de heno”, “ d_2 : gripe”, “ d_3 : alimentos tóxicos”, etc. La ocurrencia de los desórdenes causa la aparición de algunas manifestaciones, como son “ m_1 : congestión nasal”, “ m_2 : fiebre”, “ m_3 : diarrea”, etc.

Cada caso de diagnóstico provee un conjunto de valores, estos son σ_j , los cuales están asociados a alguna de las manifestaciones m_j : un valor medido normal de temperatura corporal puede ser codificado como $\sigma_2 = 0$, la presencia de congestión nasal se codifica como $\sigma_1 = 1$, etc.

Si un conjunto borroso es definido, diciendo, las lecturas de temperatura del cuerpo (definiendo el concepto lingüístico de “temperatura anormal.” a través de una función de pertenencia significativa), entonces el σ_2 medido estaría en el rango entre cero y uno: por ejemplo, una “temperatura ligeramente anormal”, como 37.2°C , podría ser interpretada como $\sigma_2 = 0,25$.

El objetivo de diagnóstico es estimar el vector de severidad de desórdenes actual D a partir de las observaciones y el conocimiento a priori (posiblemente incierto) que se describirá en la siguiente sección. En este trabajo, la meta es expresar el conocimiento como un conjunto de restricciones en D y M , envolviendo también algunas variables artificiales. Intencionalmente, estas restricciones serán lineales, con el fin de poder usar algoritmos de programación lineal en la solución del problema de minimización.

8.4. Metodología de diagnóstico (caso binario)

Vamos a considerar primero el caso binario, donde los desórdenes, manifestaciones y observaciones o están ausentes o presente, representados por valores numéricos entre $[0, 1]$.

El bloque de construcción básico del conocimiento de diagnóstico es formado por asociaciones entre los desórdenes y síntomas. De acuerdo con (Dubois 1995), dado un desorden d_i , el conjunto de manifestaciones presentes que son causadas únicamente por d_i serán denotas por $M(d_i)^+$, y aquellas definidas como ciertamente ausentes cuando únicamente d_i se encuentra presente serán denotadas por $M(d_i)^-$. La *situación normal* se asociará a todas las manifestaciones m_j las cuales serán igual a cero (cambiando las definiciones de algunas manifestaciones por su negación lógica, si es necesario).

Asociación desorden-manifestación. La asociación entre desórdenes d_i y manifestaciones m_j se expresará por una “variable de asociación” auxiliar, denotada como s_i^j que cumple las siguientes restricciones:

- Si la manifestación m_j ciertamente es causada por d_i cuando únicamente d_i se encuentra presente ($m_j \in M(d_i)^+$), s_i^j cumplirá, por definición, la restricción de igualdad:

$$s_i^j = d_i \quad (8.1)$$

- Si la manifestación m_j ciertamente se encuentra ausente cuando únicamente d_i se encuentra presente ($m_j \in M(d_i)^-$), entonces

$$s_i^j = 0 \quad (8.2)$$

- Si la presencia o ausencia de la manifestación m_j cuando d_i se encuentra presente es incierta, (por ejemplo, $m_j \in U(d_i)$, $U(d_i) = \neg(M(d_i)^+ \cup M(d_i)^-)$), las siguientes restricciones en s_i^j se añadirán:

$$s_i^j \leq d_i \quad (8.3)$$

Indicando que en situación “normal” m_j se encuentra ausente, pero que bajo d_i se encuentra presente algunas veces.

Superposición. Una vez que se han fijado las restricciones en los coeficientes de relación s_i^j , el estado inferido de la presencia o ausencia de las manifestaciones m_j será dado por la disyunción sobre todos los posibles desórdenes presentes, que se podría expresar de la siguiente forma:

$$m_j = \min(1, \sum_{i=1}^k s_i^j) \quad (8.4)$$

Sin embargo, (8.4) no es lineal. Con el fin de preservar la linealidad, se agregarán las siguientes desigualdades a la base de conocimiento (BC) en lugar de (8.4):

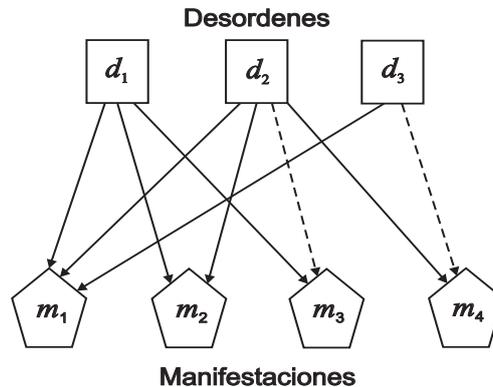


Figura 8.1: Una base de conocimiento básica. (—→) la flecha sólida: asociaciones ciertas; (---→) flecha discontinua: asociaciones inciertas.

$$s_i^j \leq m_j \quad \forall i \quad (8.5)$$

$$m_j \leq \sum_i s_i^j \quad (8.6)$$

La ecuación (8.5) indica que la intensidad de la manifestación es mayor o igual a la severidad del desorden que causa esta; la segunda indica que al menos uno de los desórdenes debe estar presente para que m_j no sea cero, En este contexto, la interpretación de $s_i^j = 1$ es “ d_i se encuentra presente y esta causando m_j ”.

Al conjunto de restricciones presentado, que envuelve desórdenes, manifestaciones y variables de asociación, será denotado como *base de conocimiento* (BC).

Ejemplo 2 Vamos a considerar el conocimiento a cerca de tres desórdenes y cuatro manifestaciones dado por:

$$\begin{aligned} M(d_1)^+ &= \{m_1, m_2, m_3\}, M(d_1)^- = \{m_4\}, U(d_1) = \emptyset \\ M(d_2)^+ &= \{m_1, m_2, m_4\}, M(d_2)^- = \emptyset, U(d_2) = \{m_3\} \\ M(d_3)^+ &= \{m_1\}, M(d_3)^- = \{m_2, m_3\}, U(d_3) = \{m_4\} \end{aligned}$$

donde podemos notar: 1) “el desorden único d_1 esta presente, causando m_1, m_2, m_3 y no causa m_4 ”; 2) “la presencia de d_2 causa m_1, m_2 y m_4 ; con respecto a m_3 , se encuentra algunas veces presente y otras ausente (caso incierto)”; 3) “la presencia de d_3 causa m_1 pero, ciertamente no causa ni m_2 ni m_3 ; la presencia de m_4 con d_3 es incierta”. Una representación gráfica aparece en la figura 8.1.

El conocimiento causal puede ser equivalentemente expresado por los coeficientes de relación:

$$s_1^1 = s_1^2 = s_1^3 = d_1, \quad s_1^4 = 0 \quad (8.7)$$

$$s_2^1 = s_2^2 = s_2^4 = d_2, \quad s_2^3 \leq d_2 \quad (8.8)$$

$$s_3^1 = d_3, \quad s_3^2 = s_3^3 = 0, \quad s_3^4 \leq d_3 \quad (8.9)$$

conjuntamente con las ecuaciones de superposición manifestación:

$$s_1^1 \leq m_1, \quad s_2^1 \leq m_1, \quad s_3^1 \leq m_1, \quad s_1^1 + s_2^1 + s_3^1 \geq m_1 \quad (8.10)$$

más las análogas para m_2, m_3 y m_4 .

Evidentemente, variables iguales a cero y aquellas iguales a los desórdenes podrían ser algebraicamente eliminadas. En la actual implementación: sólo el s_i^j relacionado con los síntomas inciertos es necesario conservar. En este ejemplo, la base de conocimiento equivalente simplificado estará dada por:

$$\begin{aligned}
& s_2^3 \leq d_2, \quad s_3^4 \leq d_3 \\
d_1 \leq m_1, \quad d_2 \leq m_1, \quad d_3 \leq m_1, \quad d_1 + d_2 + d_3 \geq m_1 \\
& d_1 \leq m_2, \quad d_2 \leq m_2, \quad d_1 + d_2 \geq m_2 \\
& d_1 \leq m_3, \quad s_2^3 \leq m_3, \quad d_1 + s_2^3 \geq m_3 \\
& d_2 \leq m_4, \quad s_3^4 \leq m_4, \quad d_2 + s_3^4 \geq m_4
\end{aligned}$$

Comentario: si a y b son valores lógicos, la desigualdad $a \leq b$ reproduce la tabla de verdad de la implicación binaria clásica $a \rightarrow b$. Entonces, la estructura de la base de conocimiento propuesta puede ser equivalentemente entendida como un conjunto de aserciones lógicas (\vee denota disyunción):

$$normal \rightarrow \neg m_j, \quad \forall j \quad (8.11)$$

$$s_i^j \rightarrow d_i, \quad s_i^j \rightarrow m_j \quad \forall i, j \quad (8.12)$$

$$m_j \rightarrow \bigvee_{i=1}^q s_i^j \quad \forall j \quad (8.13)$$

$$d_i \rightarrow s_i^j \quad \text{if } m_j \in M(d_i)^+ \quad (8.14)$$

$$d_i \rightarrow \neg s_i^j \quad \text{if } m_j \in M(d_i)^- \quad (8.15)$$

Note que para $m_j \in M(d_i)^-$, (8.15) vincula a $s_i^j \rightarrow \neg d_i$. Juntamente con (8.12), s_i^j debe ser cero, porque de otra manera d_i y $\neg d_i$ se estarían vinculando.

8.4.1. Incorporando observaciones

En cada caso particular de diagnóstico, la severidad de alguna de las manifestaciones (m_j) se asumirán como disponibles, denotadas como medidas σ_j . Estas nos llevaran a un conjunto adicional de restricciones que será llamado *Base de Hechos* (BH).

Con un sensor “fiable” las σ_j observadas deberían ser iguales a la intensidad de la manifestación actual m_j , como se comento previamente, por ejemplo, $m_j = \sigma_j$ es agregada a la (BH).

Con el fin de incorporar la posibilidad de mediciones inciertas, incorporamos restricciones a la BH tales como

$$\sum_{j \in H} m_j \geq 1 \quad (8.16)$$

esta indicaría que al menos una de las manifestaciones en un particular conjunto (indexado por H) debe ser verdadera, pero se desconoce cuál es.

8.4.2. Refinamientos en la metodología

Con el fin de aumentar flexibilidad en el anterior esquema de diagnóstico es necesario implementar algunos refinamientos.

Desviaciones a la hipótesis de superposición. En algunos casos, los síntomas producidos por la ocurrencia combinada de algunos desórdenes, podrían ser diferentes a los síntomas producidos por el conjunto de los desórdenes individuales. Para esto se requiere la introducción de una variable de “desorden compuesto” artificial como en (Mouzon 2001).

Ejemplo 3 bajo la hipótesis de superposición que motiva las ecuaciones (8.5) y (8.6), la ocurrencia combinada del desorden d_1 y d_2 (denotado por d_{12}) vinculará al conjunto de asociaciones desorden-manifestación generado por:

$$M(d_{12})^+ = M(d_1)^+ \cup M(d_2)^+, \quad M(d_{12})^- = M(d_1)^- \cap M(d_2)^-$$

Si este no fuere el caso, debe insertarse una variable de “desorden compuesto” artificial d_{12} (y sus variables de asociación s_{12}^j) dentro de la base de conocimiento, con las restricciones apropiadas que describan la particular ocurrencia de síntomas (por ejemplo, se debe codificar $M(d_{12})^+$ y $M(d_{12})^-$), más la condición de mutuamente exclusivo $d_1 + d_2 + d_{12} \leq 1$.

d_i **causa al menos una de las manifestaciones** m_j . En algunos casos, la incertidumbre en los efectos de un desorden puede ser expresada en términos tales como: “ d_i causa al menos una de las manifestaciones m_j , $j \in \mathcal{K}_i$ ”, donde \mathcal{K}_i es un conjunto índice de posibles consecuencias de d_i (\mathcal{K}_i es usada en (Mouzon 2000)). Esto nos llevará a declarar, juntamente con (8.3), la restricción:

$$\sum_{i \in \mathcal{K}_i} s_i^j \geq d_i \tag{8.17}$$

La igualdad en (8.17) debe ser usada sólo si una manifestación en \mathcal{K}_i pudiese aparecer, esto es, si por ejemplo \mathcal{K} fuere una colección de manifestaciones mutuamente exclusivas.

Ejemplo 4 Considere una situación donde un desorden d_1 tiene las manifestaciones inciertas m_3 y m_4 . El conocimiento de que “al menos una de ellas aparece con d_1 ” podría ser expresado por

$$s_1^3 \leq d_1, \quad s_1^4 \leq d_1, \quad s_1^3 + s_1^4 \geq d_1$$

Correlación. Algunos tipos de “interdependencia” entre desórdenes y manifestaciones puede también ser declarado a través de desigualdades más generales que envuelvan a

d_i , s_i^j y m_j , refinando las capacidades de representación. La interpretación de algunas de estas capacidades se ilustrará en el siguiente ejemplo .

Ejemplo 5. Este ejemplo considera la traducción de una serie de enunciados lingüísticos a desigualdades.

- *El enunciado “la ocurrencia de d_2 es posible sólo si d_4 ó d_5 se encuentran presentes” (suponiendo que d_2 pudiese describir un desorden que sólo ocurre en pacientes con “diabetes (d_4).^o con coagulopatía (d_5)”). La siguiente expresión representa los anteriores desórdenes:*

$$d_2 \leq d_4 + d_5 \quad (8.18)$$

Note que, convenientemente, sólo las manifestaciones especificadas para d_2 es necesario definir vía (8.1)-(8.3) debido a que (18) obliga directamente la ocurrencia de aquellas asociadas a d_4 ó d_5 en caso de que d_2 estuviese presente. Si (18) no fuese declarada, entonces necesitaríamos definir: d_{24} = “ d_2 y d_4 ”, así como d_{25} = “ d_2 y d_5 ” tal y como se hizo en el Ejemplo 3, con la lista entera de las manifestaciones asociadas. Aunque este enfoque de desórdenes múltiples no es un problema desde el punto de vista teórico, es algo inconveniente desde el punto de vista del usuario final.

- *“ m_3 no es causado por d_2 cuando d_1 se encuentra ausente” lo anterior es representado por: $s_2^3 \leq d_1$ (agregando a $s_2^3 \leq d_2$ lo cual establece que d_2 algunas veces podría causar m_3).*
- *“Cuando d_1 se encuentra presente, d_2 debe también estar presente y causar la manifestación m_3 ” esto se representa¹ por $s_2^3 \leq d_1$.*
- *“ m_1 no es causado por d_4 cuando d_1 se encuentra presente” esto se representa por: $s_4^1 \leq d_4$ junto con $s_4^1 \leq (1 - d_1)$.*
- *“ d_3 y d_5 son mutuamente exclusivos” se representa por $d_3 + d_5 \leq 1$*
- *“ m_4 no es causado por d_2 si m_5 se encuentra ausente” se representa por $s_2^4 \leq m_5$.*
- *“ d_1 causa m_2 cuando d_4 se encuentra presente” se represente por $s_1^2 \geq d_1 + d_4 - 1$.*
- *“ d_1 causa m_2 cuando d_4 se encuentra ausente” se representa por $s_1^2 \geq d_1 - d_4$.*

Note que condiciones como las recién listadas, podrían ser una alternativa en muchos casos con respecto a la necesidad considerada anteriormente, de introducir variables artificiales para los desórdenes múltiples.

¹Evidentemente, $s_i^j \geq d_k$, junto con $s_i^j \leq d_i$, vinculan $d_k \leq d_j$ (en el espíritu de (8.18), como intuitivamente se espera)

8.4.3. Diagn osis basada en la consistencia

En este tema, S representar  el conjunto de variables de decisi n s_i^j y , como se menciono anteriormente, D representar  el vector severidad de desorden y M la intensidad de manifestaci n. Entonces, el problema de diagn osis podr a ser formulado como un problema de satisfacci n de restricciones:

Encontrar el conjunto de variables de decisi n factibles (D, S, M) dadas las restricciones en BC y BH .

Cada soluci n factible (D, S, M) podr a ser considerada como una *explicaci n* de todas las posibles observaciones (usualmente, s lo los H son de inter s). Si no existe una soluci n factible, entonces el diagn osis concluir  diciendo: *desorden desconocido*.

Todas las restricciones presentes son lineales en las variables desconocidas, as  que la *programaci n lineal* (PL) de n meros enteros (de hecho binaria) (Sierksma 2001) podr a ser usada como herramienta para obtener la soluci n, al menos en principio. Sin embargo, el software para esa clase de problemas es mucho menos eficiente que las implementaciones de PL de valores num ricos reales (la PL binaria tiene una complejidad NP-completa, mientras que la PL de valores reales es probablemente polinomial). Este hecho ha motivado el uso de valores l gicos de n meros reales, por ejemplo, aquellos usados en la l gica borrosa, que se ver  en la pr xima secci n.

Tambi n es necesario aclarar, que la teor a elemental de PL, nos muestra que el conjunto factible de variables de decisi n ser  un politopo convexo. Enumerar todos los v rtices en el politopo es muy poco pr ctico. Dado esto, el problema ser  convertido a una soluci n de optimizaci n en PL, de tal forma que obtengamos una soluci n  nica. Para ello, se introducir n las variables artificiales que se estudiar n en la secci n 8.7. Con la introducci n de estas variables que caracterizar n nuestras medidas en los sensores (m_i), nuestro sistema siempre obtendr  una soluci n  nica, aunque nuestra BH sea desconocida.

Ejemplo 6. *Considere una base de conocimiento con 4 des rdenes (d_1 a d_4) y 5 manifestaciones (m_1 a m_5). Por conveniencia, el desorden d_{14} se introducir  como una variable auxiliar usada para expresar la ocurrencia combinada de d_1 y d_4 .*

Las piezas disponibles de conocimiento se describir n ahora en t rminos ling u sticos y entonces, se traducir n a desigualdades con el fin de ilustrar la metodolog a:

- *Conocimiento en d_1 : este causa a m_1 , esto s lo ocurre si d_3 se encuentra presente.*

$$s_1^1 = d_1, \quad s_1^2 = s_1^3 = s_1^4 = s_1^5, \quad d_1 \leq d_3 \quad (8.19)$$

- *Conocimiento en d_2 : este causa m_2 y, ocasionalmente, tambi n m_4 .*

$$s_2^1 = 0, \quad s_2^2 = d_2, \quad s_2^3 = 0, \quad s_2^4 \leq d_2, \quad s_2^5 = 0 \quad (8.20)$$

- *Conocimiento en d_3 : este causa m_3 y, ocasionalmente, también m_5 (pero sólo cuando d_2 se encuentra presente y causa a m_4 ó cuando d_4 se encuentra presente).*

$$s_3^1 = s_3^2 = s_3^4 = 0, \quad s_3^3 = d_3, \quad s_3^5 \leq d_3, \quad s_3^5 \leq s_2^4 + d_4 + d_{14} \quad (8.21)$$

- *Conocimiento en d_4 : este causa m_4 y m_5 y, ocasionalmente, también m_2 .*

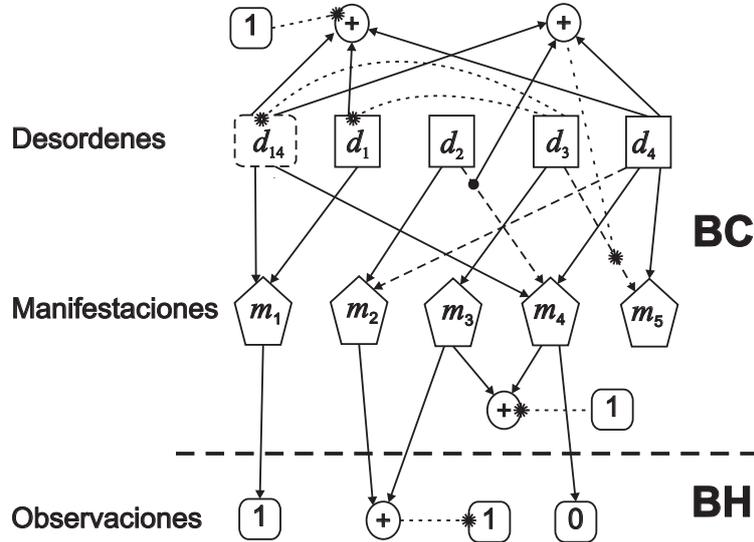


Figura 8.2: Descripción gráfica de un problema de diagnóstico. (—) la flecha sólida: asociaciones ciertas; (---) flecha discontinua: asociaciones inciertas; (···*) línea punteada con estrella: desigualdades (la estrella denota a en $a \leq b$).

$$s_4^1 = s_4^3 = 0, \quad s_4 \leq d_4, \quad s_4^4 = s_4^5 = d_4 \quad (8.22)$$

- *La ocurrencia combinada de d_1 y d_4 causa m_1 y m_4 (pero no m_2 , el cuál d_4 sólo a veces ha causado). Como d_1 causa d_3 , m_5 es aún concebible, pero no seguro (como exclusivamente sucedió con d_4).*

$$s_{14}^1 = d_{14}, \quad s_{14}^2 = s_{14}^3 = s_{14}^5 = 0, \quad s_{14}^4 = d_{14}, \quad d_{14} \leq d_3 \quad (8.23)$$

La exclusividad mutua entre d_1 , d_4 y d_{14} es también agregada a la base de conocimiento como una desigualdad:

$$d_1 + d_4 + d_{14} \leq 1 \quad (8.24)$$

- *Las manifestaciones m_3 y m_4 no pueden ocurrir simultáneamente.*

$$m_3 + m_4 \leq 1 \quad (8.25)$$

El diagrama en la figura 8.2 aproximadamente describe la relación entre las diferentes entidades en la base de conocimiento (así como un conjunto particular de observaciones como las que se muestran a continuación).

Se ha completado la base de conocimiento con las ecuaciones de superposición las cuales (las variables s_i^j son omitidas cuando se igualan a cero) son:

$$s_1^1 \leq m_1, \quad s_{14}^1 \leq m_1, \quad s_1^1 + s_{14}^1 \geq m_1 \quad (8.26)$$

$$s_2^2 + s_4^2 \geq m_2, \quad s_2^2 \leq m_2, \quad s_4^2 \leq m_2 \quad (8.27)$$

$$s_3^3 = m_3 \quad (8.28)$$

$$s_{14}^4 + s_2^4 + s_4^4 \geq m_4, \quad s_{14}^4 \leq m_4, \quad s_2^4 \leq m_4, \quad s_4^4 \leq m_4 \quad (8.29)$$

$$s_3^5 \leq m_5, \quad s_4^5 \leq m_5, \quad s_3^5 + s_4^5 \geq m_5 \quad (8.30)$$

Base de hechos. las ecuaciones que en la base de hechos se obtienen a partir de las medidas, cambiarán para cada caso de diagnóstico. Considere, por ejemplo, la siguiente información medida “ m_1 se encuentra presente; al menos una de m_2 o m_3 se encuentra presente; m_4 esta ausente; m_5 no esta medida”. Entonces la base de hechos (también ilustrada en la figura) será:

$$m_1 = 1, \quad m_2 + m_3 \geq 1, \quad m_4 = 0 \quad (8.31)$$

Inferencia. Para el caso, en el ejemplo en consideración, la presencia de m_1 vincula d_1 y, en consecuencia también a d_3 (lo cual explicaría lo observado en “ m_2 o m_3 ”). d_4 debe estar ausente porque de otra manera, m_4 estaría presente (el desorden compuesto d_{14} también se descarta por la ausencia de m_4). Con respecto a d_2 , este podría estar presente (pero no causando m_4). En la expresión $D = (d_1, d_2, d_3, d_4)$, los diagnósticos factibles son $(1,0,1,0)$ y $(1,1,1,0)$.

Comentario: Si las sumas se rempazan por las disyunciones y las desigualdades se rempazan por las implicaciones lógicas (ver(8.11)-(8.15)), entonces la metodología propuesta podría ser considerada equivalente a la inferencia basada en la lógica. Sin embargo, el objetivo de este trabajo es explotar la idea inversa (la inferencia lógica es equivalente a las restricciones algebraicas) con el fin de usar álgebra computacional y software de optimización (ampliamente difundidos en la actualidad). Por lo anterior, el enfoque lógico al menos en este capítulo no se usará.

8.5. El enfoque borroso

En el esquema de representación del conocimiento anterior, no se necesitará más la programación con números enteros si las variable de decisión pueden tomar cualquier valor real positivo, o posiblemente mejor, valores en el intervalo $[0, 1]$. Esto es con el fin de integrar el enfoque que aquí se presenta al marco de trabajo de la lógica borrosa estándar y evitar soluciones de PL no restringidas. Realmente, los límites entre $[0, 1]$ proveen una elegante interpretación en el esquema de los conjuntos borrosos, cuando:

- Las observaciones σ_i son el resultado de borrosificación de medidas del mundo real, que indican una “intensidad en la presencia” de una manifestación en particular, $0 \leq \sigma_i \leq 1$ (por ejemplo, indicando si o no la fiebre en un paciente es más o menos “alta”); obviamente, esto incrementa el poder de expresión en la base de conocimiento (considerando “grado de fiebre” en lugar de “fiebre contra no fiebre”).
- Las severidades en los desórdenes f_i pueden ser graduales (ocurrencia parcial del desorden; (Kitowski 1987)).
- La expresión (8.1) se representa como: “la intensidad de las manifestaciones que aparece con el desorden incrementa, así como la severidad en d_i también incrementa”.

En la interpretación anterior, borrosidad se refiere a “gradación”, esto en cumplimiento de las propiedades lingüísticas y no con respecto al tema de “incertidumbre” o “posibilidad”. Incertidumbre en la base de conocimiento o las observaciones se incorporarán más adelante a través de relajar ciertas restricciones, generalizando, por ejemplo, (8.3). Una discusión más completa en el conocimiento incierto aparece en la sección 8.6 y 8.7.

Conocimiento incierto. Ecuaciones (8.1), (8.2) y (8.3) son un caso particular de “restricciones de sector”:

$$a_i^j d_i \leq s_i^j \leq b_i^j d_i \quad (8.32)$$

donde a_i^j y b_i^j son parámetros definidos por el usuario. En este contexto:

- 1) se usa $a_i^j = b_i^j = 1$ para denotar un desorden “cierto” desorden \rightarrow asociación manifestación.
- 2) se usa $a_i^j = b_i^j = 0$ para denotar lo opuesto (por ejemplo, no-asociación).
- 3) se usa $a_i^j = 0$, $b_i^j = 1$ para denotar asociaciones “inciertas”.

Nada evita, sin embargo, el uso de diferentes valores para a_i^j y b_i^j , esto para satisfacer requerimientos inmediatos particulares (figura 8.3), aun estableciendo restricciones tales como:

$$a_i^j d_i + p_i^j \leq s_i^j \leq b_i^j d_i + q_i^j$$

Ejemplo 7. Por ejemplo, proponiendo “la manifestación m_3 es usualmente presente con el desorden d_5 , pero algunas veces m_3 aparece solo débilmente cuando d_5 se encuentra presente” podría ser interpretado como $0,4d_5 \leq s_5^3 \leq d_5$ lo cual se agregaría a la *BH*.

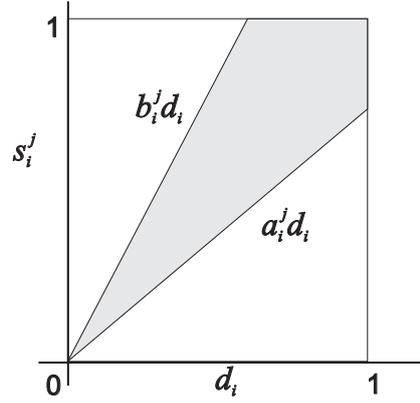


Figura 8.3: Conocimiento borroso incierto (la zona fiable aparece sombreada).

Si m_3 siempre aparece débilmente con d_5 , esta condición en la *BH* podría ser escrita como, $0,3d_5 \leq s_5^3 \leq 0,4d_5$.

En el caso borroso, (8.17) podría ser generalizado a:

$$\alpha d_i \leq \sum_{j \in J} s_i^j \leq \beta d_i \quad (8.33)$$

Ejemplo 8. La aserción “ d_1 podría causar m_1, m_2, m_3 ó m_4 : al menos una de ellas más otra débilmente, y cuando más dos de ellas serán causadas débilmente”, puede ser traducida como: $1,5d_1 \leq s_1^1 + s_1^2 + s_1^3 + s_1^4 \leq 2d_1$.

Observaciones inciertas. Algunas observaciones podrían ser inciertas o no estar disponibles. Estas pueden ser representadas por los datos de observación Σ los cuales poseen un valor lógico de intervalo $[\sigma_{l,j}, \sigma_{u,j}]$ (relacionado a los conjuntos borrosos intuitivos (Atanasov 1986; De 2001)). De esta forma, las mediciones inciertas estarán sujetas a las restricciones:

$$\sigma_{l,j} \leq m_j \leq \sigma_{u,j} \quad (8.34)$$

donde $\sigma_{l,j}$ y $\sigma_{u,j}$ son cantidades conocidas. Adicionalmente, si lo único que se conoce es que al menos una manifestación en un conjunto, por decir K , es verdadera a un grado α , entonces la restricción $\sum_{k \in K} m_k \geq \alpha$ podría agregarse para expresar tal situación, generalizando (8.16).

Ejemplo 9. Considere m_1, m_2 , las cuales son manifestaciones de temperaturas anormales (media y alta, respectivamente). Si una observación incierta se encuentra disponible como “la temperatura es media o alta”, esta se codificaría como $m_1 + m_2 \geq 1$.

Diagnosis borrosa basada en la consistencia. En este contexto, el problema de diagnosis basado en la consistencia puede entenderse como un problema factible de *PL* con números reales (Gass 2003; Sierksma 2001) donde las variables de decisión son: d_i ,

s_i^j y m_j . De allí que, la teoría elemental de PL asevera que para el conjunto factible de variables de decisión, existirá un politopo convexo. El conjunto de diagnósticos factibles, por ejemplo, el conjunto de valores factibles de D será expresado como \mathcal{D}^{feas} , representando las coordenadas de posibles desórdenes que son consistentes con las observaciones disponibles: $D \in \mathcal{D}^{feas}$ indica que existen algunos valores de s_i^j y m_j que cumplen con las restricciones especificadas en BC y en BH .

La metodología presentada en el caso borroso es sin embargo, demasiado impráctica según el escenario hasta ahora discutido. Realmente, describir \mathcal{D}^{feas} puede llegar a ser muy tedioso (posiblemente, enumerar un gran conjunto de vértices) y, aún peor, el conjunto solución \mathcal{D}^{feas} pudiese ser muy grande (por ejemplo, casi sin información cuando pocas observaciones se encuentran disponibles) o vacío. Podría ser el caso, “desorden desconocido” (restricciones no factibles), resultados con una base de conocimiento trivial “ d_1 causa m_1 y m_2 ” codificadas como $\{s_1^1 = s_1^2 = d_1, m_1 = s_1^1, m_2 = s_1^2\}$ y la base de hechos $\{m_1 = 0,5, m_2 = 0,50001\}$.

Con el fin de resolver las dificultades al momento de expresar “preferencias” en la presencia de conocimiento incierto y de agregar “suavidad” en el comportamiento de las restricciones, se discutirá en la próxima sección el esquema posibilístico.

8.6. Marco de trabajo Posibilístico

En el artículo de (Dubois 1996), se presentan los problemas de satisfacción restringida posibilística (PSR). Los autores introducen restricciones las cuales son satisfechas en cierto grado, transformando la factividad/infactividad de una solución potencial a una noción gradual: dada una PSR con variables de decisión $\delta \in \Delta$, una relación borrosa $R : \Delta \rightarrow [0, 1]$ se sugirió, esto con el propósito de representar preferencia o prioridad como un “grado de consistencia”; se creó R por la conjunción de relaciones individuales (expresando preferencia o prioridad en cada restricción individual), se definió las mejores soluciones para PSR como aquellas que satisfacen el problema global a grado máximo. En el trabajo citado, los autores establecen que “ R puede ser vista como una distribución de posibilidad, prescribiendo hasta que extensión un valor δ para las variables de decisión, puede ser considerado como apropiado de acuerdo a las restricciones existentes”.

Estas ideas serán explotadas más adelante, particularizando el operador de conjunción y la definición de relación borrosa, esto con el propósito de que resulten especialmente apropiados para los algoritmos de optimización en la programación lineal y cuadrática.

En particular, las bases de conocimiento y de hechos definirán una distribución de posibilidad $\pi : \Delta \mapsto [0, 1]$ (Dubois 1988) donde Δ denota el conjunto de todas las variables de decisión (D, M, Σ) . Entonces, se calcula la posibilidad de un evento A (subconjunto de Δ), como es usual a través de:

$$\pi(A) = \sup_{\delta \in A} \pi(\delta) \quad (8.35)$$

En particular, para un multidimensional $\Delta = \Delta_1 \times \Delta_2$, $\delta = (\delta_1, \delta_2) \in \Delta$, la posibilidad *marginal* de δ_1 se define como:

$$\pi(\delta_1) = \sup_{\delta_2 \in \Delta_2} \pi(\delta_1, \delta_2) \quad (8.36)$$

Por consiguiente, los cálculos posibilísticos son problemas de optimización.

recíprocamente, considere una función de coste $J : \Delta \mapsto \mathbb{R}^+$ (por ejemplo, verificando $J(\delta) \geq 0$ para toda $\delta \in \Delta$), así que existe $\delta_0 \in \Delta$ tal que $J(\delta_0) = 0$. Entonces, una distribución de posibilidad puede definirse en Δ vía:

$$\pi(\delta) = e^{-J(\delta)} \quad \delta \in \Delta \quad (8.37)$$

así que la posibilidad de un evento A es dada por:

$$\pi_X(A) = e^{-\inf_{\delta \in A} J(\delta)} \quad (8.38)$$

En lo subsecuente, un evento A será descrito por un conjunto de restricciones sobre δ . de esta manera, problemas de optimización de restricciones numéricas pueden ser interpretados en términos posibilísticos: el coste $J_X(s)$ se interpretará como posibilidad-log. Por definición, a los valores infactibles de las variables de decisión se les asignará posibilidad cero.

Vamos a discutir ahora como la distribución de posibilidad vista en este punto puede ser generada, en el contexto de la metodología descrita en las secciones anteriores.

8.7. Bases de conocimiento posibilístico

Hasta ahora, el problema de diagnóstico ha sido tratado bajo un esquema de programación lineal factible, esto es, cada conjunto de variables de decisión (D, M, Σ) es, o factible o infactible.

Por medio de la introducción de algunas variables artificiales y un índice de coste, las ideas anteriores habilitarán graduando a los diferentes diagnósticos candidatos como más o menos “posibles”.

Desórdenes básicos. Cada desorden básico d_i contribuirá al índice de coste vía el término lineal αd_i . La interpretación posibilística-log de tales términos es que: “la posibilidad *a priori* de un particular desorden que tiene una severidad de $0 \leq d_i$ es $e^{-\alpha_i d_i}$ ”.

De esta forma se construye un componente, que es el índice de coste de posibilidad *a priori*:

$$J_D = \alpha_1 d_1 + \alpha_2 d_2 + \dots \quad (8.39)$$

De esta manera, se asume que los desórdenes son no-interactivos (su posibilidad *a priori* es independiente de la de otros desórdenes) y a una situación normal (todos los $d_i = 0$) se le asigna posibilidad uno²

El término J_d puede ser modificado incluyendo términos de producto cruz tales como $\alpha_{ij} d_i d_j$, los cuales indican que la ocurrencia simultánea de d_i y d_j (por ejemplo, ambas diferentes de cero) son “menos posibles”. Este es el caso donde la asunción de no-actividad se relaja; se Requeriría usar programación cuadrática para resolver el problema de optimización. Después en el Ejemplo 15, se incorporará ese tipo de posibilidad.

Restricciones del conocimiento. Todas las igualdades y desigualdades que conforman nuestra base de conocimiento descrita en la sección 8.4 pueden ser reformuladas como igualdades a través de variables artificiales apropiadas. Entonces, esas variables artificiales se pueden incluir en el índice de coste para generar una distribución de posibilidad. Este procedimiento será llamado “restricciones suavizadas” y se describe en detalle a continuación.

Restricción de igualdades suavizadas. Considere una restricción de igualdad $a = b$. Una versión relajada (“suavizada”) de tal restricción puede ser escrita como:

$$a = b + \epsilon - \nu, \quad \epsilon \geq 0, \quad \nu \geq 0 \quad (8.41)$$

²Una guía alternativa para configurar un índice de coste apropiado podría ser una interpretación probabilística: una posibilidad *a priori* podría ser traducida (aproximadamente) a una densidad de probabilidad *a priori*. Una amplia discusión sobre la relación entre posibilidad y probabilidad se encuentra en (Dubois 2001b).

Asumiremos que una probabilidad *a priori* de los desórdenes d_i que tienen una severidad f_i es gobernada por una distribución de densidad proporcional a $e^{-\lambda_i f_i}$, donde λ_i es un coeficiente conocido. Asumiremos también que varios desórdenes son estadísticamente independientes. Entonces la densidad de probabilidad conjunta sería proporcional a $L(f_0, \dots, f_k) = e^{-\sum_{i=1}^k \lambda_i f_i}$. Con el fin de configurar una regla aproximada para determinar λ_i , consideremos el caso borroso con una función de densidad exponencial, que produce la expresión:

$$P(d_i \geq a) = \frac{\int_a^1 e^{-\lambda_i \psi} d\psi}{\int_0^1 e^{-\lambda_i \psi} d\psi} = \frac{e^{-\lambda_i a} - e^{-\lambda_i}}{1 - e^{-\lambda_i}} \quad (8.40)$$

Por ejemplo, si “desorden significativo” es entendido como uno con severidad más alta que 0.66, la aserción “10 % de los casos de diagnóstico exhiben un desorden d_i con un grado significativo” resultará en definir a $\lambda_i = 2,8$ (que se obtuvo resolviendo para λ_i , esto es, $P(d_i \geq 0,66) = 0,1$). Si la frecuencia fue 5 %, $\lambda_i \approx 4,1$; si esta fuere 1 %, $\lambda_i \approx 6,8$, uno en mil, $\lambda_i \approx 10,4$. Aserciones de probabilidad como la Gaussiana, por ejemplo, funciones de densidad proporcionales a $e^{-\lambda_i f_i^2}$ pueden ser integradas transparentemente dentro de este enfoque (usando programación cuadrática).

siendo ϵ y ν variables artificiales. Si ϵ y ν han sido penalizadas en un índice de optimización (típicamente con términos de índice de coste, como, ϵ o ϵ^2), la interpretación posibilística de (8.41) puede ser: “ $a = b$ es posible, lo menos posible es la situación donde a difiera de b ”.

Ejemplo 10. Una asociación $s_1^1 = d_1$ puede ser “suavizada” declarando

$$s_1^1 = d_1 + \epsilon - \nu, \quad \epsilon \geq 0, \quad \nu \geq 0$$

y construyendo un índice de coste

$$J(s_1^1, d_1, \epsilon, \nu) = \nu^2 + 7\epsilon$$

De esta forma, la posibilidad de $(s_1^1 = 0,5, d_1 = 0,5)$ es 1 (el coste de J es cero), lo de $(s_1^1 = 1, d_1 = 0,5)$ es $e^{-7*0,5} = 0,03$, y lo de $(s_1^1 = 0, d_1 = 0,5)$ es $e^{-0,5^2} = 0,78$. En términos lingüísticos: con el desorden d_1 ocurriendo a una intensidad media, es “bastante posible” para la manifestación m_1 no aparecer, y es “casi imposible” para esta aparecer a un grado más alto que el de d_1 . La figura 8.4 describe la distribución de posibilidad. Debido a la concepción de J , el suavizamiento ha sido hecho intencionalmente no-simétrico (sería simétrico con $J = 7\epsilon + 7\nu$).

Desigualdades. Una restricción de desigualdad “suavizada” no es otra cosa que una igualdad sin penalización en sus variables artificiales.

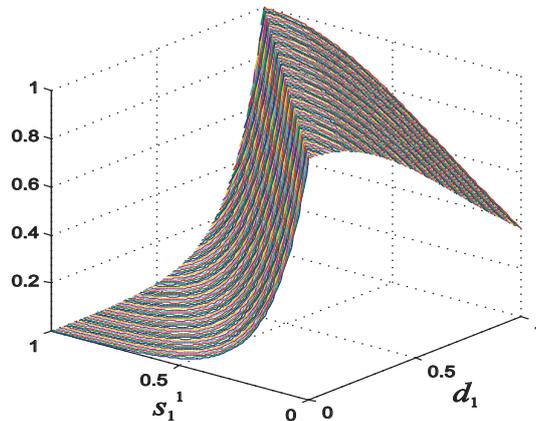


Figura 8.4: Distribución de posibilidad proveniente de restricciones de igualdades “suavizadas asimétricamente”.

Ejemplo 11. Considerando el Ejemplo 10, si el índice de coste fuese:

$$J(s_1^1, d_1, \epsilon, \nu) = 7\epsilon$$

todos los pares (s_1^1, d_1) con $s_1^1 \leq d_1$ tendrían posibilidades igual a 1 (coste cero). La posibilidad de $(s_1^1 = 1, d_1 = 0,5)$ es aun $e^{-7*0,5} = 0,03$.

Ejemplo 12. El uso de más variables artificiales permiten una mayor flexibilidad en la representación propuesta. Para este caso, expresando:

$$s_1^1 = d_1 + \epsilon_1 + \epsilon_2 - \nu_1 - \nu_2 \quad (8.42)$$

$$0 \leq \epsilon_1 \leq 0,2d_1, \quad 0 \leq \epsilon_2, \quad 0 \leq \nu_1 \leq 0,2d_1, \quad 0 \leq \nu_2 \quad (8.43)$$

con un asociado:

$$J(s_1^1, d_1, \epsilon_1, \epsilon_2, \nu_1, \nu_2) = 0,4\epsilon_1 + 7\epsilon_2 + 0,4\nu_1 + 3\nu_2$$

se induce una distribución de posibilidad. La posibilidad del par (s_1^1, d_1) es calculada, a partir de (8.38), usando el valor mínimo del índice de coste sobre el conjunto factible de variables de decisión $(\epsilon_1, \epsilon_2, \nu_1, \nu_2)$. Como todas las restricciones anteriores son lineales, el problema puede ser resuelto por programación lineal. Para el caso, la posibilidad de $(s_1^1 = 0,55, d_1 = 0,5)$ es $e^{-0,4*0,05} = 0,98$, que corresponde a la posibilidad del singleton en el espacio de la variable de decisión dado por $(s_1^1 = 0,55, d_1 = 0,5, \epsilon_1 = 0,05, \epsilon_2 = \nu_1 = \nu_2 = 0)$. Con los valores anteriores de posibilidad, haciéndolos variar en un rango apropiado para s_1^1 y d_1 , se puede calcular la distribución de posibilidad mostrada en la figura 8.5. La configuración propuesta en este ejemplo puede ser usada para suavizar restricciones de sector (8.32).

Manifestaciones y Observaciones. Si el valor medido de la manifestación m_j es σ_j , la proposición básica $m_j = \sigma_j$ puede ser relajada con variables adicionales, esto de acuerdo a la “fiabilidad” de cada sensor, conforme a una distribución de posibilidad en (m_j, σ_j) asociada a algún término de índice de coste J_M .

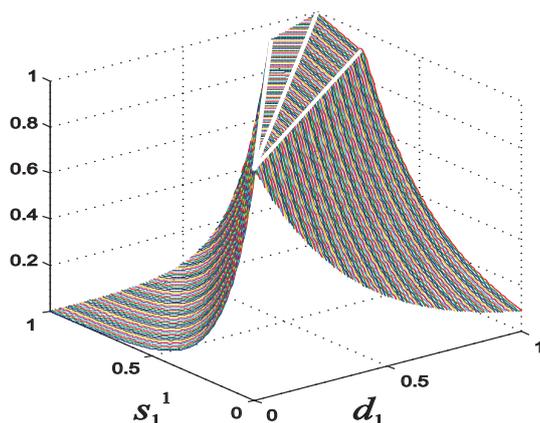


Figura 8.5: Distribución de posibilidad con 4 variables artificiales: restricciones de sector suavizadas.

Ejemplo 13. análogamente para (8.12), una medición intervalar puede ser descrita vía:

$$m_j = \sigma_j + \epsilon_1 - \nu_1 + \epsilon_2 + \nu_2, \quad 0 \leq \epsilon_1 \leq 0,05, \quad 0 \leq \nu_1 \leq 0,05, \quad 0 \leq \epsilon_2, \quad 0 \leq \nu_2$$

y un índice de coste incorporando el término $5\epsilon_2 + 5\nu_2$. La restricción describe una medida intervalar, donde la posibilidad de $\sigma_j \in [m_j - 0,05, m_j + 0,05]$ es uno y la posibilidad de que σ_j este fuera del intervalo referido depende del peso del índice de coste. Por ejemplo, con $\sigma_j = 0,1$, la posibilidad de $m_j = 0,95$ sería $e^{-5*0,8} = 0,018$, la posibilidad de $m_j = 0,14$ sería uno.

Conjuntos de ecuaciones: el costo total asociado al problema de diagnóstico.

Considere un conjunto de ecuaciones (restricciones suavizadas) con sus variables artificiales asociadas $\delta_1, \delta_2, \text{etc.}$ tal y como se definió anteriormente, con un índice de coste individual $J_1(\delta_1) + J_2(\delta_2), \text{etc.}$ El índice de coste definido por:

$$J(\delta_1, \delta_2, \dots) = J_1(\delta_1) + J_2(\delta_2) + \dots \quad (8.44)$$

será la elección para definir una distribución de posibilidad en el espacio del producto, interpretando el operador de conjunción posibilístico en (Dubois 1996) como un producto³ algebraico, transformado a una suma después de que se obtiene su logaritmo.

Como resultado, se puede construir un índice de coste total con razonable facilidad añadiendo todos los componentes de las restricciones individuales más las posibilidades *a priori* de los desórdenes. De esta manera, se formará una base de conocimiento posibilístico a partir de un conjunto de restricciones, donde las variables involucradas son:

- 1) Las severidades de los desórdenes d_i
- 2) Las variables de asociación s_i^j
- 3) Las variables de manifestaciones m_j
- 4) Los datos de mediciones σ_j
- 5) Las variables artificiales ϵ, ν necesarias para “suavizar” las restricciones básicas de la base de conocimiento y la base de hechos.

El valor del índice de coste refleja la posibilidad-log de una combinación particular de las variables de decisión anteriores. Las próximas secciones proveen ejemplos detallados del procedimiento propuesto.

8.7.1. Resultados del diagnosis posibilístico

Una vez que el conjunto de mediciones σ_j este disponible, la salida del sistema de diagnóstico deberá ser la combinación de coste mínimo (posibilidad máxima) de las

³La elección del producto algebraico como el operador de agregación se basa en heurísticas: se necesitan las restricciones más difíciles de satisfacer para poder explicar una hipótesis, en vez de eso se escoge la menos difícil de satisfacer que es la posibilidad de un resultado de diagnóstico (en lugar de buscar sólo a la restricción “más difícil de satisfacer” si el *mínimo* fuese usado como operador de agregación). De cualquier manera, si J_i fuese lineal y $J(s_1, s_2, \dots)$ fuese definida como el *máximo* $J_i(s_i)$, esto correspondería al operador de conjunción *mínimo* en el espacio posibilístico. Si esa fuese la elección, entonces los cálculos de posibilidad serían un problema de solución conocida de *minimax* lineal, fácil de resolver (añadiendo una nueva variable γ , las restricciones $J_i(\delta_i) \leq \gamma$ y minimizando γ como término único en el índice de coste).

variables de decisión, después de haber fijado σ_j a los valores obtenidos. El valor de posibilidad obtenido es, por lo tanto, la posibilidad “a priori” de encontrar dichas mediciones.

Sin embargo, en una situación donde algunos resultados de diagnóstico múltiple pudiesen ser razonablemente posibles, existiría la necesidad de informar al usuario final sobre las opciones alternativas con las que cuenta.

Convenientemente, podemos diseñar fácilmente distribuciones de posibilidad marginal: la posibilidad marginal (8.36) de $d_i = \delta$, (donde δ es un valor numérico conocido) se calcula sumando $d_i = \delta$ a la BH y minimizando J sobre el resto de variables de decisión. Tal posibilidad marginal puede ser diseñada en un rango de valores para δ en el intervalo $[0,1]$.

Además, si una definición del producto para la posibilidad condicional⁴ de un evento “A” dado un evento “B” fuese usado (por ejemplo, $\pi(A|B) := \pi(A \cap B)/\pi(B)$), se podrían calcular fácilmente las posibilidades condicionales.

De hecho, vamos a designar a “B” como el conjunto de variables de decisión compatible con las restricciones establecidas en la base de conocimiento y en la base de hechos $\{BC, BH\}$; y “A” será designada como un evento descrito por un conjunto de restricciones $\{Q\}$. Entonces:

$$\pi(A|B) := \frac{\pi(A \cap B)}{\pi(B)} = \frac{\exp(\min_{\text{restricciones}\{BC, BH, Q\}} J)}{\exp(\min_{\text{restricciones}\{BC, BH\}} J)} \quad (8.45)$$

Obteniendo los logaritmos, la posibilidad-log condicional es dada por el incremento de costo que resulta de agregar las restricciones $\{Q\}$.

En las próximas secciones se ilustrará el uso de la posibilidad condicional por medio de ejemplos.

8.8. Ejemplos adicionales

Ejemplo 14. Vamos a considerar el problema de diagnóstico binario en (Yamada 2002) §5, donde un conjunto de 3 enfermedades y 3 manifestaciones (binarias) son definidas. Se denotaran los desórdenes como: $\{d_1$: fiebre de heno, d_2 : gripe, d_3 : intoxicación alimentaria $\}$. Las manifestaciones se denotan como: $\{m_1$: congestión nasal, m_2 : fiebre alta, m_3 : diarrea $\}$.

La base de conocimiento en la referencia citada provee las posibilidades condicionales para cada uno de los $3 \times 3 \times 2 = 18$ casos (se consideran ambas manifestaciones y sus negaciones). Para este caso,

⁴Aparte de la fórmula de producto considerada aquí, han sido propuestas en la literatura (Dubois 2001b; Cooman 1997) diferentes definiciones para la posibilidad condicional.

$$\pi(\text{congestion nasal}|\text{fiebre de heno}) = 1, \quad \pi(\overline{\text{congestion nasal}}|\text{fiebre de heno}) = 0,3$$

denotan que la posibilidad de que la fiebre de heno cause una congestión nasal es 1, pero su necesidad es 0.7 (la posibilidad de no causar el síntoma es de 0.3). En el enfoque actual, la posibilidad condicional provista nos da información a cerca de la variable de asociación s_1^1 : la aserción anterior es traducida al ámbito de las restricciones:

$$s_1^1 = d_1 - \epsilon, \quad 0 \leq \epsilon \leq d_1 \quad (8.46)$$

junto con un término de adición en el índice de coste total como: $-\ln(0.3)\epsilon$, esto es, 1.204ϵ .

Como otro ejemplo, la aserción

$$\pi(\text{diarrea}|\text{gripe}) = 0,4, \quad \pi(\overline{\text{diarrea}}|\text{gripe}) = 1$$

se traduce como:

$$s_2^3 \leq d_2, \quad s_2^3 = \epsilon_5 \quad (8.47)$$

agregando un término en el índice de coste $-\ln(0.4)\epsilon_5 = 0,92\epsilon_5$.

Siguiendo esta metodología, se puede expresar completamente la base de conocimiento posibilístico de (Yamada 2002) desarrollada a continuación:

$$\begin{aligned} s_1^1 &= d_1 - \epsilon_1, & s_2^1 &= \epsilon_3, & s_3^1 &= 0 \\ s_1^2 &= \epsilon_2, & s_2^2 &= d_2 - \epsilon_4, & s_3^2 &= \epsilon_6 \\ s_1^3 &= 0, & s_2^3 &= \epsilon_5, & s_3^3 &= d_3 - \epsilon_7 \\ 0 \leq \epsilon_1 &\leq d_1, & 0 \leq \epsilon_2 &\leq d_1, & 0 \leq \epsilon_3 &\leq d_2, & 0 \leq \epsilon_4 &\leq d_2 \\ 0 \leq \epsilon_5 &\leq d_2, & 0 \leq \epsilon_6 &\leq d_3, & 0 \leq \epsilon_7 &\leq d_3 \end{aligned}$$

junto con las ecuaciones de superposición

$$\begin{aligned} m_1 &\leq s_1^1 + s_2^1 + s_3^1, & m_2 &\leq s_1^2 + s_2^2 + s_3^2, & m_3 &\leq s_1^3 + s_2^3 + s_3^3 \\ s_i^j &\leq m_j, & i &= 1, \dots, 3, & j &= 1, \dots, 3 \end{aligned}$$

con un índice de coste:

$$J = 0,51d_1 + 0d_2 + 0,92d_3 + 1,2\epsilon_1 + 1,61\epsilon_2 + 0,36\epsilon_3 + 1,2\epsilon_4 + 0,92\epsilon_5 + 0,69\epsilon_6 + 2,3\epsilon_7$$

En el caso particular considerado en (Yamada 2002), se tiene el conocimiento de datos de observación a cerca de la presencia de *fiebre de heno*, $d_1 = 1$, así como de *fiebre alta*, $m_2 = 1$, esto es, $BH = \{d_1 = 1, m_2 = 1\}$.

Dada la base de conocimiento, la posibilidad de los datos medidos es $e^{-0,51} = 0,6$, que corresponden al valor mínimo de J obtenido después de haber definido completamente todas las restricciones en BC y BH. Estaremos usando $\pi(\{BC, BH\}) = 0,6$ como notación. El conjunto de variable de decisión con más posibilidad para producir $J = 0,51$ es (sólo se muestran variables con valor no cero):

$$d_1, d_2, m_1, m_2, s_1^1, s_2^2 \quad (8.48)$$

puntualizando que, posiblemente, *fiebre alta* (m_2) es causada por *gripe* (d_2), como $s_2^2 = 1$, y que la presencia de m_1 (*congestión nasal*) es causada por d_1 (*fiebre de heno*), como $s_1^1 = 1$.

Para tener una mejor descripción del diagnosis resultante, la posibilidad marginal para *no gripe* (dadas las observaciones corrientes) pueden calcularse al agregar $d_2 = 0$ a las restricciones de igualdades en la BH. El coste mínimo es $J = 2,12$, esto se explica como “ $d_1 = d_3 = \epsilon_6 = 1$, el resto de las variables son cero”, interpretado como “*fiebre de heno* y *intoxicación alimentaria*, *intoxicación alimentaria* es responsable de la *fiebre alta* que se encuentra presente”. La explicación alternativa “ $d_1 = \epsilon_2 = 1$, el resto de las variables son cero” también tiene el mismo coste (“*fiebre de heno* esta causando *fiebre*”). La exponencial del coste alcanzado puede ser considerado como la posibilidad $e^{-2,12} = 0,12$ de no gripe.

La posibilidad condicional de $d_2 = 1$ (condicionada por la base de conocimiento y de hechos disponible) es 1, porque los rasgos de la explicación más posible son $d_2 = 1$. La posibilidad condicional de $d_2 = 0$ es dada por:

$$\pi(\{d_2 = 0\}|\{BC, BH\}) = \frac{\pi(\{d_2 = 0\} \cup \{BC, BH\})}{\pi(\{BC, BH\})} = \frac{0,12}{0,6} = 0,2$$

Las posibilidades condicionales para todas las variables de decisión pueden ser calculadas de la misma forma. Por ejemplo, $\pi(\{m_1 = 1\}|\{BC, BH\}) = 1$, $\pi(\{m_1 = 0\}|\{BC, BH\}) = 0,3$, etc.. Los resultados numéricos en este caso son coincidentes con aquellos encontrados en (Yamada 2002); sin embargo, en el caso general no se busca igualdad en los resultados.

Ejemplo 15. Este ejemplo ilustra brevemente las ventajas del enfoque posibilístico en situaciones donde las mediciones son tan pocas que es difícil distinguir entre los diferentes desórdenes con manifestaciones similares. También, el ejemplo mostrará la gran utilidad del enfoque en situaciones de inconsistencia que derivan de mediciones defectuosas.

Considere la base de conocimiento en el Ejemplo 2 describiendo la relación entre tres desórdenes y cuatro manifestaciones anormales, refinada con la información posibilística siguiente:

- Desorden d_1 causa m_1 , m_2 y m_3 , pero no m_4 . La posibilidad-log *a priori* de d_1 es $-2,8d_1$ (lo cual puede ser interpretado como “ d_1 ocurre significativamente en el 10% de los casos”, conforme a (8.40))
- d_2 causa m_1 , m_2 , y m_4 con sus efectos en m_3 siendo parcialmente incierto: $0,5d_2 \leq s_2^3 \leq d_2$. La posibilidad-log *a priori* de d_2 es $-4,1d_2$ (“ d_2 ocurre significativamente en el 5% de los casos”)
- d_3 causa m_1 , pero no m_2 ni m_3 ; sus efectos sobre m_4 son inciertos (en la mayoría de las ocasiones m_4 se encuentra presente pero la posibilidad de que m_4 no aparezca cuando d_3 se encuentre presente es 0.1; traducida como $d_3^4 = d_3 - \epsilon$ junto con $-\ln(0,1)\epsilon = 2,3\epsilon$ en el índice de coste). La posibilidad-log *a priori* de d_3 es $-4,5d_3$.

La base de conocimiento discutida en el Ejemplo 2 ha sido modificada en la siguiente forma:

$$\begin{aligned} s_1^1 = s_1^2 = s_1^3 = d_1, & \quad s_2^1 = s_2^2 = s_2^4 = d_2, & \quad 0,5d_2 \leq s_2^3 \leq d_2 \\ s_3^1 = d_3, & \quad s_3^2 = s_3^3 = 0, & \quad s_3^4 = d_3 - \epsilon \end{aligned}$$

junto con las ecuaciones de superposición, las ecuaciones de mediciones inciertas $m_1 = \sigma_1 + \epsilon_1 - \nu_1$, $m_2 = \sigma_2 + \epsilon_2 - \nu_2$, $m_3 = \sigma_3 + \epsilon_3 - \nu_3$, $m_4 = \sigma_4 + \epsilon_4 - \nu_4$ y restricciones sobre todas las variables que se encuentran en el intervalo $[0,1]$. El índice de coste de posibilidad-log ha sido definido como:

$$J_1 = 2,8d_1 + 4,1d_2 + 4,5d_3 + 2,3\epsilon + 10 \sum_{i=1}^4 (\epsilon_i + \nu_i)$$

donde la posibilidad de una medición totalmente contraria al valor actual de una manifestación⁵ se propone que sea e^{-10} .

Además, términos de producto cruz se han añadido penalizando $20d_i d_j$ para indicar preferencia por los resultados que contengan desórdenes únicos; también se han agregado penalizaciones de producto cruz en los términos de error en los sensores para la ϵ positiva y la ν negativa, para penalizar la combinación de errores de sensor con el mismo signo ($2\epsilon_i \epsilon_j$ y $2\nu_i \nu_j$), juzgando *a priori* el menos posible. El índice de coste total es, entonces

⁵De esta manera, ninguna observación tendrá posibilidad cero y se evitaban los problemas de no factibilidad. Esto es paralelo al enfoque Bayesiano donde se evitan los eventos de probabilidades cero, esto con el propósito de poder calcular probabilidades condicionales significativas en cualquier situación.

$$J = J_1 + 20(d_1d_2 + d_1d_3 + d_2d_3) + 2(\epsilon_1\epsilon_2 + \epsilon_1\epsilon_3 + \epsilon_1\epsilon_4 + \epsilon_2\epsilon_3 + \epsilon_2\epsilon_4 + \epsilon_3\epsilon_4) + 2(\nu_1\nu_2 + \nu_1\nu_3 + \nu_1\nu_4 + \nu_2\nu_3 + \nu_2\nu_4 + \nu_3\nu_4)$$

Caso 1: Con el fin de ilustrar la salida de diagnóstico cuando la presencia de muchos desórdenes es posible, consideremos que la única medición disponible es m_1 , por ejemplo, $\sigma_1 = 0,9$. Como m_1 es producida por todos los desórdenes, todos son parcialmente posibles. El vector del desorden más posible que resulta de la optimización de J es $(0,9,0,0)$, con una posibilidad de $e^{-2,8*0,9} = 0,08$. Las posibilidades marginales de cada desorden se muestran en la figura 8.6: Estas indican que existe alguna posibilidad (0.025) con respecto a aceptar $d_2 = 0,9$ como un diagnóstico, así como algunas evidencias (0.017) sobre la posibilidad de $d_3 = 0,9$. Ante la escasa información sobre mediciones, el resultado esta de acuerdo con las posibilidades *a priori*, como intuitivamente se esperaba. La posibilidad condicional escalando las gráficas en la figura 8.6 proponiendo que la máxima altura (máxima posibilidad) ahora sea igual a 1: $d_1 = 0$ es completamente posible, pero la posibilidad condicional de $d_2 = 0,9$ es 0.31, y la de $d_3 = 0,9$ es 0.22.

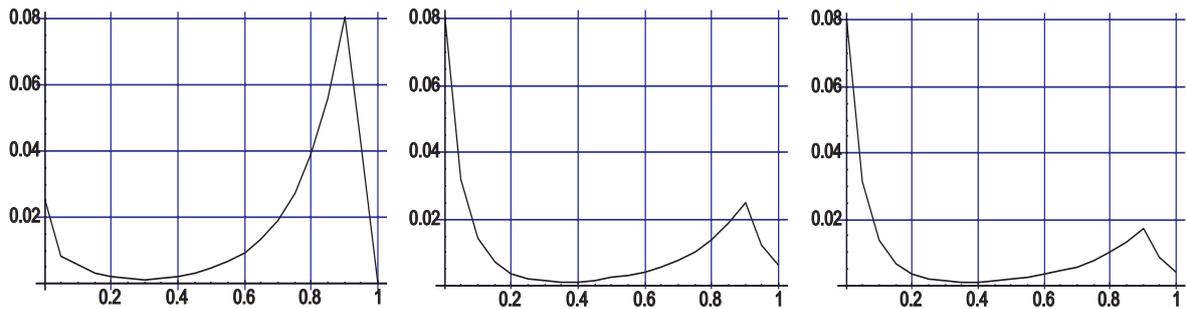


Figura 8.6: Posibilidades marginales en el caso 1. Abscisas: severidad (de izquierda a derecha: d_1, d_2, d_3); ordenadas: posibilidad

Caso 2: $\sigma_1 = 0,9, \sigma_4 = 0,02$. En este caso, el diagnosis más probable es como el anterior: $d_1 = 0,9, d_2 = d_3 = 0$, con una posibilidad de 0.066. La figura 8.7 ilustra las posibilidades marginales que se obtuvieron, donde ahora d_2 básicamente es descartada por la nueva medida σ_4 . Existe, sin embargo, una situación donde $d_3 = 0,9$ evidencia ser la causa (debido a que d_3 en algunos casos esta de acuerdo con los patrones de comportamiento de los sensores) con una posibilidad de 0.002.

Caso 3: Consideremos ahora $\sigma_1 = 0,8, \sigma_2 = 0,77, \sigma_3 = 0,05$. Esta es una situación “casi imposible” (la posibilidad pico es 0.001). En realidad, no hay un desorden en la base de conocimiento que parezca ser capaz de proveer tal patrón de comportamiento: m_2 y m_3 deben tomar el mismo valor para todos los desórdenes codificados.

La baja posibilidad de combinación que tienen las medidas presentes, debe de usarse como una advertencia de una situación inesperada. Sin embargo, la inferencia concluye que $d_2 = 0,77$ es el diagnosis más posible (ver figura 8.8), junto con un fallo en el sensor 3 con lectura de 0.34 unidades menos que las esperadas (note que m_2^3 es parcialmente

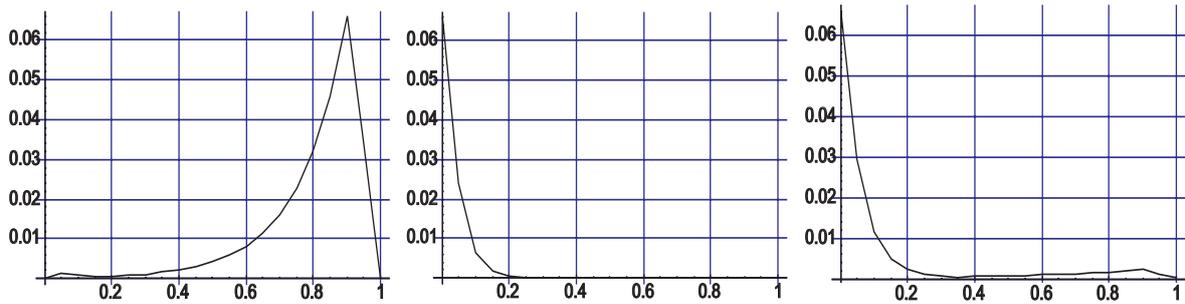


Figura 8.7: Posibilidades marginales en el caso 2

incierto, pero $d_2 = 0,77$ debe producir al menos $m_3 \geq 0,385$, el cual no esta de acuerdo con el valor medido). El mismo patrón podría estar siendo producido por una situación menos posible, tal como $d_1 = 0,77$ y un fallo de sensor 3 más grande, como lo muestra el marginal de d_1 .

Note como, transparentemente, los fallos de sensores son integrados en el esquema así como *alertando* que el patrón actualmente presente es uno realmente “raro”: al cual se le debe dar particular atención o un trato fuera de lo común⁶.

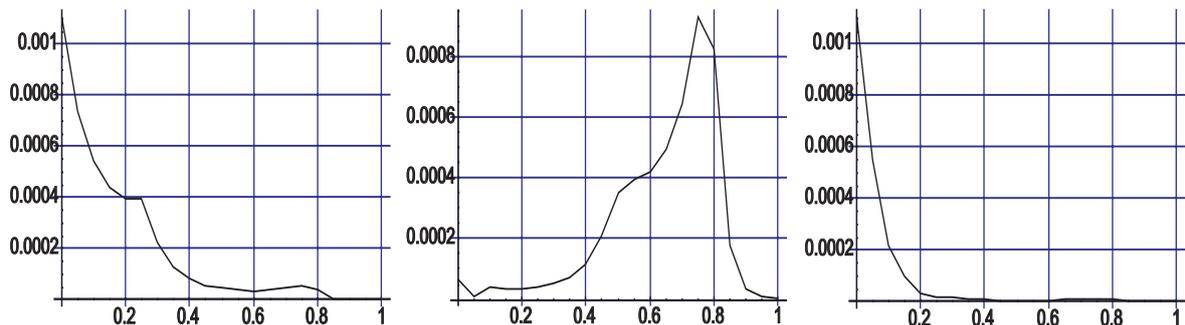


Figura 8.8: Posibilidades marginales en el caso 3

⁶Existen ciertas alternativas para estos casos:

- Considerando que σ_3 transporta una información parcialmente correcta a cerca de s_3 siendo esta “baja”, por lo que se acepta al desorden 2 (0.77) como el diagnóstico.
- Considerando que σ_3 es una medición totalmente incierta y, por lo tanto se presentaría el resultado de diagnóstico inferido sólo de σ_1 y σ_2 , removiendo $\{\sigma_3 = 0,05\}$ de la base de hechos (lo cual señalaría a d_1 como el desorden más relevante).
- Obteniendo σ_3 otra vez (revisando los procedimientos de mediciones con el fin de mejorar su fiabilidad) o una manifestación relacionada diferente la cual podría proveer información adicional.
- considerando σ_3 como una pieza de datos fiable y, por lo tanto, concluir que estamos ante un desorden desconocido (por ejemplo, si al repetir la medición se confirma el valor antes medido, que aunque no es el que esperábamos resulta ser real).

Claro, en un caso práctico, la elección de la alternativa más adecuada dependerá de la aplicación particular y del modo de fallo del sensor.

8.9. Discusión comparativa

La equivalencia entre desigualdades e implicaciones, y entre adiciones y conjunciones ha sido discutido previamente, indicando que el enfoque binario debe ser tratado a través de la inferencia lógica. Con respecto a la metodología para el conocimiento borroso y el incierto, pueden considerarse otras comparaciones relevantes, como se discute a continuación.

Comparación con los enfoques relacionales borrosos. La lógica borrosa se usa en (Sanchez 1976; Tsukamoto 1977) con el propósito de establecer una “relación borrosa” R entre desórdenes y manifestaciones como una representación del conocimiento básica. En particular, de ambas borrosidad (gradualidad) e incertidumbre requeriría definir relaciones borrosas inciertas, como las vistas en (De 2001). Para llevar a cabo el diagnóstico se usan operadores composicionales. Sin embargo, el significado de tales relaciones borrosas y la elección de operadores de composición no es muy clara después de todo; citando a (Dubois 1995), “En la literatura las ecuaciones relacionales borrosas para diagnóstico, normalmente son una causa de preocupación en los problemas representacionales”. Considerando este trabajo, la distribución de posibilidad implícitamente describe vía las restricciones y el coste, el poder ser interpretadas como una clase de relación borrosa incierta. Otra contribución del trabajo actual es ilustrar que las relaciones medición-manifestación-desorden puede ser expresada vía restricciones lineales e índices de coste lineal o cuadrático, y que al mismo tiempo puede conservar ambas cosas un razonablemente útil *poder expresivo* (correlaciones, desórdenes múltiples, abducción, muestreo de sensores con mediciones espurias) y una semántica *entendible*. La punto clave es la adición de variables artificiales o auxiliares.

Comparación con el razonamiento causal abductivo. El razonamiento causal se basa en la definición de las manifestaciones causadas ($M(d_i)^+$) y las no causadas ($M(d_i)^-$) por un desorden, consideradas en la sección 8.6. Entonces, para tratar con las diferentes alternativas de un diagnóstico consistente, se usan algunos criterios “abductivos” (Cayrac 1995), particularmente en (Dubois 1995) “diagnóstico (binario) posibilístico con incertidumbre graduada”. En el trabajo actual se extienden esos resultados debido a que incorpora ambas cosas incertidumbre graduada y severidad graduada. Además, los refinamientos introducidos en la sección 8.4.2 permiten descripciones más finas de las interrelaciones entre desórdenes y manifestaciones que sólo los conjuntos $M(d)^+$ y $M(d)^-$ originalmente usados en (Dubois 1995), al mismo tiempo conserva maleabilidad computacional.

Con respecto a la comparación con (Dubois 1995), el enfoque presentado por Yamada también considera la posibilidad como una incertidumbre graduada de eventos binarios, por lo que, el enfoque presentado en esta tesis puede ser considerado más flexible. El Ejemplo 4 considerado en una sección previa muestra como la traducción de las bases de conocimiento de Yamada a ecuaciones lineales, puede realizarse fácilmente transformando posibilidades más bajas que 1 a pesos relacionados por preferencia.

Es de notarse también, la integración transparente en el trato de errores de sensor

ilustrada en el Ejemplo 15, que puede ser considerada ventajosa con respecto a otras aportaciones en la literatura.

Sin embargo, existen otras posibilidades de abducción (cardinalidad mínima, preferida para fallos únicos, para fallos cuyas manifestaciones “ciertas” han sido totalmente observadas, *etc.* (Cayrac 1995; Mouzon 2000)) las cuales no han sido consideradas, en favor del uso del enfoque Bayesiano de las posibilidades *a priori* y *a posteriori* que se sugiere para trabajos futuros.

Problemas computacionales. El principal objetivo de esta contribución ha sido evitar los enfoques basados en lógica o en conjuntos (aun si estos resulten interesantes y fructíferos, y posiblemente equivalentes en muchos aspectos a la metodología propuesta) prefiriendo explotar las equivalencias algebraicas, esto con el propósito de usar software de optimización estándar. La programación lineal o cuadrática puede resolver problemas de optimización con cientos y miles de variables de decisión en un tiempo muy razonable, por esto, la metodología propuesta puede escalar a bases de conocimiento de diagnóstico complejas. También el software necesario se encuentra ampliamente disponible; por ejemplo, la base de conocimiento y de hechos puede ser literalmente tecleada en software de álgebra de cálculos simbólicos tal como *Mathematica*®.

Conclusiones del capítulo

En este capítulo se ha presentado una metodología para codificar el razonamiento de diagnóstico basado en el conocimiento como un problema de optimización. Considerando los casos binarios, la metodología es capaz de incorporar estructuras de conocimiento de enfoques basados en la inconsistencia así como algunos basados en la posibilidad, al menos aproximadamente. Además, este trabajo incorpora las interesantes características siguientes: a) la habilidad de incorporar transparentemente restricciones de “correlación” como desigualdades envolviendo d_i , m_i , s_i^j lo cual provee en algunos casos más expresividad que en los enfoques referidos; b) la habilidad para poder tratar naturalmente desórdenes múltiples, particularmente fallos (poco comunes) de sensores; c) el uso simultáneo de posibilidad (abducción) y gradualidad (borrosidad), produciendo distribuciones de posibilidad a cerca del verdadero valor (severidad) de los desórdenes; d) la habilidad de expresar el conocimiento de diagnóstico como un problema de programación lineal o cuadrática, para lo cual existe código muy eficiente. La disponibilidad de código eficiente hace la presente metodología apropiada para casos a gran escala, al menos en el caso lineal.

Capítulo 9

Aplicación del diagnóstico posibilístico borroso a la detección de fallos en motores de combustión interna mediante análisis de aceite

9.1. Introducción

En este capítulo se presenta una aplicación a la aportación principal de esta tesis, que es el “*diagnóstico de desórdenes basado en el razonamiento borroso possibilístico*”. Primero se presenta, una explicación sobre el diagnóstico de motores de combustión interna por el análisis de aceite usado. Después se presenta, la interpretación de la base de conocimientos provista por el experto, que en la metodología propuesta ese conocimiento puede expresarse en un diagrama causa-efecto. El conocimiento a utilizar es expresado a través de reglas borrosas tales como:

Si la severidad del desorden “ i ” es d_i , $0 \leq d_i \leq 1$, y es el único desorden que esta ocurriendo en el sistema, entonces la intensidad de las manifestaciones m_{i_1}, \dots, m_{i_p} es aproximadamente “ d_i ”, y la intensidad del resto de ellas es cero.

Este conocimiento se convierte en un sistema de ecuaciones lineales que luego serán usadas por el algoritmo de optimización lineal. Se introduce el uso las variables auxiliares vistas en el capítulo 8 y el índice de coste donde se especifica las posibilidades de cada desorden. A partir de muestras de aceites usados se obtienen nuestras medidas σ_i (observaciones) que después de un preproceso con reglas de inferencia borrosas son convertidas en nuestras manifestaciones m_j . Una vez colectada la información necesaria hacemos uso del optimizador. Este optimizador fue desarrollado en un software de aplicación llamado MATHEMATICA, del cual tomamos principalmente la función “*Minimize*”, la cual al ejecutarse nos dará los resultados de diagnóstico que posteriormente analizaremos. Por último estudiaremos los resultados de posibilidad condicional que también obtendremos.

9.2. Diagnóstico mediante análisis de aceite

El análisis del lubricante usado en los motores de combustión interna es una de las herramientas del mantenimiento predictivo; esto es, el mantenimiento realizado en base a un deterioro de un equipo (en este caso un motor), señalado por la variación de un parámetro controlado e indicativo del funcionamiento o rendimiento de la máquina.

Después de la II Guerra Mundial aparece como técnica aplicable y comienza siendo utilizada dentro del campo ferroviario norteamericano para el control del estado de los motores y poco a poco comienza a extenderse a otros campos de aplicación y a otros países, actualmente está implantado en el campo aeronáutico, flotas de transportes y militar, donde se conoce generalmente como JOAP (*Joint Oil Analysis Program*) (Tormos 2002).

Los objetivos del mantenimiento predictivo son: aumento en la fiabilidad de los equipos, aumento de la vida eficaz de los mismos, y garantizar la seguridad de instalaciones así como de manipuladores y usuarios. En este mismo tenor el mantenimiento según condición (*Condition Monitoring*) que se logra a través del análisis del lubricante usado en los motores de combustión interna, logra beneficios en lo referente a: reducción de costos de mantenimiento, mejora en seguridad y productividad.

El análisis de lubricante usado es comparable con el análisis médico de sangre. Como las pruebas sanguíneas, el aceite lubricante contiene una buena cantidad de información a cerca del medio donde este fue usado. Pero el factor crítico en el análisis de aceite, es la interpretación de resultados de las diferentes pruebas a que el aceite es sometido en su análisis (Sala 2006b).

El experto en diagnóstico, interpreta los resultados de las pruebas del aceite y genera un sumario preciso a cerca de las condiciones de la máquina y de las condiciones del aceite. La interpretación de los resultados del análisis de aceite toma en cuenta principalmente: los límites de desgaste de los fabricantes del motor, características del aceite según el fabricante, literatura industrial típica, el estado actual de desgaste del equipo, los niveles de contaminación del aceite usado y las condiciones generales del mismo. En el sistema de diagnóstico propuesto en base al análisis de aceite, se presentan algunos fallos que tienen que ver con fallos exclusivos del aceite. Ejemplo de estos fallos son: aceite incorrecto (no adecuado para determinado motor), aceite viejo (un aceite que ha sido usado por un periodo de tiempo más largo del que debiera).

También existen fallos que están contenidos en el aceite (como su análisis lo confirma) pero que no son intrínsecos del aceite, sino que fueron ocasionados por condiciones particulares del motor. Ejemplo de estos fallos son: acumulación excesiva de partículas de cobre. Esto pudo ser ocasionado por desgaste en cojinetes de algún eje del motor. Si se encuentran exceso de partículas de cromo, esto puede significar fallo en los anillos del pistón. Como se puede notar en los anteriores juicios, se hace necesaria la opinión del experto en diagnóstico de motores.

9.3. Definición de la base de conocimiento

El conocimiento básico a ser incorporado en el sistema experto de diagnóstico¹ se conforma de los siguientes tópicos.

El sistema se centra en el diagnóstico de máquinas automotrices (camiones, autobuses y en general equipo para construcción de caminos), por lo que, los diferentes parámetros a ser medidos serán:

- *Propiedades del aceite*: viscosidad, *Total Base Number* (TBN) y detergencia
- *Contaminantes del aceite*: Compuestos insolubles, dilución de combustible, hollín, ingestión de polvo (silicón), agua y glicol.
- *Elementos metálicos*: hierro, cobre, plomo, cromo, aluminio, estaño, níquel, sodio y boro.

Se pueden desarrollar otras mediciones sobre la muestra de aceite, pero con estos parámetros básicos se logra un buen diagnóstico. En sistemas desarrollados para otro tipo de máquinas se pueden elegir otros parámetros tomando en cuenta las particularidades de ese tipo de máquinas.

La selección de variables listadas anteriormente se basa en el hecho de que el monitoreo de condición por el análisis de aceite en máquinas Diesel se subdivide de acuerdo a sus tres características principales: degradación de aceite, contaminación, y partículas resultantes del desgaste.

Degradación. Los principales parámetros que caracterizan la degradación del aceite son, viscosidad, *Total Base Number* (TBN) y detergencia. Viscosidad es sin lugar a dudas la más importante propiedad del aceite y provee información sobre la condición física del aceite. El sistema de diagnóstico considerado usa como medida la viscosidad cinemática a 100 °C en centistokes (cSt). Para evaluar la viscosidad en un aceite usado, se toma como referencia el valor de viscosidad original que presenta un aceite nuevo.

La comparación entre el volumen del TBN de un aceite nuevo y aquel de un aceite usado permite determinar si o no el aceite usado es capaz aún de neutralizar los residuos de ácido. Estos ácidos son producidos por combustión (sulfuro en gasolina), oxidación del aceite y por los aditivos para aceite. Un volumen de TBN muy bajo puede resultar debido a: excesiva oxidación del aceite, cuando el aceite ha estado en servicio por demasiado tiempo, o cuando el nivel del aceite se mantuvo muy bajo, o debido a sistema de enfriamiento defectuoso, que provocó sobre calentamiento; el uso de una de gasolina con alto

¹La base de conocimientos y de hechos para este sistema experto de diagnóstico fue proporcionada por un experto en este campo, quien es, Bernardo Tormos del grupo de Motores Térmicos CMT, de la Universidad Politécnica de Valencia, quien también aportó su valiosa opinión con respecto al desempeño del sistema de diagnóstico que aquí se propone.

contenido de sulfuro; el uso de un lubricante inapropiado; o debido a la contaminación del aceite por la ingestión de gasolina o agua.

La detergencia en el aceite permite mantener la máquina sin depósitos en las partes calientes y evitar la formación de lodo. El nivel de detergencia en un aceite usado es medido por un índice que varía entre 100 (detergencia ideal) y 0 (sin detergencia).

Contaminantes. Las causas de contaminación del aceite son muchas, y pueden clasificarse de acuerdo a la fuente que la produce. Existe contaminación que viene de fuera del sistema tal como: polvo (sílice) o líquidos (agua o mezclas con otros aceites). También hay impurezas que provienen de los procesos industriales o ambientes en los que la máquina trabaja, por ejemplo en manufactura se producen rebabas de soldadura, en el ensamblaje se involucra generalmente el polvo, silicones o polvo para pulir, a la vez que el mantenimiento puede introducir impurezas a través de trapos sucios o juntas deterioradas.

El lubricante en si mismo puede producir o contener contaminantes como son: partículas de desgaste, lodo (deterioración del aceite), hollín, ácidos (oxidación del aceite, sulfuro de la gasolina), gasolina, anticongelante, deterioración de los empaques y sellos. Un análisis minucioso del contenido en contaminantes puede indicar los problemas de una máquina y las causas que los originaron, de ahí la importancia de este tipo de medidas.

Desgaste. El análisis de partículas de desgaste indica el estado de desgaste de la máquina. Las predicciones de un desgaste anormal en máquinas Diesel proveen opciones importantes de otra manera indisponibles para los responsables en la toma de decisiones. Se pueden evitar daños secundarios identificando y removiendo a tiempo componentes dañados además de los subsecuentes ahorros en mantenimiento cuando fallos de consecuencias catastróficas son predichos con antelación.

El monitorizado de las condiciones de desgaste cuantitativo es un problema difícil y complejo. Las medidas son afectadas por diferentes factores que deben ser tomados en cuenta para ser compensados de manera que podamos tener una buena interpretación de las medidas. Se aplica una metrología apropiada para evaluar las condiciones de desgaste en máquinas de combustión interna. Estas técnicas metrológicas son incluidas en un modulo de preproceso de datos.

Esta metrología usa datos de las medidas de partículas de desgaste espectrométricos y datos de mantenimiento típicos para obtener parámetros más representativos de las condiciones de desgaste que toman en cuenta las condiciones de operación de la máquina que afectan la concentración de partículas de desgaste en las medidas.

Finalmente, la evaluación de esta proporción de desgaste (normal, medio o alto) es hecha bajo criterios estadísticos considerando las características individuales de cada máquina tales como: fabricante, metalurgia, etc.

9.4. Preproceso de medidas y escenario de diagnóstico

9.4.1. Preproceso de medidas

Cuando se adquiere información de una máquina en particular, algunas observaciones tienen el mismo significado para todas las máquinas a diagnosticar. Sin embargo, algunas otras necesitan el uso de datos históricos para generar desviaciones “normalizadas”, que toman en cuenta información estadística de una marca o modelo de máquina en particular, o de una unidad en particular con características especiales.

De esta forma, la concepción de la base de reglas, puede resultar más general (aplicada sin modificaciones a un gran número de casos) si los datos se escalan adecuadamente y se cambian de sitio antes de la inferencia.

En algunas medidas, el procedimiento envuelve normalizar la desviación a partir de la media, expresándola en unidades de varianza, y generando una cantidad adimensional. Los datos estadísticos son calculados a partir de una base de datos de análisis previos, que se clasificaron por marca, modelo (y también de registros históricos de la misma máquina).

Otras variables son transformadas a unidades de ingeniería, que proveen una mejor interpretación que las unidades originales en la que los datos se presentan (por ejemplo, la viscosidad del aceite se expresa como un porcentaje de un valor de referencia, tomado de las características de un aceite fresco, en lugar de las medidas en centiStokes).

También, con el propósito de considerar el comportamiento real de la máquina, se toman en cuenta el consumo de aceite y las adiciones de aceite nuevo, lo que conduce a obtener una concentración de elementos de desgaste compensada más representativa, del estado real de desgaste de la máquina.

9.4.2. Escenario de diagnóstico

La lista de los síntomas a considerar es la siguiente:

Síntoma	Abreviatura
viscosidad anormal	va
viscosidad alta	vh
viscosidad baja	vl
hollín anormal	soot
<i>tbn</i> anormal	tbn
periodo del aceite viejo	oldoil

Síntoma	Abreviatura
óxido o nitrito anormal	oxinit
silicón anormal	sil
desgaste 1	wear1
desgaste 2	wear2
<i>cu</i> anormal (para desgaste)	cu1
<i>cu</i> muy alto (para contaminación de agua)	cuvh
agua anormal	wat
<i>na</i> anormal	na
<i>bo</i> anormal	bo
<i>nabo</i> anormal (cociente de concentracion de sodio y boro)	nabo
dilución	dil

La lista de los desórdenes o fallos es:

Desorden	Abreviatura
aceite dañado	wo
problema de combustión (hollín)	cp1
problema de combustión 2 (tbn)	cp2
problemas de inyección	inj
aceite viejo	old
muestra contaminada	sc
ingestión de polvo	dust
contaminación de aceite con silicón	oilsil
contaminación de agua	water
desgaste (medido directamente)	wear

9.5. Construcción de las reglas de inferencia

La relación desorden síntoma en este sistema de diagnóstico se presenta en el cuadro 9.1. De esta relación obtenemos nuestras reglas de inferencia. En el desorden *old1* la presencia del *tbn* es incierta, sólo se presenta el 20% de las veces. Según la técnica de optimización propuesta en el capítulo 8, las reglas de inferencia propuestas en el cuadro 9.1 generan las ecuaciones lineales de coeficientes de relación y de superposición que se listan a continuación:

- *Conocimiento causal expresado por los coeficientes de relación $s_i^j = d_i$* , donde d_i corresponden a los diferentes desórdenes y s_i^j a las diferentes manifestaciones o síntomas (m_j) que se presentan en relación a los diferentes desórdenes existentes. En el sistema de diagnóstico propuesto resultan las siguientes ecuaciones de coeficientes de relación:

$$\begin{array}{lll}
s_{wo1}^{vh} = d_{wo1}, & s_{wo2}^{vl} = d_{wo2}, & s_{cp1}^{vh} = d_{cp1} \\
s_{cp1}^{soot} = d_{cp1}, & s_{cp2}^{tbn} = d_{cp2}, & s_{inj}^{vl} = d_{inj} \\
s_{inj}^{dil} = d_{inj}, & s_{old}^{vh} = d_{old}, & s_{old}^{soot} = d_{old} \\
s_{old}^{tbn} \leq d_{old}, & s_{old}^{oldoil} = d_{old}, & s_{sc}^{sil} = d_{sc} \\
s_{dust}^{sil} = d_{dust}, & s_{dust}^{wear1} = d_{dust}, & s_{oilsil1}^{sil} = d_{oilsil1} \\
s_{oilsil1}^{cu1} = d_{oilsil1}, & s_{oilsil2}^{wear2} = d_{oilsil2}, & s_{oilsil2}^{sil} = d_{oilsil2} \\
s_{wat1}^{na} = d_{wat1}, & s_{wat1}^{bo} = d_{wat1}, & s_{wat1}^{nabo} = d_{wat1} \\
s_{wat2}^{cu1} = d_{wat2}, & s_{wat2}^{cuvh} = d_{wat2}, & s_{wat2}^{wat} = d_{wat2} \\
s_{wear1}^{cu1} = d_{wear1}, & s_{wear2}^{wear1} = d_{wear2}, & s_{wear3}^{wear2} = d_{wear3}
\end{array}$$

Cuadro 9.1: Relación desórdenes\síntomas

Desorden	Síntomas\desórdenes presentes
wo1	vh
wo2	vl
wo	wo1, wo2
cp1	vh, soot
cp2	tbn
inj	vl, dil
old	vh, soot, 0.2tbn, oldoil
sc	sil
dust	sil, wear1
oilsil1	sil, cu1
oilsil2	sil, wear2
oilsil	oilsil1, oilsil2
wat1	na, bo, nabo
wat2	cu1, cuvh, wat
wat	wat1, wat2
wear1	cu1
wear2	wear1
wear3	wear2
wear	wear1, wear2, wear3

- *Ecuaciones de superposición*, agrupadas con respecto a cada nodo de manifestaciones o síntomas (m_{vh}, m_{vl}, \dots):

$$\begin{array}{l}
s_{wo1}^{vh} \leq m_{vh}, \quad s_{cp1}^{vh} \leq m_{vh}, \quad s_{old}^{vh} \leq m_{vh} \\
s_{wo1}^{vh} + s_{cp1}^{vh} + s_{old}^{vh} \geq m_{vh}
\end{array}$$

$$s_{wo2}^{vl} \leq m_{vl}, \quad s_{inj}^{vl} \leq m_{vl}, \quad s_{wo2}^{vl} + s_{inj}^{vl} \geq m_{vl}$$

$$\begin{array}{l}
s_{cp1}^{soot} \leq m_{soot}, \quad s_{old}^{soot} \leq m_{soot} \\
s_{cp1}^{soot} + s_{old}^{soot} \geq m_{soot}
\end{array}$$

$$s_{cp2}^{tbn} \leq m_{tbn}, \quad s_{old}^{tbn} \leq m_{tbn}$$

$$s_{cp2}^{tbn} + s_{old}^{tbn} \geq m_{tbn}$$

$$s_{old}^{oldoil} = m_{oldoil}$$

$$s_{sc}^{sil} \leq m_{sil}, \quad s_{dust}^{sil} \leq m_{sil}, \quad s_{oilsil1}^{sil} \leq m_{sil}, \quad s_{oilsil2}^{sil} \leq m_{sil},$$

$$s_{sc}^{sil} + s_{dust}^{sil} + s_{oilsil1}^{sil} + s_{oilsil2}^{sil} \geq m_{sil}$$

$$s_{dust}^{wear1} \leq m_{wear1}, \quad s_{wear2}^{wear1} \leq m_{wear1}, \quad s_{dust}^{wear1} + s_{wear2}^{wear1} \geq m_{wear1}$$

$$s_{oilsil2}^{wear2} \leq m_{wear2}, \quad s_{wear3}^{wear2} \leq m_{wear2}, \quad s_{oilsil2}^{wear2} + s_{wear3}^{wear2} \geq m_{wear2}$$

$$s_{oilsil1}^{cu1} \leq m_{cu1}, \quad s_{wat2}^{cu1} \leq m_{cu1}, \quad s_{wear1}^{cu1} \leq m_{cu1}$$

$$s_{oilsil1}^{cu1} + s_{wat2}^{cu1} + s_{wear1}^{cu1} \geq m_{cu1}$$

$$s_{wat2}^{cuvh} = m_{cuvh}$$

$$s_{wat2}^{wat} = m_{wat}$$

$$s_{wat1}^{na} = m_{na}$$

$$s_{wat1}^{bo} = m_{bo}$$

$$s_{wat1}^{nabo} = m_{nabo}$$

$$s_{inj}^{dil} = m_{dil}$$

- *Las ecuaciones de desórdenes compuestos*, son los desórdenes de un mismo tipo. En este sistema de diagnóstico tenemos restricciones para los desórdenes compuestos tales como:

$$d_{wo} \geq d_{wo1}, \quad d_{wo} \geq d_{wo2}$$

$$d_{oilsil} \geq d_{oilsil1}, \quad d_{oilsil} \geq d_{oilsil2}$$

$$d_{wat} \geq d_{wat1}, \quad d_{wat} \geq d_{wat2}$$

$$d_{wear} \geq d_{wear1}, \quad d_{wear} \geq d_{wear2}, \quad d_{wear} \geq d_{wear3}$$

En algunos casos, un desorden puede producir diferentes conjuntos de manifestaciones, dependiendo del contexto (por ejemplo, d_{wear}). En este caso, a cada uno de los conjuntos de síntomas posibles se le asigna un identificador de desorden auxiliar para describir una subdivisión de desorden (subdesórdenes como: d_{wear1} , d_{wear2} , \dots), y la severidad presentada al usuario es el valor máximo que presentan los subdesórdenes, porque según la restricción, el desorden siempre será mayor o igual que los subdesórdenes.

Una manera en la que podemos apreciar la relación causa-efecto en un sistema de diagnóstico como es el que aquí se estudia, es a través de un medio gráfico, por lo que en la figura 9.1 se presenta un diagrama de relación causa-efecto, que nos muestra el tipo de relaciones que existen entre los desórdenes y manifestaciones o síntomas. Esta relación también se expresa en el cuadro 9.1.

Las líneas continuas nos especifican una relación directa entre el desorden y el síntoma, mientras que las líneas discontinuas (punteadas) nos especifican una relación incierta. Los cuadrados representan los desórdenes, mientras que los pentágonos las manifestaciones. El signo *más* identifica a los desórdenes compuestos. Una interpretación de como la figura 9.1 describe las reglas sería:

- 1) *Aceite viejo* (desorden *old*) produce *viscosidad alta* (manifestación *vh*), *hollín* (manifestación *soot*), y una manifestación de un aceite que *no se ve joven* (*oldoil*) y algunas veces un *TBN* anormal.
- 2) *Desgaste* (*wear*) produce *cobre* anormal o *metal1* (*wear1*) anormal o *metal2* (*wear2*) anormal.
- ...

9.6. Diagnóstico de fallos

En este sistema experto para el diagnóstico por el análisis de aceite usado, se propone la utilización del enfoque posibilístico borroso que mira al sistema experto como un problema de optimización con restricciones lineales. Aunque se pudiesen usar otras herramientas de software en esta aplicación, como es MATLAB (cuenta con funciones de optimización), se ha elegido el uso del software llamado MATHEMATICA, por resultar más adecuado para un entorno de sistema experto. Explicaremos primero, como se obtiene nuestras observaciones σ_j que es la información primaria para el sistema de diagnóstico.

9.6.1. Adquisición de observaciones σ_j

Las observaciones se adquieren como se explicó anteriormente a través de un preproceso de los datos del análisis del aceite usado. Para esto los datos se colocan en un fichero tipo texto, donde cada reglón de los 29 que tiene corresponde a un dato (ver cuadro 9.2), que luego es leído por un comando de MATHEMATICA. En el cuadro 9.2 se encuentran las lecturas de 5 casos (5 muestras, cada una con sus correspondientes 29 datos). Los valores σ_j u observaciones se obtienen de valores numéricos de las medidas (29 datos) y a través de las definiciones de conjuntos borrosos establecidos por el experto para cada manifestación anormal. Las reglas que definen a los conjuntos borrosos necesarios para obtener los valores σ_j se encuentran en el programa que se muestra a continuación.

(*SetDirectory: find file to read*)

```

SetDirectory["C : \"];
ss = Import["medidas1.txt", "Table"];
Array[meas, 15];
meas = {ss[[27, 2]], ss[[27, 2]], ss[[9, 2]], ss[[24, 2]], 0., ss[[22, 2]], 0., 0., 0., 0.,
ss[[4, 2]], ss[[18, 2]], 0., 0., ss[[13, 2]]};

```

(*viscosity high and low : fuzzification, saturate to 1 *)

```

If[meas[[1]] > 1, meas[[1]] = 1];
If[meas[[1]] < -1, meas[[1]] = -1];
meas[[1]] = Max[0., meas[[1]]];

```

Cuadro 9.2: Casos de medidas de 5 muestras de aceite usado

medidas	caso 1	caso 2	caso 3	caso 4	caso 5
Muestra	101306034	101491737	101400298	101318200	101306034
TIEMPO	1	0.3	1.2	0.5	1
AG	0	0	0.01	0	0
AGUA	0.08	0.08	0.08	0.08	1
AL	0.06	0	0.15	1	0.06
BA	0	0	0	0	0
BO	119.6	104.9	53.3	34.1	119.6
CA	0	0	0	0	0
CAR	0.03	0.03	0.95	0.02	0.03
CR	0	0	0.225	1	0
CU	0	0	1.29	1.04	0
DET	0.03	0.03	0.525	0.02	0.03
DI	0.1	0.95	0.1	0.1	0.1
FE	0.08	0	0.81	1	0.08
IN	0.06	0.03	0.6	0.06	0.06
MG	0	0	0	0	0
MN	0	0	0.08	0.04	0
NA	0.01	0.01	0.02	1	0.01
NI	0	0	0.13	0.17	0
PB	0.07	0	0.4	0.475	0.07
PO	0	0	0	0	0
SIL	0	0.03	0.05	1	0
SN	0	0	0.15	0.14	0
TBN	0.13	0.14	0.19	0.09	0.13
TI	0	0	0.02	0.01	0
VA	0	0	0.03	0	0
VISC	-0.12	-0.95	1.29	0.05	-0.12
ZN	0	0	0	0	0
NABO	0.005	0.0048	0.028	29.8826	0.005

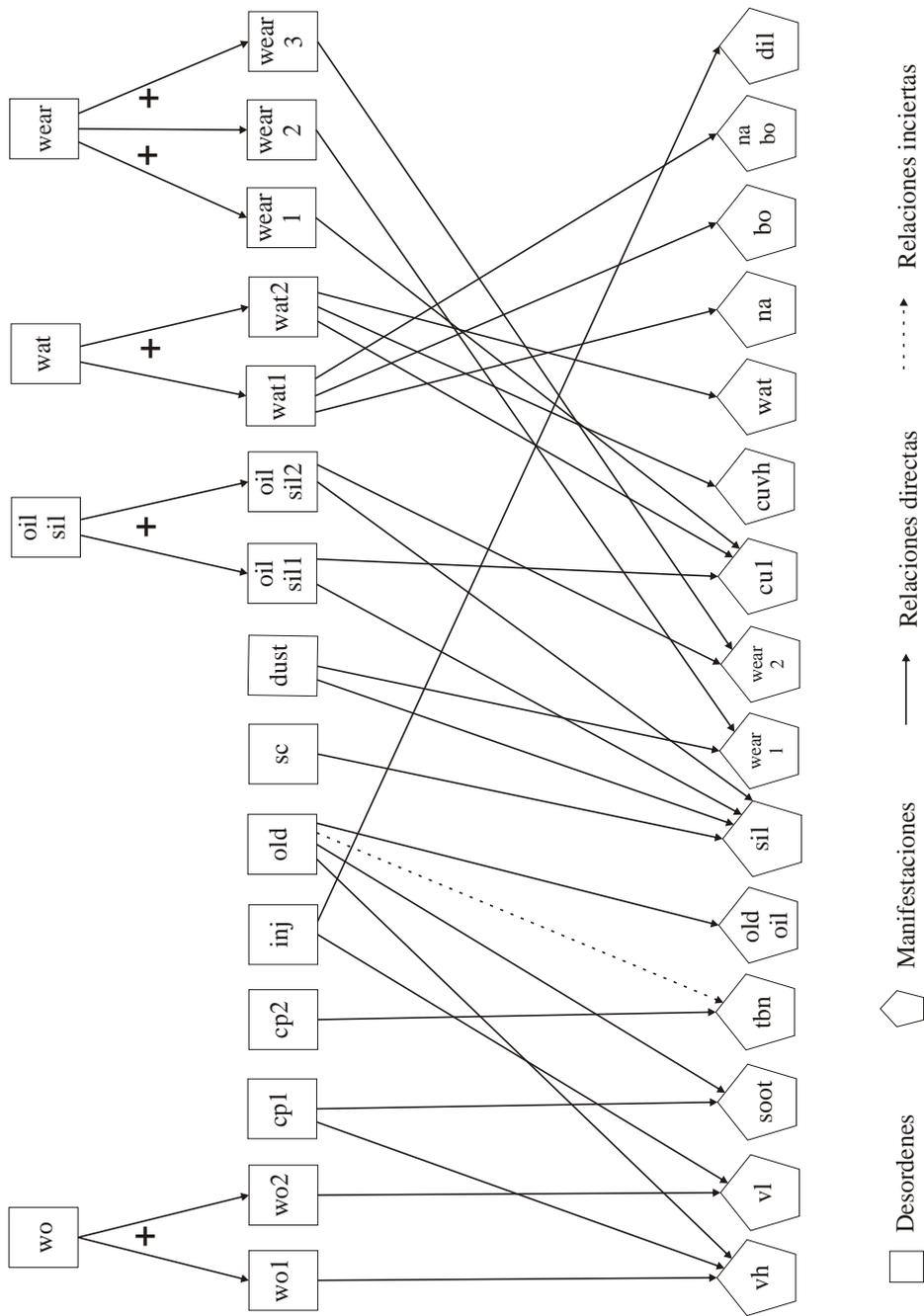


Figura 9.1: Diagrama de relación causa-efecto del sistema de diagnóstico por el análisis de aceites usados

$meas[[2]] = -Min[0., meas[[2]]];$

(*usage time*)

$tiempo = ss[[2, 2]];$

$If[tiempo < .9, meas[[5]] = 0,$

$If[tiempo < 1.2, meas[[5]] = (tiempo - 0.9)/(1.2 - 0.9), meas[[5]] = 1]];$

(*wear 1 and wear 2*)

$meas[[7]] = Max[Max[ss[[14, 2]], ss[[5, 2]], ss[[10, 2]]];$

$meas[[8]] = Max[ss[[20, 2]], ss[[23, 2]]];$

(*cooper issues *)

$cu = ss[[11, 2]];$

$If[cu < 0.9, cu1 = 0; cu2 = 0; ,$

$If[cu < 1.25, cu1 = (cu - 0.9)/(1.25 - .9); cu2 = 0; ,$

$If[cu < 2.25, cu1 = 1; cu2 = (cu - 1.25)/(2.25 - 1.25); , cu1 = 1;$

$cu2 = 1;]]];$

$meas[[9]] = cu1;$

$meas[[10]] = cu2;$

(* nabo *)

$nabo = ss[[29, 2]];$

$If[nabo < 0.5, meas[[14]] = 0,$

$If[nabo < 4, meas[[14]] = (nabo - .5)/(4 - .5), meas[[14]] = 1]];$

(*boro *)

$bo = ss[[7, 2]];$

$If[bo < 50, meas[[13]] = 0,$

$If[bo < 70, meas[[13]] = (bo - 50)/(70 - 50), meas[[13]] = 1]];$

(*it is the "normal" boro ... *)

$meas[[13]] = 1 - meas[[13]];$

(* this way Sigma observations were obtained*)

$SigmaVh = meas[[1]]; SigmaVl = meas[[2]]; SigmaSoot = meas[[3]];$

$SigmaTbn = meas[[4]]; SigmaOldoil = meas[[5]]; SigmaSil = meas[[6]];$

$SigmaWear1 = meas[[7]]; SigmaWear2 = meas[[8]];$

$SigmaCu1 = meas[[9]]; SigmaCuwh = meas[[10]]; SigmaWat = meas[[11]];$

$SigmaNa = meas[[12]]; SigmaBo = meas[[13]]; SigmaNabo = meas[[14]];$

$SigmaDil = meas[[15]];$

La salida del programa de adquisición viene a ser las variables de observación σ_i , con las cuales obtendremos las manifestaciones m_j , que luego se usarán en el comando de minimización para de esta manera tener el diagnóstico de fallos.

9.6.2. Función de optimización lineal

La función que principalmente se usa de MATHEMATICA es la función “*Minimize*”, esta es una función de programación lineal cuyos parámetros son los siguientes: *Minimize* [*index*, *constraints*], *{decision variables}*]. Se minimiza *index* (*J*) que es el índice de coste, esta minimización dependerá de *constraints* que son las restricciones lineales, donde *decision variables* (*x, y, ...*) son las variables solución del problema de optimización.

En nuestra aplicación los parámetros para el comando “*Minimize*” toman los siguientes nombres:

Minimize[(*J*), {*RelationCoefficients*, *SuperpositionEq*, *DisorderAdds*, *PositivityConstraints*, *Manifestations*}, (*DecisionVariables*)]

J es índice de coste o posibilidad de que los desórdenes se presenten. Las restricciones lineales están comprendidas en las variables, *RelationCoefficients*, *SuperpositionEq*, *DisorderAdds*, *PositivityConstraints*, y *Manifestations*. Las variables solución se encuentran en *DecisionVariables*. A continuación discutiremos cada uno de los grupos usados por la función de optimización.

- *RelationCoefficients*, esta variable contiene todos los coeficientes de relación $s_i^j = d_i$

RelationCoefficients = {*Swo1vh* == *dwo1*, *Swo2vl* == *dwo2*, *Scp1vh* == *dcp1*,
Scp1soot == *dcp1*, *Scp2tbn* == *dcp2*, *Sinjvl* == *dinj*,
Sinjdil == *dinj*, *Soldvh* == *dold*, *Soldsoot* == *dold*,
Soldtbn == *dold - v, ...*};

Aquí *Swo1vh* == *dwo1* corresponde a la relación $s_{wo1}^{vh} = d_{wo1}$, (MATHEMATICA usa el signo == para asignar un valor a los componentes de la lista) aunque las relaciones son 27 sólo se muestran algunas. La que nos interesa es la última de la lista *Soldtbn* == *dold - v* que corresponde a la relación incierta $s_{old}^{tbn} \leq d_{old}$ esto si *v* es multiplicada por cero en el índice de coste. Si el peso que multiplica a *v* no fuese cero, esto indicaría que “cuando *dold* se encuentre presente, en la mayoría de los casos el *TBN* sería anormal, pero existe la posibilidad de que fuese normal”. En esta relación *v* se dejo para ejemplificar el uso de las variables artificiales vistas en la sección 8.7 en el tema: suavizando la restricción de igualdades.

- *Índice de coste J*, en este índice incluimos todos los desórdenes y las variables auxiliares, porque son a estos parámetros que el experto en diagnóstico debe atribuir

una posibilidad *a priori* que sería el peso que multiplica a las variables.

$$\begin{aligned} Jvariables &= (0 * v + 5 * e1 + 5 * v1 + \dots + 5 * e15 + 5 * v15); \\ Jdisorders &= (1 * dwo + 1 * dcp1 + 1 * dcp2 + 1 * dinj + 1 * dold + 1 * dsc \\ &\quad + 1 * ddust + 1 * doilsil + 1 * dwat + 1 * dwear); \end{aligned}$$

$$J = (Jvariables + Jdisorders);$$

La variable J la subdividimos en $Jvariables$, que corresponde a las variables auxiliares y en $Jdisorders$, que corresponde a los desórdenes. Las variable v vistas en el párrafo anterior, están multiplicada por cero ($\pi = 1$). Desde el punto de vista del algoritmo de minimización v es obligada a tener un valor muy grande, de manera que la desigualdad en la relación $s_{old}^{tb} \leq d_{old}$ esta asegurada. Podríamos haber usado simplemente la desigualdad, sin embargo, con el uso de las variables de relajación, podemos graduar o borrosificar la posibilidad del coeficiente de relación.

Las variables artificiales (ϵ_j, ν_j) $e1, e2, \dots, e15$ y $v1, v2, \dots, v15$ se encuentran multiplicadas por un peso de 5, son usadas para suavizar el comportamiento que se tenga de las manifestaciones (m_j) con respecto a los desórdenes (d_i). Se ampliará este tema más adelante en la explicación de las medidas (*Manifestations*).

Los desórdenes básicos d_i se encuentran en $Jdisorders$, como se puede observar todos los desórdenes están en situación normal con posibilidad igual a 1 esto es debido a que el experto que ha provisto la base de conocimientos y de hechos, hasta el momento no ha dado una apreciación² de posibilidad para los desórdenes, pero sin embargo, se les puede asignar una posibilidad *a priori*, lo cual esta contemplado para versiones futuras de este sistema de diagnóstico. También es de notar que los subdesórdenes no se encuentran incluidos en el índice porque en realidad pertenecen a un solo desorden, el que los agrupa.

- *SuperpositionEq*, se incluyen todas las ecuaciones de superposición

$$\begin{aligned} SuperpositionEq &= \{Swo1vh \leq mvh, Soldvh \leq mvh, Scp1vh \leq mvh, Swo1vh + \\ &\quad Soldvh + Scp1vh \geq mvh, Swo2vl \leq mvl, \dots\}; \end{aligned}$$

Como se vió anteriormente el número de ecuaciones de superposición para este sistema son 35, sólo incluimos estas que se muestran arriba (manifestaciones del nodo vh) para ejemplificar la nomenclatura usada para definir las ecuaciones. En caso de ser necesario estas desigualdades se pueden suavizar por el uso de variables auxiliares. De existir ecuaciones de superposición inciertas es necesario el uso de las variables auxiliares e y v .

²Las posibilidades (π) ofrecen una alternativa abductiva cuando algunos desórdenes son consistentes. Sin embargo, la configuración actual de los extractos del laboratorio de análisis de aceites usados, ha provisto de un conjunto de mediciones bastante amplio y completo, de tal forma que en todas las muestras probadas existió sólo un mínimo global del índice. No obstante, la posibilidad de encontrar nuevas metodologías de análisis con menos información de muestras se encuentra bajo estudio.

- *DisorderAdds*, en esta variable se incluye los desórdenes compuestos

$$\begin{aligned} DisorderAdds = \{dwo \geq dwo1, dwo \geq dwo2, doilsil \geq doilsil1, doilsil \geq \\ doilsil2, dwat \geq dwat1, dwat \geq dwat2, dwear \geq dwear1, \\ dwear \geq dwear2, dwear \geq dwear3\}; \end{aligned}$$

- *PositivityConstraints*, en esta variable restringimos a toda variable usada a que tenga un valor no negativo, esto es, todas las variables deben ser mayor o igual a cero

$$\begin{aligned} PositivityConstraints = \{v \geq 0, e1 \geq 0, v1 \geq 0, \dots, dwo1 \geq 0, dwo2 \geq 0, dcp1 \\ \geq 0, \dots, Swo1vh \geq 0, Swo2vl \geq 0, Scp1vh \geq 0, \dots\}; \end{aligned}$$

Las únicas variables que no pueden ser restringidas son las manifestaciones m_j .

- *DecisionVariables*, aquí debemos incluir todas las variables usadas en el parámetro *constraints* de “Minimize”, en caso de faltar una o existir error de escritura, la función “Minimize” no se ejecutará

$$\begin{aligned} DecisionVariables = \{dwo1, dwo2, \dots, Swo1vh, Swo2vl, \dots, \\ moldoil, msil, \dots, v, e1, v1, \dots\}; \end{aligned}$$

- *Manifestations*, contiene todas las manifestaciones o síntomas necesarios para el diagnóstico, los cuales son nuestras entradas de datos

$$\begin{aligned} Manifestations = \{mvh == SigmaVh + e1 - v1, mvl == SigmaVl + e2 - v2, \\ \dots, mdil == SigmaDil + e15 - v15\}; \end{aligned}$$

Como podemos notar todas las manifestaciones m_j están relajadas por el uso de las variables auxiliares e_j y v_j . Al agregar estas variables logramos dos objetivos: primero, es que en caso de existir incertidumbre en las lecturas de los sensores o instrumentos o métodos de adquisición de las medidas σ_j , el experto en diagnóstico quien es el poseedor de este conocimiento *a priori*, lo puede especificar en el índice de costes J dándoles mayor posibilidad o menor según convenga a las variables e_i y v_j , de manera que pueda expresar el grado de incertidumbre existente en alguna de las medidas. Para nuestro caso, nuestras variables artificiales fueron pesadas con 5 en el índice de coste, lo cual significa que aunque las posibilidades son pocas ($\pi = 0,007$), todavía es posible encontrar algún error en las medidas.

El segundo objetivo es lograr que el algoritmo de optimización lineal converja en una solución aunque sus entradas (manifestaciones) sean casos nuevos, esto es, no previstos en la base de reglas de inferencia, logrando así, una estimación de un posible nuevo desorden, o en contribuir con el conocimiento del experto para determinar que algo no previsto esta sucediendo o que algo anda mal en determinado punto del proceso de diagnóstico.

Para lograr obtener las manifestaciones m_j (que pueden ser equivalentes a las observaciones σ_j según se vió en el capítulo 8), primero debemos preprocesar las medidas del aceite usado, y así obtener nuestras observaciones o medidas σ_j que se obtuvieron en la subsección anterior.

9.6.3. Discusión sobre algunos resultados

Una vez que se ha explicado sobre el programa de diagnóstico de fallos por el análisis de aceites, se discuten algunos de sus resultados. Contamos con 5 ficheros tipo texto donde colectamos las muestras de aceite usado de 5 casos, en el cuadro 9.2 se muestran estos datos.

Después del preproceso de las medidas de el cuadro 9.2 se generan las respectivas observaciones σ_j las cuales hemos anotado en el cuadro 9.3. Recordemos que con nuestras observaciones obtenemos nuestras manifestaciones ($m_j = \sigma_j + e_j - v_j$). En nuestra tabla 9.3 el valor de nuestra observación será igual al de nuestra manifestación, esto sucederá cuando $e_j = 0$ y $v_j = 0$. Cuando las variables auxiliares tengan algún valor diferente de cero nuestras manifestación no corresponderán a nuestras observación. El índice de coste nos dará indicios de lo anterior porque tendrá valores muy elevados alertándonos de que existen algunos errores en en las medidas (de sensores, de medición o de sistema) según se verá más adelante en el análisis de resultados (si el índice de coste es alto, primeramente podría indicar que algunos desórdenes se encuentran presentes además de los errores en las medidas).

Una vez obtenidas nuestras manifestaciones se alimentan al sistema de diagnóstico. Después de la optimización los resultados de los desórdenes encontrados, se especifican en el cuadro 9.4. Analizando los resultados de diagnóstico mostrados en el cuadro 9.4 podemos llegar a las siguientes conclusiones:

- *Caso 1*, en este básicamente se detectan desórdenes irrelevantes en combustión, inyección, y desgaste, los cuales en realidad son poco significativos por lo que es conveniente desestimarlos. Las lecturas de la edad del aceite (*oldoil*) en el cuadro 9.3 parecen ligeramente incorrectas (deducimos del *tbn*, *soot* y viscosidad *vl*, que el aceite parece ser más joven que el actual 0.333 *oldoil*): de cualquier manera, los resultados de los desórdenes no parecen ser originados por un aceite viejo. Como concluimos, la muestra de “*caso 1*” nos dice que la máquina que usa ese aceite esta en buenas condiciones y que ningún desorden se disparo en forma relevante.

En casos como este, que ligeramente se disparan algunas reglas, también podemos observar pequeñas variaciones en el índice de coste J , por la contribución del ajuste del las variables auxiliares, causado por el algoritmo de minimización, el cual infiere que en realidad no existe un desorden significativo y que por lo tanto, existen pequeños errores en las entradas de manifestaciones m_j .

El coste óptimo que corresponde a los valores de las variables de decisión asociados con los errores en las medidas son dados por (sólo se muestran las variables con un valor diferente a cero)

$$e1 = 0,03, v5 = 0,273, v11 = 0,08, v12 = 0,01$$

los cuales respectivamente corresponden a las manifestaciones o medidas: viscosidad alta, edad del aceite (*oldoil*), agua y sodio. Esto significa que unas medidas

Cuadro 9.3: Donde se muestran las observaciones σ_j

σ_j	caso 1	caso 1b	caso 2	caso 3	caso 4	caso 5
vh	0	0.03	0	1	0.05	0
vl	0.12	0.12	0.95	0	0	0.12
soot	0.03	0.03	0.03	0.95	0.02	0.03
tbn	0.13	0.09	0.14	0.19	0.09	0.13
oldoil	0.333	0.06	0	1.0	0	0.333
sil	0	0	0.03	0.05	1	0
wear1	0.08	0.08	0	0.81	1	0.08
wear2	0.07	0.07	0	0.4	0.475	0.07
cu1	0	0	0	1	0.4	0
cuvh	0	0	0	0.04	0	0
wat	0.08	0	0.08	0.08	0.08	1
na	0.01	0	0.01	0.02	1	0.01
bo	0	0	0	0.835	1	0
nabo	0	0	0	0	1	0
dil	0.1	0.1	0.95	0.1	0.1	0.1

“perfectamente consistentes” debe tener unas lecturas como las que se muestran en la columna de “caso 1b” de el cuadro 9.3. La observación más “inconsistente” es sobre la edad del aceite, como se comento previamente, esta debe ser considerada como un error en las medidas que atraiga la atención del experto (quien en este caso considero que la desviación es irrelevante y el diagnosis correcto).

- *Caso 2*, en este, es clara la evidencia de problemas de inyección (0.95), más otro aún incipiente de combustión que fue disparado por una ligera desviación del *tbn*. Paradójicamente este desorden se juzga como más posible que el “caso 1” (donde la situación es “normal”) esto es debido a las ligeras discrepancias en las lecturas sobre la edad del aceite (se le ha dado un peso alto a los errores de sensor).
- *Caso 3*, exhibe la detección de características de un viejo y deteriorado aceite más un desgaste muy significativo. El boro esta fuera de los valores normales de especificación, pero para este no existe alguna regla que haya sido diseñada para capturar tal comportamiento. El sistema de diagnóstico considera la observación del “*bo*” como un “error en las medidas” (El valor de J es alto) pero esto no cambia el diagnosis final: aceite viejo y desgaste (posiblemente debido a que no se ha hecho cambio de aceite a tiempo). Si el error en las medidas se remueve, el resultado de diagnóstico es el mismo, y el índice de coste es $J = 3,05$, lo cual denota una situación mucho más posible³.

³El análisis del errores en las medidas, usualmente debe ser llevado a cabo bajo la supervisión de un experto. La ventaja más importante de esta aportación, es que permite el análisis de los errores en las medidas, mientras los enfoques basados en lógica usualmente fallan “catastróficamente” (el AND de 45 síntomas activos y 1 inactivo es cero: si el último fue un error en las medidas, el diagnóstico ha sido arruinado).

Cuadro 9.4: Resultado de desórdenes encontrados por el sistema

Desorden	caso 1	caso 2	caso 3	caso 4	caso 5
wo1	0.02	0	0	0.03	0.02
wo2	0.02	0	0	0	0.02
wo	0.02	0	0	0.03	0.02
cp1	0	0	0.05	0.02	0
cp2	0.1	0.14	0	0.09	0.1
inj	0.1	0.95	0	0	0.1
old	0.03	0	0.95	0	0.03
sc	0	0	0	0	0
dust	0	0.03	0	0.68	0
oilsil1	0	0	0.025	0.16	0
oilsil2	0	0	0.025	0.16	0
oilsil	0	0	0.025	0.16	0
wat1	0	0	0.02	1	0
wat2	0	0	0.04	0.08	0
wat	0	0	0.04	1	0
wear1	0	0	0.935	0.32	0
wear2	0.08	0	0.81	0.32	0.08
wear3	0.07	0	0.4	0.32	0.07
wear	0.08	0	0.935	0.32	0.08
COSTE	2.44	1.72	7.13	3.20	7.05
π	0.087	0.18	0.0008	0.041	0.0008

- *Caso 4*, estas lecturas indican que las manifestaciones observadas son compatibles con “ingestión de polvo” (*dust*) más “contaminación de agua 1” *wat1* (sodio y boro anormal) más “contaminación de aceite por silicón” (*oilsil* medio), más “desgaste” (*wear* medio). La posibilidad de los resultados fueron de algún modo bajos, sin embargo suficientes para atraer la atención del experto.

Básicamente, el experto concluyó que el diagnóstico es correcto, pero ahondando en los detalles de la base de datos, se encontró que esta muestra fue obtenida de una máquina que opera en la maquinaria de una mina de postacio, así que esta fue expuesta a altas cantidades de partículas de polvo suspendidas en el aire con composiciones anormales. Las duras condiciones en el medio ambiente de trabajo de la máquina, realmente causan el incremento de desgaste y también la alta concentración de químicos en el aceite, de modo que la máquina requiere un mantenimiento más frecuente.

- *Caso 5*, este análisis es hecho con los mismos datos del caso de caso 1, sólo que hemos modificado la observación del agua (*wat*) para ejemplificar como el índice de coste nos puede ayudar a determinar cuando algo anda mal en las medidas u observaciones.

Se ha cambiado deliberadamente AGUA=1 en el cuadro 9.2. Esta manifestación no esta especificada en alguna regla como causante de desorden por si misma, para que se presente el desorden “contaminación de agua” (d_{wat2}) además del síntoma “agua anorma” (m_{wat}) deben estar presentes los síntomas “cobre anormal” (m_{cu1}) y “cobre muy alto” (en la contaminación del agua m_{cuwh}), por lo que cuando se tiene tan sólo la lectura de m_{wat} , el sistema considera que existe un error en la medida del agua.

En este caso, el índice de coste tiene el valor de $J = 7,05$, donde $v_{11} = 1$ que corresponde a m_{wat} , por lo que esta manifestación contribuye con 5 en el índice J , evidenciando de esta manera el error en la medida del agua.

9.7. Conocimiento incierto y posibilidad condicional

9.7.1. Conocimiento incierto

En la escritura de las reglas de inferencia se hace necesario borrosificar algunas relaciones, para darle el sentido más acercado a la realidad que sea posible, para de esta manera, dar cumplimiento a las propiedades lingüísticas de nuestra base de conocimientos. En el enfoque que aquí se presenta, se relajan algunas restricciones por el uso de variables auxiliares, como es el caso de nuestra regla de inferencia de el cuadro 9.1 donde en el desorden “old”, la presencia del síntoma “tbn” es incierta y sólo se presenta el 20 % de las veces. En las ecuaciones de coeficientes de relación, esta incertidumbre fue expresada como $S_{old}^{tbn} = d_{old} - v$.

Este conocimiento también puede ser expresado por medio de una interpretación lingüística borrosa y decir: La manifestación del *tbn* es usual cuando el desorden *dold* se encuentra presente, pero algunas veces el *tbn* aparece sólo con una gradualidad del 20 %.

Según se vió en la sección 8.5, este conocimiento podría ser interpretado por medio de la siguientes desigualdades: $0,2d_{old} \leq S_{old}^{tbn} \leq d_{old}$. En nuestro programa de diagnóstico la variable *RelationCoefficients* toma la siguiente forma:

$$\begin{aligned} \text{RelationCoefficients} = \{ & Swo1vh == dwo1, Swo2vl == dwo2, Scp1vh == dcp1, \\ & Scp1soot == dcp1, Scp2tbn == dcp2, Sinjvl == dinj, \\ & Sinjdil == dinj, Soldvh == dold, Soldsoot == dold, \\ & 0,2 * dold \leq Soldtbn, Soldtbn \leq dold, \dots \}; \end{aligned}$$

Los resultados observados en los desórdenes con la anterior modificación para todas las medidas σ_i , son exactamente los mismos que los que se obtuvieron en el cuadro 9.4. La ventaja de aplicar el conocimiento borroso incierto en la interpretación lingüística de las manifestaciones, es que no tenemos que ajustar las posibilidades de ocurrencia de e y v en el índice de coste J .

9.7.2. Posibilidad condicional

En nuestro sistema de diagnóstico la base de reglas (BC), esta formada por el conjunto de ecuaciones de coeficientes de relación, ecuaciones de superposición y ecuaciones de restricciones lineales. Nuestra base de hechos (BH) esta formada por las observaciones σ_i . Para obtener la posibilidad condicional, necesitamos saber cuál será el valor de $\pi(BC, BH)$.

Para el “caso 1”, con las observaciones σ_i que se tienen, resultó un índice de coste $J = 2,44$ por lo tanto para este caso $\pi(BC, BH : caso 1) = 0,087$. La posibilidad condicional, nos sirve para poder determinar dadas las condiciones actuales (BC, BH) cuál podría ser el comportamiento de un determinado desorden, si variásemos su posibilidad de ocurrencia en el intervalo $[0,1]$. De acuerdo a la base de hechos y la base de conocimiento actual, podemos determinar cuales son los desórdenes que tienen más posibilidades de tener un comportamiento diferente al que observan en estos momentos, lo cual es evidencia de incertidumbre en el diagnóstico.

Por ejemplo en el “caso 1”, los resultados del cuadro 9.4 muestran a “dinj” y “dwear2” tener un desarrollo aún incipiente, todos los demás desórdenes se encuentran ausentes. Para obtener la posibilidad condicional de uno o varios desórdenes, con respecto a nuestra base de conocimiento y nuestra base de hechos actuales, podemos hacer uso del teorema de Bayes. De esta forma, variando el comportamiento del o de los desórdenes entre los valores de $[0, 1]$, obtendremos la posibilidad condicional deseada.

Para obtener la posibilidad condicional de “dinj” usamos la siguiente expresión:

$$\pi(\{dinj\}|\{BC, BH\} : caso 1) = \frac{\pi(\{dinj\} \cup \{BC, BH\})}{\pi(\{BC, BH\} : caso 1)}$$

A continuación se presenta un análisis de la posibilidad condicional de los desórdenes de cada caso de medidas. En las gráficas de la figura 9.2 donde las abscisas son la severidad del desorden y las ordenadas la posibilidad, se analiza el comportamiento de la posibilidad condicional de los desórdenes del “caso 1”.

- *Caso 1*, podemos observar en las gráficas de “dinj”, “dwear2” y “dwear3” de la figura 9.2, (desórdenes incipientes del caso actual según tabla 9.4), que su comportamiento es similar con una posibilidad moderada, en cambio la gráfica para el resto de los desórdenes, (desórdenes actualmente ausentes), como sus posibilidades de presentarse dadas las condiciones actuales es mucho menor.
- *Caso 2*, este se presenta en la figura 9.3, donde podemos observar para “dinj” (desorden presente) una posibilidad alta y para “dcp2” (desorden incipiente) una posibilidad moderada, todos los desórdenes restantes según las condiciones actuales, observan una posibilidad baja de presentarse, excepto “dwo2”. Si observamos,

para este, a una severidad del 100% tendremos una posibilidad casi del 40% lo cual aparentemente es mucha, si consideramos que es un desorden que actualmente se encuentra ausente.

Si analizamos en el diagrama de la figura 9.1 el síntoma o manifestación necesaria para que “dwo2” se presente es viscosidad baja “mvl”, la cuál también es necesaria para que “dinj” se presente. Por lo anterior podemos concluir que es coherente que exista una alta posibilidad condicional de que ocurra “dwo” según las condiciones actuales.

Tal vez la pregunta en este análisis es: ¿por qué si “mvl” dispara el desorden “dinj” no dispara el desorden “dwo2”, si la manifestación que lo causa se encuentra presente?. La razón es, que el algoritmo de optimización lineal dadas las condiciones actuales, determina que sólo ocurrirá “dinj” y no ambos. Recordemos que en el índice de coste la posibilidad que multiplica a cada desorden es la misma ($\pi = 0,37$). Si quisiésemos que para este caso ambos desórdenes se disparasen, entonces tendríamos que aumentar la posibilidad que multiplica a “dwo2” en el índice de costes J .

- *Caso 3*, las gráficas de posibilidad condicional de este, se muestran en la figura 9.4. Aquí las posibilidades altas son para los desórdenes actualmente presentes “dold”, “dwear1”, “dwear2” y “dwear3”. Sin embargo existen dos desórdenes que también muestran una “*posibilidad condicional*” de ocurrencia alta, los cuales son: “dwol” y “dcp1”, desórdenes que aunque no se encuentran presentes pueden ser causados por “mvh”, manifestación que también causa el desorden “dold”.
- *Caso 4*, finalmente, las gráficas de “*posibilidad condicional*” para este último caso de análisis, se muestran en la figura 9.5. Como recordaremos esta es una muestra que proviene de una máquina que trabaja en una mina de potasio y que se determino en el análisis de resultados que el aceite usado por ella se encuentra contaminado con polvo y agua.

Se observan en las gráficas una posibilidad alta para los desórdenes que se diagnosticaron como presentes “ddust”, “doilsil1”, “dwat1”, “dwear1”, “dwear2” y “dwear3”. Sin embargo se observa una alta posibilidad de ocurrencia de un desorden “no presente” actualmente, este es “dsc=0” (muestra contaminada) que es el diagnóstico óptimo según la estimación *a priori* actual en el índice de coste J (la severidad más posible en “dsc” es el valor considerado en el cuadro 9.4). En la figura 9.5 la “*posibilidad condicional*” de “dsc=1” permanece alta (0.44): lo cual significa que existe incertidumbre en el diagnóstico producido. El “*caso 5*” no se analiza, debido a que los resultados de sus desórdenes son los mismos que los del “*caso 1*”.

Conclusiones del capítulo

Convertir la base de conocimiento lingüístico de un sistema experto a un sistema de optimización por ecuaciones lineales, resulta una técnica mucho más adecuada para un ordenador (por la disponibilidad de rápidos software de optimización comerciales) que interpretar las reglas de inferencia lingüísticas. Otra ventaja de esta aportación, es que el experto en diagnóstico puede hacer uso de su conocimiento *a priori* y decidir, que desórdenes él considera menos (o más) factibles de ocurrencia, asignando una posibilidad adecuada a cada uno de los desórdenes en el índice de costes J . También, por el uso de las variables auxiliares en nuestras entradas de información al sistema, podemos determinar cuando una medida (error de sensor) es incorrecta o cuando algún caso no previsto en nuestras reglas se presenta. La información que provee el sistema nos puede llevar a la modificación de información posibilística de las reglas (errores en las mediciones “demasiado” frecuentes, puede significar que algunas de las manifestaciones son más “inciertas” de lo que originalmente se pensaba).

Además, el análisis de posibilidad condicional nos advierte de inconsistencias en el diagnóstico, esta incertidumbre puede atraer la atención del experto quien puede refinar la base de conocimiento y de hechos o hacer ajustes en la posibilidad *a priori* del índice de coste J para mejorar el diagnóstico. Por lo anterior, se debe aclarar que la herramienta aquí propuesta se considera sólo apta como ayuda en la toma de decisiones, debido a que: 1) las reglas detectan errores en las medidas (en inglés, *outliers*) y por lo tanto requieren la atención del experto, 2) el experto posee información adicional a cerca de las condiciones de trabajo de la máquina (mina de potasio, etc.) que lo ayudan a tener una mejor idea de la situación.

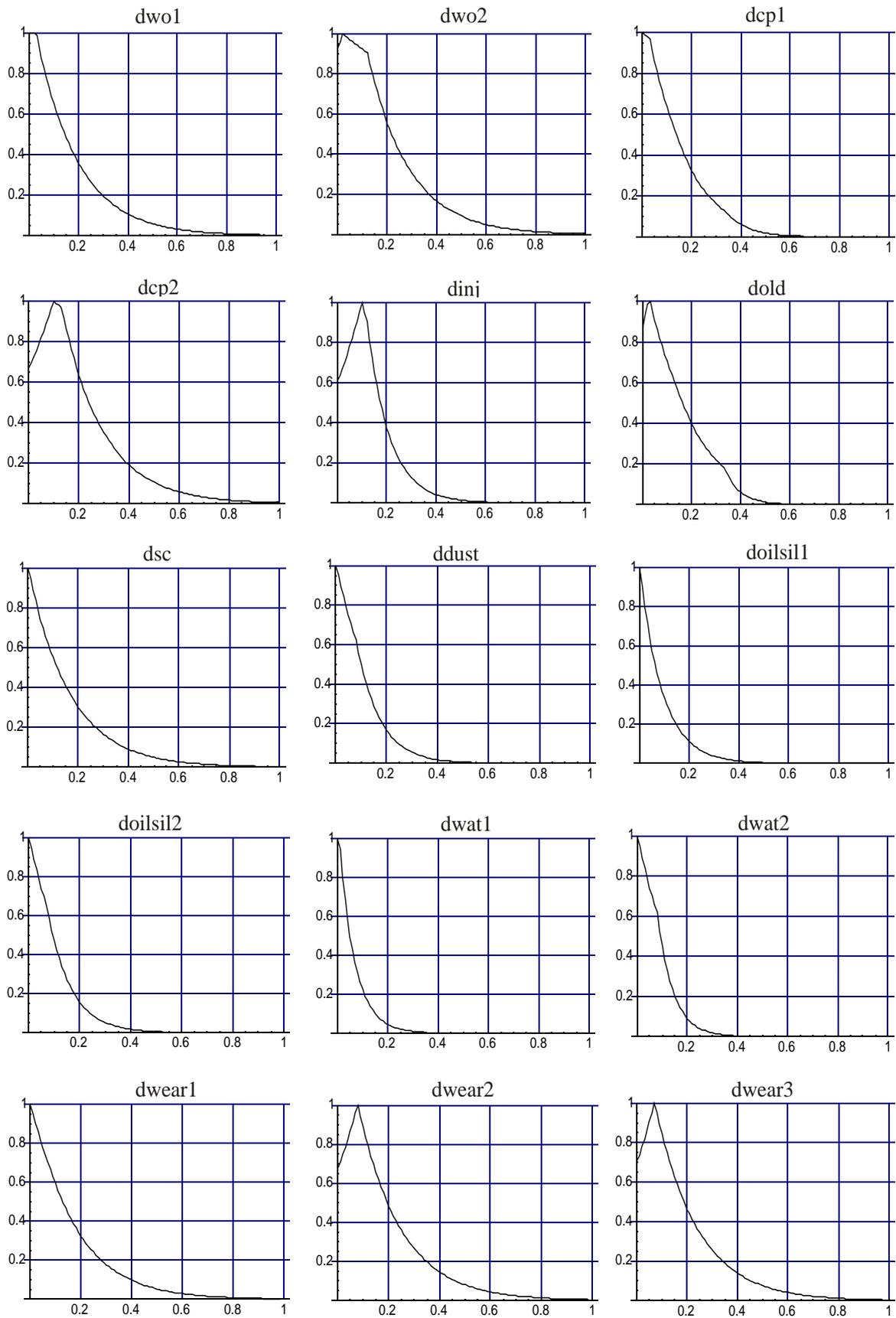


Figura 9.2: Gráficas de posibilidad condicional de los desórdenes del caso *caso 1*. Abscisas: severidad; ordenadas: posibilidad

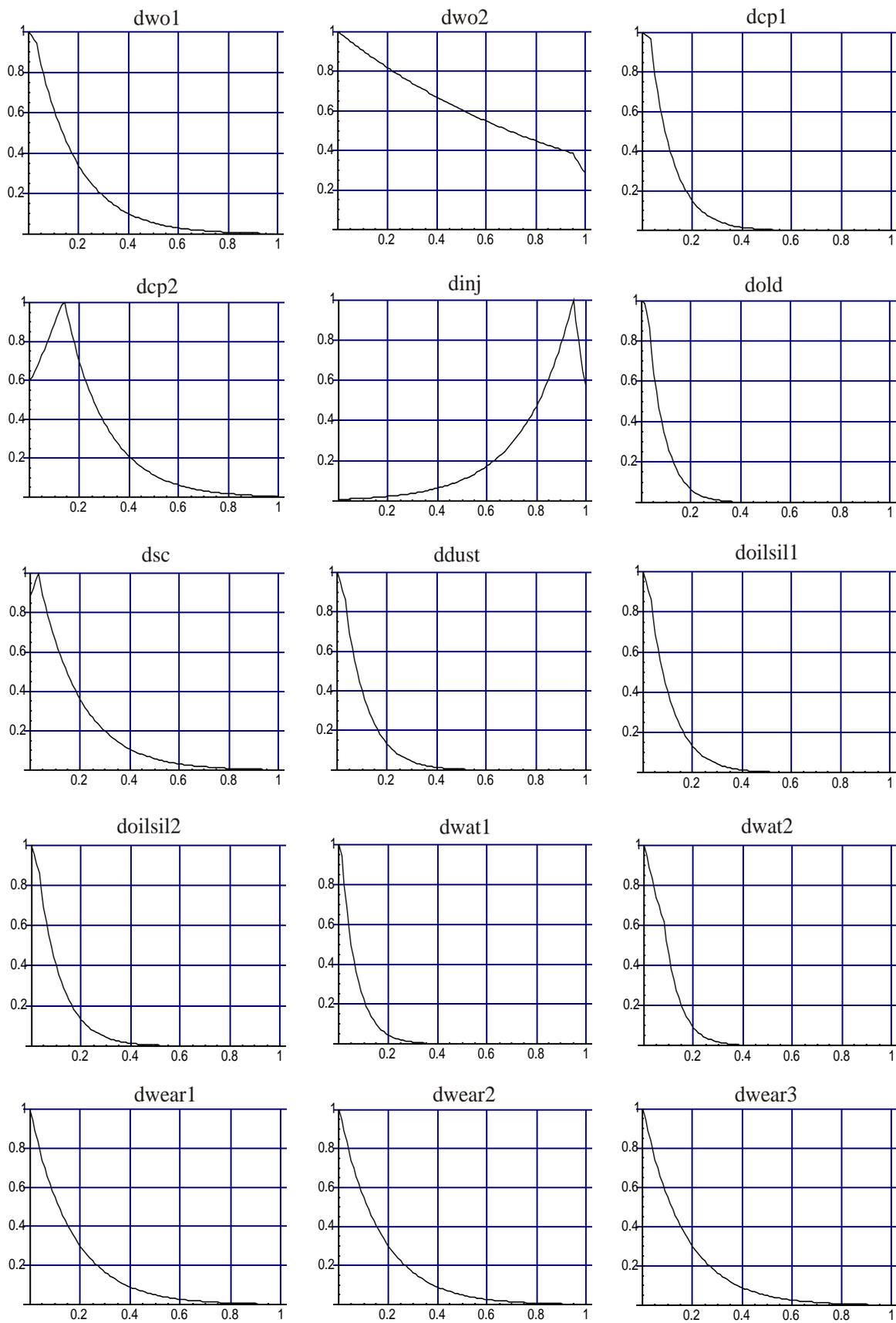


Figura 9.3: Gráficas de posibilidad condicional de los desórdenes del caso *medidas2*.
 Abscisas: severidad; ordenadas: posibilidad

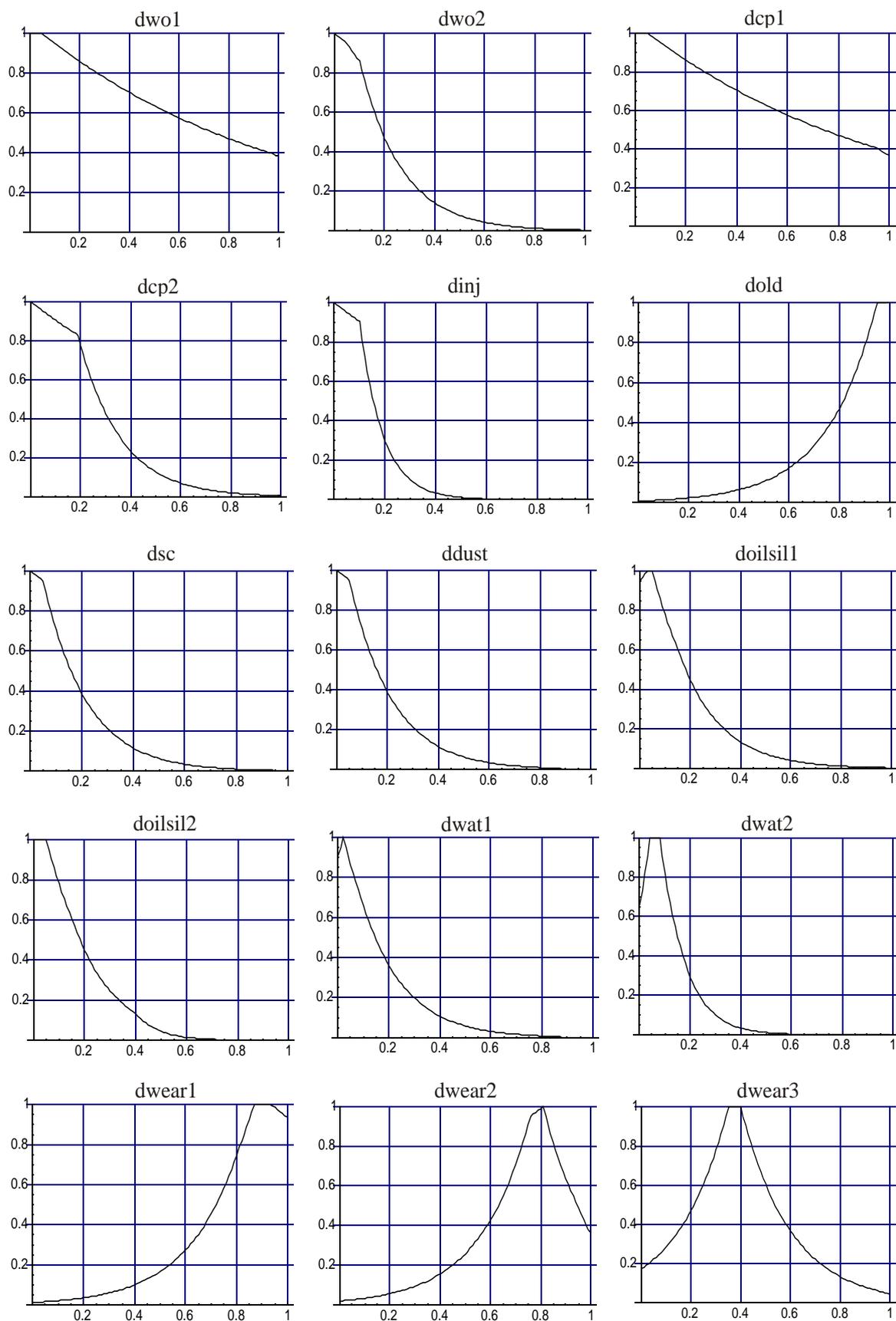


Figura 9.4: Gráficas de posibilidad condicional de los desórdenes del caso *medidas3*. Abscisas: severidad; ordenadas: posibilidad

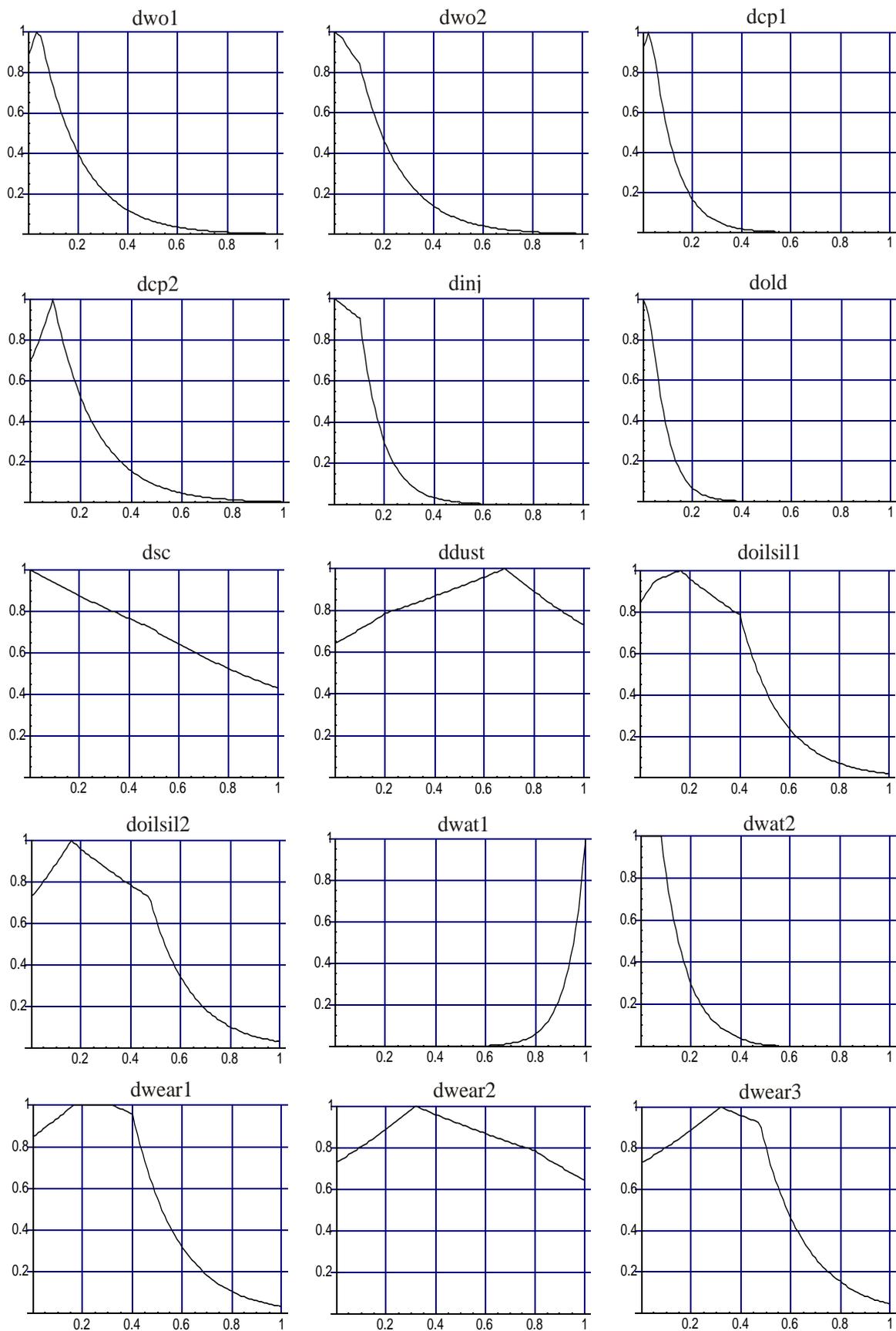


Figura 9.5: Gráficas de posibilidad condicional de los desórdenes del caso *medidas4*. Abscisas: severidad; ordenadas: posibilidad

Conclusiones

Capítulo 10

Conclusiones

10.1. Conclusiones del trabajo

En este trabajo se hizo un estudio del estado del arte en el diagnóstico de fallos. Estudiamos desde las generalidades del diagnóstico de fallos, la lógica borrosa los sistemas expertos hasta las técnicas de inteligencia Artificial usadas en el diagnóstico. En estas últimas se expusieron los métodos relacionales en los que se basa la aportación principal de esta tesis.

También se presentó lo referente a las aportaciones que en la etapa inicial de investigación se desarrolló, presentando estos avances en diferentes congresos. Hicimos uso de los sistemas expertos intervalares borrosos, los sistemas expertos que combinan lógica borrosa con probabilidad y las redes Bayesianas. De estos avances concluimos que aunque los resultados resolvían en parte la problemática de la incertidumbre presente en los sistemas de diagnóstico, tienen como desventaja lo complejo de su implementación. Además la presentación de resultados es compleja, tanto que no aportan una diagnosis final, si no que requieren aún de más investigación. Para ello es necesario auxiliarse en técnicas de tomas de decisiones, lo que añade más complejidad a esas técnicas sin garantía de optimizar el diagnóstico.

Se hizo un estudio profundo de la fundamentación teórica de la aportación principal de esta tesis, que es: un sistema experto posibilístico borroso visto como un problema de optimización lineal. Como convertir nuestros enunciados lingüísticos que componen nuestra base de reglas en un sistema experto borroso, a un conjunto de ecuaciones lineales a través de técnicas relacionales. Como aplicar estas ecuaciones a un algoritmo de optimización lineal. Como interpretar el índice de coste en forma posibilística. Como analizar los resultados de optimización. Además, el análisis de error en las medidas y las aportaciones en el diagnóstico, que ofrece la posibilidad condicional.

Con este profundo análisis, se demostró lo útil que puede resultar convertir un sistema experto borroso, a un problema de optimización lineal. Cómo, el hecho de convertir

un sistema borroso a un conjunto de ecuaciones lineales, no implica que perdamos la representatividad de la borrosidad.

Por último se hizo una aplicación de la técnica de sistema experto posibilístico borroso a un sistema de diagnóstico de motores de combustión interna por el análisis de aceites usados. De este último resultó un artículo que ha sido aceptado para su publicación y exposición en el *IEEE International Conference on Fuzzy Systems*, FUZZ-IEEE 2007.

Podemos concluir en esta aplicación, que resultó mucho más fácil para el experto en diagnóstico, entender los resultados. También, apporto información extra a cerca de los errores en las medidas. Con esto, el experto en diagnóstico, puede ir refinando su base de conocimiento al especificar, que medidas en realidad serán afectadas por una mayor incertidumbre de la que se pensaba.

10.2. Trabajos futuros

Actualmente, se encuentra en desarrollo la automatización del sistema de diagnóstico para motores de combustión interna por el análisis de aceite usado. Esto se logrará mediante un software que facilite la interfaz de usuario. Este software de alto nivel, leerá el resultado de la función de minimización de MATHEMATICA y presentará los resultados al usuario en forma más clara y fácil de interpretar. Lo anterior se hace debido a que los resultados ofrecidos por la función de minimización que actualmente se ha utilizado en el capítulo 9 no son tan claros para el usuario final. Con esto, se busca contribuir con el departamento de Motores Térmicos CMT de la UPV.

Otro de los trabajos futuros de investigación que queda abierto, es la aplicación del diagnóstico posibilístico borroso propuesto en esta aportación, a procesos industriales dinámicos y no tan solo estáticos, como el estudiado en esta tesis. Resultará interesante probar que respuesta se obtiene de los algoritmos de optimización, usados en la actualidad por los software comerciales, aplicados ahora a sistemas dinámicos.

También queda abierta la investigación, para aplicar el diagnóstico posibilístico borroso estudiado, a sistemas complejos que contengan capas de nodos ocultas, en estructuras similares a las usadas en redes Bayesianas y en redes neuronales. Con nodo de capa oculta, nos referimos a una situación creada por un desorden. Esta situación será una etapa intermedia que genere después un síntoma.

desorden \rightarrow situación intermedia \rightarrow síntoma

Además, se propone incorporar aprendizaje supervisado y no supervisado, de forma que la estructura del sistema de diagnóstico pueda ir creciendo automáticamente.

Por último, Se propone buscar adecuar esta técnica, a otros procesos industriales además del análisis de aceites estudiado a fondo en esta aportación.

Bibliografía del autor en el tema de la tesis

Sistema experto Borroso para el diagnóstico

presentado en: *Electro 2002, Automática e Informática Industrial, Instituto Tecnológico de Chihuahua, Méx.*

escrito por: *Ramírez J., Sala A. y Morant F.*

volumen: 1, páginas: 533-537.

Diagnóstico de sistemas mediante combinación de lógica borrosa y probabilidad

presentado en: *Electro 2004, Instituto Tecnológico de Chihuahua, Méx.*

escrito por: *Ramírez J. y Sala A.*

volumen: 1, páginas: 341-346.

Learning Bayesian Networks for Systems Diagnosis

presentado en: *Proceedings of the Electronics, Robotics and Automotive Mechanics Conference (CERMA 2006), IEEE Computer Society.*

escrito por: *Ramírez J. y Sala A.*

volumen: 2, páginas: 123-128.

An Optimization Approach to Fuzzy Diagnosis: Oil Analysis Application

presentación futura en: *International Conference on Fuzzy Systems, FUZZ-IEEE 2007.*

escrito por: *Sala A., Ramírez J., Tormos B. y Yago M.*

Aceptado para su exposición y publicación.

Bibliografía

- Abbasbandy, S. 2004. “The nearest trapezoidal fuzzy number to a fuzzy quantity.” *Applied Mathematics and Computation, Elsevier* 156 (2): 381–386 (Sep).
- Abbasbandy, S. y Amirfakhrian M. 2006. “The nearest approximation of a fuzzy quantity in parametric form.” *Applied Mathematics and Computation, Elsevier* 172 (1): 624–632 (Jan.).
- Albertos, P. 1992. “Fuzzy controllers.” Edited by L. Krijgsman A. y Vingerhoeds R. Boullart, *Application of artificial intelligence in process control*. Pergamon Press, 343–367.
- Amirkhanzadeh, R. et al. 2005. “A mixed-signal current-mode fuzzy logic controller.” *AEU - International Journal of Electronics and Communications, Elsevier* 59 (3): 177–184 (June).
- Angeli, C. 1999. “Online expert system for fault diagnosis in hydraulic systems.” *Expert Systems* 16:115–120.
- Astrom, K. J. A. y Arzen K. E. 1986. “Expert control.” *Automatica* 22 (3): 277–286.
- Atanasov, K. 1986. “Intuitionistic fuzzy sets.” *Fuzzy Sets and Systems* 20:87–96.
- Avazpour, A. y Moradi M. 2006. “The direct correlation functions of hard Gaussian overlap and hard ellipsoidal fluids.” *Physica B: Condensed Matter, Elsevier* 392 (1-2): 242–250 (Nov.).
- Ayers, D. Chellappa R. 2000. “Scenario recognition from video using a hierarchy of dynamic belief networks.” *Pattern Recognition, 2000. Proceedings. 15th International Conference on* 1 (Sept.): 835 – 838.
- Ayoubi, M. 1996. “Fuzzy systems design based on hybrid neural structure and application to the fault diagnosis of technical processes.” *Control Engineering Practice* 4 (1): 35–42.
- Baldwin, J.F. y Di Tomaso E. 2003. “Inference and learning in fuzzy bayesian networks.” *Fuzzy Systems, 2003. FUZZ '03. The 12th IEEE International Conference on* 1 (May): 630 – 635.
- Basir, O. y Yuan X. 2005. “Engine fault diagnosis based on multi-sensor information fusion using Dempster-Shafer evidence theory.” *Information Fusion, Elsevier*, Oct.
- Belacel, N. et al. 2007. “Learning multicriteria fuzzy classification method PROAFTN from data.” *Computers and Operations Research* 34 (7): 1885–1898 (July).

- Bilgic, T. y Turksen I. 1997. "Measurement of membership functions: Theoretical and empirical work." Edited by Dubois D. y Prade H., *Handbook of fuzzy sets and systems*, Volume I. University of Toronto (Canada).
- Black, M. 1937. "Vagueness: an exercise in logical analysis." *Philosophy of Science* 4:427–455.
- Blanke, M. Kinnaert M. Lunze J. y Staroswiecki M. (editores). 2003. *Diagnosis and Fault-Tolerant Control*. Springer, London.
- Bohs, L. et al. 1994. "Volumetric Flow Measured in a Laminar jet Phantom Using a Two-Dimensional Speckle Tracking System." *IEEE, Ultrasonics Symposium*.
- Bolgelt, Christian y Kruse Rudolf. 2000. *Abductive inference wiht probabilistic networks*. Kluwer, MA, USA.
- Bondia, J. Sala A. Pico J. Sainz M. A. 2006. "Controller Design Under Pole-Placement Specification: An Interval Arithmetic Approach." *IEEE Tran. On Fuzzy Systems* 14, no. 6 (Dec.).
- Bonte, M.H.A. et al. 2007. "Determining the von Mises stress power spectral density for frequency domain fatigue analysis including out-of-phase stress components." *Journal of Sound and Vibration, Elsevier* 302 (1-2): 379–386 (April).
- Boston, R. J. 2000. "A Signal Detection System Based on Dempster-Shafer Theory and Comparison to Fuzzy Detection." *IEEE Transactions on Systems, Man, and Cybernetics-part C: Applications and Reviews* 30, no. 1 (February).
- Bo-Suk, Y. y Kwang J. K. 2006. "Application of Dempster-Shafer theory in fault diagnosis of induction motors using vibration and current signals." *Mechanical Systems and Signal Processing, Elsevier* 20 (2): 403–420 (Feb.).
- Brucq, Denis de, et al. 2002. "Identical Foundation of probability theory and fuzzy sets theory." *IEEE SIF*.
- Buchanan, B. y Shortliffe E. 1984. *Rule-based expert systems, Readings (MA)*. USA: Addison-Wesley.
- Cairó, L. y Llibre J. 2007. "Polynomial first integrals for weight-homogeneous planar polynomial differential systems of weight degree 3." *Journal of Mathematical Analysis and Applications* 331 (2): 1284–1298.
- Caro, Lucas y Araabi B.Nadjar. 1999. "Generalization of the Dempster-Shafer Theory: A fuzzy-Valued Measure." *IEEE Transactions on Fuzzy Systems* 7, no. 3 (June).
- Carrasco, E.F. y Rodriguez J. et al. 2004. "Diagnosis of acidification state in an anaerobic wastewater treatment plant using a fuzzy-based expert system." *Control Endineering Practice* 12 (1): 59–64.
- Castillo, E. et al. 1997. *Expert Systems and Probabilistic Network Models*. Springer, London.
- Cayrac, D. 1995. "Tesis Doctoral, Diagnostic opérationnel exploitant des connaissances incomplètes." Institut de Recherche en Informatique de Toulouse, Cedex.

- Cayrac, D. Dubois D. y Prade H. 1996. "Handling uncertainty with possibility theory and fuzzy sets in a satellite fault diagnosis application." *IEEE Trans. on Fuzzy Systems* 4 (3): 251–269.
- Chacón, J. 2001. *Tesis Doctoral, Diagnóstico de fallos mediante la utilización de información incompleta e incierta*. UPV Valencia España.
- Chakravarthy, S. et al. 2002. "Noncontact Level Sensing Technique Using Computer Vision." *IEEE Trans. Instrum. Meas.* 51, no. 2 (April).
- Chang, C. Combs J. y Stachowitz R. 1990. "A report on the expert systems validation associate." *Expert Systems with Applications* 1 (3): 217–230.
- Chen-Fu, Chien Shi-Lin Chen y Yih-Shin Lin. 2002. "Using Bayesian Network for Fault Location on Distribution Feeder." *IEEE TRANSACTIONS ON POWER DELIVERY* 17, no. 13 (JULY).
- Chiang, L.H. et al. 2001. *Fault Detection and Diagnosis in Industrial Systems*. Springer-Vergal.
- Cooman, De G. 1997. "Possibility theory II: conditional possibility." *International Journal of General Systems* 25 (4): 325–351.
- Cooper, V.A. et al. 2007. "Calibration of conceptual rainfall-runoff models using global optimisation methods with hydrologic process-based parameter constraints." *journal of Hydrology* 334 (3-4): 455–466 (Feb.).
- Cuena, J. 1985. *Lógica informática*. Alianza-Informática.
- Czogala, E. Mrozek A. y Pawlak Z. 1995. "The idea of a rough fuzzy controller and its application to the stabilization of a pendulum-car system." *Fuzzy Sets and Systems* 72 (1): 61–73.
- Davis, P. et al. 2007. "A physical probabilistic model to predict failure rates in buried PVC pipelines." *Reliability Engineering y System Safety* 92 (9): 1258–1266 (Sep.).
- De, S.K. et al. 2001. "An application of intuitionistic fuzzy sets in medical diagnosis." *Fuzzy Sets and Systems* 117 (2): 325–351.
- de Miguel, Luis J. y Blázquez L. Felipe. 2005. "Fuzzy logic-based decision-making for fault diagnosis in a DC motor." *Engineering Applications of Artificial Intelligence, Elsevier* 18 (4): 423–450 (June).
- Dempster, A. P. 1967. "Upper and lower probabilities induced by multivalued mapping." *Ann. Math. Stat.* 38:325–339.
- Detroja, K.P. et al. 2006. "A possibilistic clustering approach to novel fault detection and isolation." *Journal of Process Control, Elsevier* 16 (10): 1055–1073 (Dec.).
- Dey, S. y Stori J. A. 2005. "A Bayesian network approach to root cause diagnosis of process variations." *International Journal of Machine Tools and Manufacture, Elsevier* 45 (1): 75–91 (Jan.).
- Driankov, D. Hellendoorn H. y Reinfrank M. 1996. *An Introduction to Fuzzy Control*. Springer.

- Dubois, D. Fargier H. y Prade H. 1996. "Possibility theory in constrain satisfaction problem: handling priority, preference and uncertainty." *Applied Inteligence* 6 (4): 287–309.
- Dubois, D. Grabrich M. De Mouzon O. y Prade H. 2001a. "A possibilistic framework for single-fault causal diagnosis under uncertainty." *Int. J. General Systems* 30 (2): 167–192.
- Dubois, D. y Prade H. 1982. "On several representations of an uncertain body of evidence." Edited by Gupta M.M. y Sanchez E. New York:NorthHolland, *Fuzzy Information and Decision Processes*. 167–182.
- . 1988. *Possibility theory an approach to computerized processing of uncertainty*. New York: Plenum.
- . 1995. "Fuzzy relation equation and causal reasoning." *Fuzzy Sets and Systems* 45 (2): 119–134.
- . 1997. "The three semantics of fuzzy sets." *Fuzzy Sets and Systems* 90 (2): 142–150.
- . 2000a. *Fundamentals Of Fuzzy Sets*. Kluwer Academic.
- . 2004. "Possibilistic logic: a retrospective and prospective view." *Fuzzy Sets and Systems, Elsevier* 144:3–23.
- . 2001b. "Possibility theory, probability theory and multiple-valued logics: a clarification." *Annals of Mathematics and Artificial Inteligence* 32 (1): 35–66.
- Dubois, Didier y Prade Henri (editores). 2000b. *Fundamentals Of Fuzzy Sets*. Kluwer Academic.
- Facchinetti, G. 2006. "Evaluations of fuzzy quantities." *Fuzzy Sets and Systems, Elsevier* 157 (7): 892–903 (April).
- Frank, H. 1996. "A new axiom system of fuzzy logic." *Fuzzy Sets and Systems* 77:203–205.
- Frank, P. M. y Ding X. 1997. "Survey of robust residual generation and evaluation methods in observer-based fault detection systems." *J. Proc. Conr. Elsevier Science* 7 (6): 403–424.
- Freeman, R. y Kokotovic P. 1996. *Robust nonlinear control design. state-space and Lyapunov techniques*. Boston: Birkhauser.
- García, E. 2000. *Tesis Doctoral, Descomposición modular de diagnosticadores de fallos basados en modelos de eventos*. UPV Valencia España.
- García, E. y Morant F. 1999. "Modular Fault Diagnosis Based on Discret Events system for a mixer chemical process." *IEEE Proceedings*. Departament of Engineering Systems Computers and Automation, Polytechnic University of Valencia, Spain.
- Gass, S.I. 2003. *Linear Programming: methods and applications*. Dover, 5th edition.
- Gertler, J. 1995. "Generating Directional Residuals with Dynamic Parity Equations." *Automatica* 31 (4): 627–635.

- Gharamani, Z. 2001. “An introduction to hidden markov models and bayesian networks.” *Int. J. of Patterns Recognition and Artif. Intell.* 15 (1): 9–24.
- Giles, R. 1988. “The concept of grade of membership.” *Fuzzy Sets and Systems* 25:297–323.
- . 1997. “Lukasiewicz logic and fuzzy set theory.” *Int. J. Man-Machine Studies* 8:313–327.
- Giordani, P. y Kiers H. A.L. 2006. “A comparison of three methods for principal component analysis of fuzzy interval data.” *Computational Statistics and Data Analysis, Elsevier* 51 (1): 379–397 (Nov.).
- Gong, Shaogang y Xiang Tao. 2003. “Recognition of Group Activities using Dynamic Probabilistic Networks.” *Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV03)*.
- Guang, Y. y Ning W. 2004. “Fuzzy controller with neuron self-tuning parameters.” *WCICA 04. Fifth World Congress on Intelligent Control and Automation, IEEE* 3 (June): 2491 – 2494.
- Gupta, A. y Prasad, B. 1988. *Principles of expert systems*. IEEE Press.
- Haider, Sajjad. 2003. “On computing marginal probability intervals in inference networks.” *Systems, Man and Cybernetics, 2003. IEEE International Conference on* 5 (Oct.): 4724– 4729.
- Hellendoorn, H. y Thomas C. 1993. “defuzzification in fuzzy controllers.” *Journal of Intelligent and Fuzzy Systems* 1 (2): 109–123.
- Henna, V. y Ottomanelli M. 2006. “Handling uncertainty in route choice models: From probabilistic to possibilistic approaches.” *European Journal of Operational Research, Elsevier* 175 (3): 1526–1538 (Dec.).
- Hégarat-Masclé, Sylvie, Bloch Isabelle y Vidal-Madjar D. 1997. “Application of Dempster-Shafer Evidence Theory to Unsupervised Classification in Multisource Remote Sensing.” *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING* 35, no. 4 (JULY).
- Hisdal, E. 1995. *Reconciliation of the yes-no versus grade of membership dualism in human thinking. In Approximate reasoning in expert systems*. Amsterdam: North-Holland.
- Hui, Zhang et al. 2003. “The research on rolling element bearing fault based on wavelet packets transform.” *Industrial Electronics Society, IECON 03. The 29th Annual Conference of the IEEE* 2 (Nov.): 1745 – 1749.
- . 2004. “Extraction of failure characteristics of rolling element bearing based on wavelet transform under strong noise.” *Industrial Technology, IEEE ICIT 04. IEEE International Conference on* 2 (Dec.): 713 – 717.
- Isermann, R. 1989. *Digital control systems*. Springer-Verlag.
- . 1994. “On Fuzzy Logic Application for Automatic Control, Supervision and Fault Diagnostic.” *Third European Congress on Fuzzy and Intelligent Technologies EUFIT 95, Aachen* 2:738–753.

- . 1997a. “Supervision, Fault-Detection and Fault- Diagnosis Methods An Introduction.” *Control Engineering Practice* 5 (5): 639–652.
- Isermann, R. y Ballé P. 1997b. “Trends in the application of model-based fault detection and diagnosis of thecnical processes.” *Control Eng. Practice* 5 (5): 709–119.
- Isermann, Rolf. 1998. “On Fuzzy Logic Applications for Automatic Control, Supervision, and Fault Diagnosis.” *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART A: SYSTEMS AND HUMANS* 28, no. 2 (MARCH).
- Jager, R. 1995. “Fuzzy logic in control.” Unpublished doctoral dissertation, Delft University of Technology, Delft (Netherlands).
- Jang, J.-S. R. Sun C.-T. y Mizutani E. 1997. *Neuro-Fuzzy and Soft Computing*. Prentice Hall.
- Jianbin, Wei y Cheng-Zhong Xu. 2006. “Consistent proportional delay differentiation: A fuzzy control approach.” *Computer Networks, Elsevier* 51 (8): 2015–2032 (Oct.).
- Jian-Da, Wu et al. 2006. “Development of an expert system for fault diagnosis in scooter engine platform using fuzzy -logic inference.” *Expert Systems with Applications, Elsevier* 33 (4): 1063–1075 (Sep.).
- Jin-Fu, Li y Cheng-Wen Wu. 2002. “Efficient FFT network testing and diagnosis schemes.” *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on* 10 (3): 267 – 278 (June).
- Kitowski, J. y Bargiel M. 1987. “Diagnosis of faulty states in complex physical systems using fuzzy relational equations.” Edited by E. Sanchez and Zadeh L.A., *Approx. Reasoning in Intelligent Systems, Decision and Control*. Pergamon, 175–194.
- Knigh, Kevin. 1991. *Artificial Intelligence*. Prentice Hall.
- Korbicz, Józef. 2006. “Neuro-fuzzy networks and their application to fault detection of dynamical systems.” *Engineering Applications of Artificial Intelligence, Elsevier*, November.
- Kosko, B. 1992. *Neural networks and fuzzy systems. A dynamical systems approach to machine intelligence*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Kreinovich, V. et al. 2007. “Ellipsoids and ellipsoid-shaped fuzzy sets as natural multivariate generalization of intervals and fuzzy numbers: How to elicit them from users, and how to use them in data processing.” *Information Sciences, Elsevier* 177 (2): 388–407 (Jan.).
- Kruse, R. et al. 1991. *Uncertainty and vagueness in Knowledge Based Systems: Numerical Methods (Artificial Intelligence)*. Springer-Verlag.
- Langari, R. y Berenji H. 1992. “Fuzzy logic in control engineering.” Edited by D. y Sofge D. White, *Handbook of intelligent control. Neural, fuzzy and adaptive approaches*. New York, NY: Van Nostrand Reinhold.
- Lardon, Laurent et al. 2004. “On-line diagnosis and uncertainty management using evidence theory experimental illustration to anaerobic digestion processes.” *Journal of Process Control* 14:747 –763.

- Larsson, J. 1990. "An expert system for frequency response analysis." *In Proc. of IFAC 11th triennial world congress. Tallinn, Estonia.*
- Li, X. et al. 2002. "Multifrequency CW-based time-delay estimation for proximity ultrasonic sensors." *IEE Proc.-Radar Sonar Navig.* 149, no. 2 (April).
- Lindley, D. 1987. "The probability approach to the treatment of uncertainty in artificial intelligence and expert systems." *Statistical Science* 2 (1): 17–24.
- Liu, E. William K. Felt E. Sangiovanni-Vicentelli A. 1994. "Analog Testability Analysis and Fault Diagnosis Using Behavioral Modeling." *Custom Integrated Circuit Conference IEEE.*
- Liu, Edwina y Zhang Du. 2002. "Diagnosis of Component Failures in the Space Shuttle Maine Engines Using Bayesian Belief Networks: A Feasibility Study." *Proceedings of the 14th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'02).*
- Lo, C.H. et al. 2007. "Fuzzy -genetic algorithm for automatic fault detection in HVAC systems." *Applied Soft Computing, Elsevier* 7 (2): 554–560 (March).
- Lukasiewicz, J. 1920. "Logike trojwartoscieweg." *Ruch. Filosofiece*, vol. 169.
- Mamdani, E. y Efstathiou J. 1984. "An analysis of formal logics as inference mechanisms in expert systems." *Int. J. Man-Machine Studies* 21 (3): 213–227.
- Manevitz, L. y Yousef M. 2007. "One-class document classification via Neural Networks." *Neurocomputing, Elsevier* 70 (7-9): 1466–1481 (March).
- Mast, T. M. et al. 1999. "Bayesian Belief Networks for Faults Identification in Aircraft Gas Turbine." *Proceedings of the IEEE International Conference on Control Applications.*
- Mehrambod, N. et al. 2003. "Sensor Fault Detection and Identificación via Bayesian Belief Networks." *Proceeding of American Control Conference, IEEE.*
- Moazeni, S. 2006. "Fuzzy shortest path problem with finite fuzzy quantities." *Applied Mathematics and Computation, Elsevier* 183 (1): 160–169 (Dec.).
- Molochí, Arce Manuel y Sala Garrido Ramón. 1999. *Decisiones de optimización.* Tirant lo blanch, Valencia.
- Montgomery, C. Douglas y Runger C. George. 1996. *Probabilidad y estadística aplicada a la ingeniería.* McGraw-Hill.
- Mouzon, O. Dubois D. y Prade H. 2000. "Using consistency and abduction based indices in possibilistic causal diagnosis." *In Proc. IEEE Int. Fuzzy Systems Conf.*, pp. 729–734.
- . 2001. "Twofold fuzzy sets in single and multiple fault diagnosis, using information about normal values." *In Proc. IEEE Int. Fuzzy Systems Conf.*, pp. 1563–1566.
- Nguyen, T. Perkins W. Laffey T. y Pecora D. 1987. "Knowledge base verification." *AI Magazine* 8 (2): 69–75.

- Nilsson, N. 2001. *Inteligencia Artificial una nueva síntesis*. McGraw-Hill Interamericana de España, S.A.U.
- Ogata, K. 1998. *Ingeniería de Control Moderna tercera edición*. Prentice Hall.
- Pan, Xia. 2005. “Notes on shift effects for T2-type charts on multivariate ARMA residuals.” *Computers and Industrial Engineering, Elsevier* 49 (3): 381–392 (Nov.).
- Patel, A. V. 2002. “Analytical structures and analysis of fuzzy PD controllers with multifuzzy sets having variable cross-point level.” *Fuzzy Sets and Systems, Elsevier* 129 (3): 311–334 (Aug.).
- Pawar, Prashant M. y Ganguli Ranjan. 2003. “Genetic fuzzy system for damage detection in beams and helicopter rotor blades.” *Computer Methods in Applied Mechanics and Engineering, Elsevier* 192 (16-18): 2031–2057 (April).
- Pawlak, Z. 1982. “Rough sets.” *Int. J. Inform. Comput. Science* 11:341–356.
- Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.
- Pedrycz, W. 1993. *Fuzzy control and fuzzy systems (second ed.)*. New York, NY: J.Wiley and Sons.
- Pedrycz, W. y Reformat M. 1997. “Rule-based modeling of nonlinear relationships.” *IEEE Transactions on Fuzzy Systems* 5 (2): 256–269.
- Peng, H. y Ding C. 2003. “Structure search and stability enhancement of Bayesian networks.” *ICDM 2003. Third IEEE International Conference on Data Mining*, Nov., 621 – 624.
- Peng, Y. y Reggia J.A. 1990. *Abductive inference models for diagnostic problem-solving*. Springer.
- Preece, A. y Shingal R. 1994. “Foundation and application of knowledge base verification.” *International Journal of Intelligent Systems* 9:683–701.
- Przytula, K.W. y Thompson D. 2000. “Construction of Bayesian networks for diagnostics.” *Aerospace Conference Proceedings IEEE* 5 (March): 193 – 200.
- Rai, V.K. y Mohanty A.R. 2006. “Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert-Huang transform.” *Mechanical Systems and Signal Processing, Elsevier*.
- Ramírez, J. Sala A. y Morant F. 2002. “Sistema experto Borroso para el diagnóstico.” *Electro 2002, Automática e Informática Industrial* 1:533–537. Instituto Tecnológico de Chihuahua, Méx.
- Ramírez, J. y Sala A. 2004. “Diagnóstico de sistemas mediante combinación de lógica borrosa y probabilidad.” *Electro 2004* 1:341–346. Instituto Tecnológico de Chihuahua, Méx.
- . 2006. “Learning Bayesian Networks for Systems Diagnosis.” *Proceedings of the Electronics, Robotics and Automotive Mechanics Conference (CERMA'06), IEEE Computer Society* 2:123–128.
- Rescher, N. 1969. *Many-valued logics*. New York: McGraw-Hill.

- Russell, S. y Norving P. 1996. *Inteligencia Artificial Un enfoque moderno*. Prentice Hall Hispanoamérica.
- . 2003. *Artificial Intelligence: a modern approach, 2nd edition*. Prentice-Hall.
- Sahin, F. et al. 2007. “Fault diagnosis for airplane engines using Bayesian networks and distributed particle swarm optimization.” *Parallel Computing, Trends in Parallel Computing, Elsevier* 33 (2): 124–143 (March).
- Sainz, G.I. et al. 2005. “Fault detection and fuzzy rule extraction in AC motors by a neuro- fuzzy ART-based system.” *Engineering Applications of Artificial Intelligence, Elsevier* 18 (7): 867–874 (Oct.).
- Sala, A. 1995. “Análisis de la estructura del conocimiento en bases de reglas borrosas.” Valencia, Spain: DISCA-UPV (Tech. Rep.).
- . 2000. “Ingeniería de Control.” Editorial de la Universidad Politécnica de Valencia.
- . 2006a. “Encoding fuzzy possibilistic diagnostic as a constrained optimisation problem.” *Information Sciences*. Under Review.
- . 2001a. *Identificación y control Iterativo. II CURSO DE ESPECIALIZACIÓN EN AUTOMÁTICA CEA-IFAC, La Manga del Mar Menor (Murcia)*. CEA-IFAC Comité Español de Automática.
- Sala, A. Tormos B. et al. 2006b. “Fuzzy inference engine for condition monitoring (diagnosis). Oil analysis application.” *Preprint submitted to Elsevier Science*.
- Sala, A. y Albertos P. 1998. “Fuzzy Systems evaluation: The inference error approach.” *IEEE Trans. on Syst. Man and Cybernetics* 28B (2): 268–275.
- . 2001b. “Inference error minimisation: fuzzy modelling of ambiguous function.” *Fuzzy Sets and Systems* 121 (1): 95–111.
- Sanchez, E. 1976. “Inverses of fuzzy relations, application to possibility distribution and medical diagnosis.” *Fuzzy Sets and Systems* 2 (1): 75–86.
- Sanchez, E. N. et al. 2006. “Combining fuzzy, PID and regulation control for an autonomous mini-helicopter.” *Information Sciences, Elsevier* 177 (10): 1999–2022 (Oct.).
- Schrage, L.E. 1998. *linear, integer and quadratic programing with LINDO*. Course Technology.
- Shafer, G. 1976. *Mathematical Theory of Evidence*. Princeton, N.J., Princeton Univ. Press.
- Shafer, G. y Pearl J. (editores). 1990. *Reading in uncertain reasoning*. Morgan Kauffman, San Mateo (CA), USA.
- Shao-Zhong, Z. et al. 2002. “Construction and Application of Bayesian Networks in Flood Decision Supporting System.” *Proceedings of the First International Conference on Machine Learning and Cybernetics, IEEE*, Nov., 718–722.
- Sheppard, John W. 2003. “Accounting for False Indication in a Bayesian Diagnostic Framework.” *AUTOTESTCON 2003. IEEE Systems Readiness Technology Conference. Proceedings*, Sept., 273 – 279.

- Sierksma, G. 2001. *Linear and Integer Programming: Theory and Practice*. Marcel Dekker Pub., New York, 2nd edition.
- Slamani, Mustapha y Kaminska Bozena. 1992. "Analog Circuit Fault Diagnosis Based on Sensitivity Computation and Functional Testing." *Design and Test of Computers, IEEE*.
- Smith, C. et al. 2007. "An approach to vibration analysis using wavelets in an application of aircraft health monitoring." *Mechanical Systems and Signal Processing, Elsevier* 21 (3): 1255–1272 (April).
- Spurbeck, M. S. y Schreier P. J. 2006. "Causal Wiener filter banks for periodically correlated time series." *Signal Processing, Elsevier* 87 (6): 1179–1187 (Oct).
- Sterritt, R. et al. 2000. "Exploring dynamic Bayesian belief networks for intelligent fault management systems." *Systems, Man, and Cybernetics, 2000 IEEE International Conference on* 5 (Oct.): 3646 – 3652.
- Suwa, M. Scott A. y Shortliffe E. 1982. "An approach to verifying completeness and consistency in a rule-based expert system." *AI Magazine* 3 (4): 16–21.
- Thangaraj, D. et al. 1994. "Strem Function Distribution of Petroleum liquids in Relaxation Tanks." *IEEE Trans. Elect. Insul.*
- Tormos, Bernardo. 2002. "Contribución al diagnostico de motores Diesel basado en el análisis del lubricante usado. Tesis doctoral." Universidad Politécnica de Valencia.
- Tsukamoto, Y. y Terano T. 1977. "Failure Diagnosis by using fuzzy logic." *In Proc. of IEEE Control and Decision Conf.*, pp. 1390–1395.
- Wainer, J. y Rezender A. 1997. "A temporal extension to parsimonius covering theory." *Artificial Intelligence in Medicine* 10:235–255.
- Walker, E. 1994. "Stone algebras, conditional events and three valued logics." *IEEE Transactions on Systems, Man and Cybernetics* 24 (12): 1699–1707.
- Wang, J. et al. 2007. "Bayes classification based on minimum bounding spheres." *Neurocomputing, Elsevier* 70 (4-6): 801–808 (Jan).
- Wang, L.-X. 1994. *Adaptive fuzzy systems and control*. Englewood Cliffs, NJ: Prentice-Hall.
- Weidl, G. Madsen A.L. y Israelson S. 2005. "Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes." *Computers and Chemical Engineering, Elsevier* 29 (9): 1996–2009 (Aug.).
- Wei-Guo, Z. et al. 2007. "Possibilistic mean-variance models and efficient frontiers for portfolio selection problem." *Information Sciences, Elsevier*, Feb.
- Whei-Min, L. Chia-Hung L. y Zheng-Chi S. 2004. "Adaptive Multiple Fault Detection and Alarm Processing for Loop System With Probabilistic Network." *IEEE TRANSACTIONS ON POWER DELIVERY* 19, no. 1 (JAN.).
- Yager, Ronald R. y Filev Dimitar P. 1995. "Including Probabilistic Uncertainty in Fuzzy Logic Controller Modeling Using Dempster-Shafer Theory." *IEEE Transactions on systems, man, and cybernetics* 25, no. 8 (August).

- Yamada, K. 2002. "Possibility causality consistency problem based on asymmetrically-valued causal model." *FuzzySets and Systems* 132:33–48.
- Yen, John. 1990. "Generalizing the Dempster-Shafer Theory to Fuzzy Sets." *IEEE Transactions on systems, man, and cybernetics* 20, no. 3 (May/June).
- Yong, Hu et al. 2003. "A new method of determining the basic belief assignment in D-S evidence theory." *Proceedings of the Second International Conference on Machine Learning and Cybernetics, Xi'an*, 2-5 November.
- Yu, Yongquan. 2001. "A defuzzification method relating to fuzzification in fuzzy control systems using quadrature fuzzy parting." *Fuzzy Systems, 2001. The 10th IEEE International Conference on* 3 (Dec.): 1483 – 1486.
- Zadeh, L. 1965. "Fuzzy sets." *Information and Control* 8:338–353.
- . 1978. "Fuzzy sets as a basis for a theory of possibility." *Fuzzy Sets and Systems* 1:3–28.
- Zarei, Jafar y Poshtan Javad. 2006. "Bearing fault detection using wavelet packet transform of induction motor stator current." *Tribology International, Elsevier* 40 (5): 763–769 (july).
- Zeshui, X. y Jian C. 2005. "Some models for deriving the priority weights from interval fuzzy preference relations." *European Journal of Operational Research, Elsevier*, Dec.
- Zimmermann, H. y Zysno P. 1985. "Quantifying vagueness in decision models." *European Journal of Operational Research* 22:148–154.
- Özyer, T. et al. 2007. "Intrusion detection by integrating boosting genetic fuzzy classifier and data mining criteria for rule pre-screening." *Journal of Network and Computer Applications, Elsevier* 30 (1): 99–113 (Jan.).