Universidad Don Bosco.
Facultad de Ingeniería.
Escuela de computación.

Investigación 1

“Razonamiento con Incertidumbre”.

Catedrático: Ing. Cruz Antonio Galdámez

Materia: Sistemas Expetos

Alumnos:

Castro Quiñónez Geraldina Altagracia CQ000097
Hernández Marín Gilma Lissette HM990229

Grupo: 04

Ciudadela Don Bosco, 30 de Junio de 2004

INTRODUCCIÓN

No Página

INTRODUCCION

En el siguiente trabajo, se presenta la definición de algunos temas que ayudaran a desarrollar la materia Sistemas Expertos. Estos temas son parte de la unidad 4, por lo que algunos temas serán continuación o serán el comienzo de otros, esto provoca que no estén completos.

Se menciona parte de la historia de incertidumbre, que comenzó con los inicios de la inteligencia artificial, y luego con sistemas expertos. La incertidumbre con la probabilidad, tienen una gran importación para poder “predecir” ciertos resultados a obtener en cualquier experimento.

RAZONAMIENTO CON INCERTIDUMBRE

1. INCERTIDUMBRE.

1.1 Definición De Incertidumbre

Históricamente los conceptos relacionados con la *incertidumbre* experimental han sido pobremente sistematizados. A finales del siglo XX los organismos competentes en la armonización de las ciencias básicas y aplicadas, como por ejemplo el BIMP, IUPAP o, recientemente, la ISO, creyeron conveniente establecer un proceso normalizado para determinar y propagar las incertidumbre de los datos experiementales. La idea es establecer un *consenso internacional para la expresión armonizada de incertidumbres de medida*. Las definiciones resultan un tanto contradictorias con respecto al lenguaje que se venía usando hasta entonces, desde el mismo concepto de *incertidumbre* que debe sustituir a lo que se conocía normalmente como *error de una medida*. Los conceptos más importantes en relación con la *incertidumbre experimental* son:

Valor verdadero: Es el valor de una *magnitud física particular* objeto de medida y asignable a un cuerpo, sustancia o fenómeno. Por el valor de la masa de la Tierra. Es un valor desconocido por definición ya que lo único que puede determinarse es lo que se conoce como *valor medido* que se define a continuación.

Valor medido: Es la expresión numérica del valor de una *magnitud física particular* obtenido como resultado de una medida experimental.

Error (de la medida): Es la diferencia entre el valor medido y el valor verdadero. El *error* de una medida no puede determinarse por cuanto está relacionado con el valor verdadero (desconocido).

Incertidumbre (absoluta) (*u*): Es un valor estimado de tal forma que la banda determinada por $x + u$ y $x - u$ -donde x es el valor medido- contiene con razonable certeza el valor verdadero de la medida. También se le puede llamar *error estimado* o *incertidumbre*. Es un abuso de lenguaje intercambiar *incertidumbre* por *error*. Las últimas recomendaciones de la IUPAP, ISO, BIPM, CIPM y CGPM diferencian claramente estos conceptos y tratan de evitar su confusión.

Incertidumbre relativa (u_r): Es el cociente entre la incertidumbre absoluta y el valor medido. De su definición se deduce que es un número adimensional.

Exactitud: es lo cercano que está el valor medido al valor verdadero. Hay que tener en cuenta que, por definición, el valor verdadero de la medida no esta definido, por lo tanto el concepto *exactitud* es usualmente *cualitativo*. No obstante, la exactitud de un aparato se puede determinar en relación a un patrón de valor estándar, conocido y admitido por todos los laboratorios.

Repetitividad (de las medidas) -precisión-: Es la fidelidad de los valores experimentales de *una misma* magnitud física medidos bajo *idénticas* condiciones experimentales. Aquí el concepto *idéntico* incluye el mismo observador, instrumento de medida, lugar y procedimiento así como la cercanía en el tiempo. El concepto de *precisión* se define muchas veces como *repetitividad*. Es decir, un aparato *preciso* es un aparato *repetitivo* (diferentes medidas de una misma magnitud bajo las mismas condiciones conducen al mismo resultado). Tanto *repetitividad* como *precisión* son conceptos cualitativos.

Reproducibilidad (de las medidas): Es la fidelidad de los valores experimentales de *una misma* magnitud física medidos bajo *diferentes* condiciones experimentales.

Sensibilidad: Es el menor cambio en la medida que puede detectar un aparato de medida.

1.2 Desarrollo Histórico de Incertidumbre.

- *Inicios de IA* (50's, 60's) - no se consideró el uso de ``números'', en general y probabilidad, en particular, ya que el enfoque era en manejo simbólico vs. el resto de computación que era numérico. Se trabajó en mundos ``artificiales'' que en general no requerían manejar incertidumbre.

  A principios de los 60's se creía que lógica iba a ser suficientemente poderosa para resolver problemas reales.
- *Sistemas Expertos* (70's) - surgió la necesidad de manejo de incertidumbre en los primeros sistemas expertos en el mundo ``real''. Por diversas razones, no se consideró el uso de probabilidad y surgieron mecanismos alternativos *ad-hoc* como el de *MYCIN* (medicina) y *Prospector* (exploración minera). En este tiempo, desligados de IA, surgen nuevas teorías para representar incertidumbre como lógica difusa y teoría de *Dempster-Shafer.*
- *Resurgimiento de probabilidad* (80's) - resurge el uso de probabilidad para manejo de incertidumbre con el desarrollo de las redes Bayesianas. También los investigadores de IA ``descubren'' otras teorías y las aplican en sistemas expertos.
- *Diversos formalismos* (90's) - continúan varios formalismos para el manejo de incertidumbre sin haber un ``ganador'' definitivo. Paradójicamente, la mayor parte de los *shells* comerciales siguen usando técnicas *ad-hoc* a pesar de haberse publicado problemas teóricos y prácticos con este tipo de métodos.

1.3 Causas de Incertidumbre

- **Información**:
  - Incompleta (falta de análisis en medicina, falta de variables de campo en sistemas de control).
  - Poco confiable (medidores poco confiables, instrumentos imprecisos, análisis poco confiables).
  - Ruido, distorsión (ruido o distorsión en sistemas de visión, de reconocimiento de voz, de comunicaciones).
- **Conocimiento**:
  - Impreciso (si tiene dolor de cabeza posiblemente tiene gripe, el lumen es una región obscura, grande y uniforme).
  - Contradictorio (si tiene dolor de cabeza es probable que tenga gripe, pero también es posible que no tenga gripe, opiniones encontradas de diferentes expertos).
- **Representación**:
  - No adecuada (no se selecciono la representación(es) idónea(s) para la aplicación).
  - Falta de poder descriptivo (las representaciones no permiten representar adecuadamente el conocimiento del dominio, como lo expresa el experto).

Ejemplos de dominios con incertidumbre

- Diagnóstico medico
- Predicción financiera
- Exploración minera / petrolera
- Interpretación de imágenes (visión)
- Reconocimiento de voz
- Monitoreo / control de procesos industriales complejos

1.4 Efectos de Incertidumbre

Si pierden varias propiedades de los sistemas que no tienen incertidumbre, principalmente aquellos basados en lógicas o reglas, lo cual hace el manejo de incertidumbre más complejo. Las principales dos características de lógica de primer orden que, en general, ya no aplican son:

1. Modular
2. Monotónica

Modular: Un sistema de reglas es modular, ya que para saber la verdad de una regla sólo tiene que considerarla a ésta, sin importar el resto del conocimiento. Por ejemplo:

Si A entonces B

Si *A* es verdadero, *B* es verdadero independientemente otras reglas o datos.

Pero si hay incertidumbre:

Si A entonces posiblemente B, o si A entonces B con probabilidad = 0.7

Ya no puedo considerar la regla por si sola, debo tomar en cuenta otras reglas que involucren a *B*. Puede haber otra regla:

Si C entonces B con probabilidad = 0.9

Si *A* y *C* son verdaderos, cual es la probabilidad de *B*? 0.7? 0.9? C1ombinación de ambas?

Monotónicas: Un sistema es monotónico si al agregar nueva información a su base de datos, entonces no se alteran las conclusiones que seguían de la base de datos original. Por ejemplo:

Si A entonces B

Si *A* es verdadero, *B* es verdadero sin importar si más información se agrega a la memoria de trabajo. Pero si tenemos:

Si A entonces posiblemente B, o si A entonces B con probabilidad = 0.7

Ya no puedo considerar que la certeza en B no puede cambiar, debo tomar en cuanta otras reglas que involucren a B. Puede haber otra regla, como en el ejemplo anterior:

Si C entonces B con probabilidad = 0.9

Entonces en un tiempo t_1 A es verdadera y C falsa, por lo que la probabilidad de B es 0.7, pero en un tiempo t_2 C se hace verdadera y entonces cambiaría la probabilidad de B.
Ambos aspectos hacen más complejas las representaciones del conocimiento que toman en cuenta incertidumbre, lo que ha llevado al desarrollo de diversas técnicas para su manejo.

1.5 Técnicas de Incertidumbre.

- **No-numéricas**
  - Lógicas no-monotónicas
  - Sistemas de mantenimiento de verdad (TMS, ATMS)
  - Teorías de endosos

Las lógicas no monotónicas.

Estos sistemas deben ser no montónicas, lo que significa que el conjunto de conclusiones posibles puede decrecer cuando el conjunto de premisas se incrementa. A partir de un conjunto dado de información inicial, uno puede algunas veces, construir diversos conjuntos mutuamente inconsistente de conclusiones, generalmente llamadas extensiones, esta propiedad es llamada pluriextensionalidad. Es regularmente difícil caracterizar las extensiones de una forma directa y constructiva porque algunas formas de razonamiento revisable son circulares. El sistema no monotónico debe permitir obtener la extensión deseada solamente de un punto de vista semántico, las conclusiones que pueden ser obtenidas de un razonamiento revisable son generalmente consistentes con respecto a las premisas. En realidad, una pieza de razonamiento revisable no es necesariamente 'lógicamente correcta', pueden ser inferidas condiciones plausibles que no son 'válidas'.
Enfaticemos también que una estructura formal para razonamiento revisable debe de ser suficientemente flexible para tratar de razonar con premisas que

contienen información incompleta, tal estructura debe de ser suficientemente robusta para apoyar de una forma conveniente la posible evolución del conocimiento representado, debe de permitir la circunscripción dinámica de la información que es realmente relevante al problema, por ejemplo debe de permitir razonamiento a pesar de la ausencia de información completa, o de forma mas general, debe apoyar la posible evolución tanto de el dominio de aplicación como de la misma conceptualización.

- **Numéricas**
  - Empíricas (MYCIN, Prospector)
  - Métodos aproximados
  - Lógica difusa
  - Teoría de Dempster-Shafer
  - Redes Bayesianas

1.6 Cálculo de incertidumbres

La incertidumbre se calcula de forma diferente dependiendo de si el valor de la magnitud se observa directamente en un instrumento de medida (medida directa) o si se obtiene manipulando matemáticamente una o varias medidas directas (medida indirecta).

<u>En una práctica calcularemos primero la incertidumbre de las medidas directas y luego la de las indirectas.</u>

1.7 Incertidumbre de las Medidas

Todas las ciencias experimentales se fundamentan en la experiencia, y ésta a su vez en la determinación cuantitativa de las magnitudes pertinentes. En definitiva, todas las ciencias precisan de la medida, bien directa, bien indirecta de magnitudes físicas. Medir implica generalmente comparar la magnitud objeto de la medida con un patrón. El resultado de la medida se expresa con un número y una unidad, dependiendo esta última del patrón que se haya escogido.

Las medidas nunca permiten obtener el ``verdadero valor'' de la magnitud que se mide. Esto es debido a multitud de razones. Las más evidentes son las imperfecciones, inevitables en un cierto grado, de los aparatos y de nuestros sentidos. El ``verdadero valor'' de una magnitud no es accesible en la realidad y por ello resulta más propio hablar de estimaciones, medidas o aproximaciones del valor de una magnitud. Independientemente de estas consideraciones, en el ámbito de la Física se sabe que no tiene sentido hablar del *valor* de una magnitud, sino sólo de la *probabilidad* de obtener uno u otro valor en una determinada medida de esta magnitud. Esto no es el resultado de las imperfecciones de los aparatos, sino de la propia esencia de la naturaleza. Este carácter probabilístico de las magnitudes se hace patente a nivel microscópico.

La consecuencia de las consideraciones anteriores, es que toda medida es incierta o está dotada de un cierto grado de incertidumbre. Es esencial estimar ésta incertidumbre, primero porque el conocimiento de la incertidumbre aumenta la información que proporciona la medida, y segundo, porque este conocimiento permite manejar las medidas con la prudencia que dicta el conocimiento de la confianza que nos merecen.

Cuando se exprese el resultado de una medida es pues necesario especificar *tres* elementos: número, unidad e incertidumbre. La ausencia de alguna de ellas elimina o limita la información que proporciona.

2. TIPOS DE ERROR

2.1 Errores

El significado de la palabra ``error'' no es muy preciso, puesto que con frecuencia autores diferentes lo emplean con sentidos diferentes. En un sentido amplio puede considerarse el error como una estimación o cuantificación de la incertidumbre de una medida. Cuanto más incierta sea una medida, tanto mayor será el error que lleva aparejado.
Suelen distinguirse dos tipos de errores: errores sistemáticos y accidentales.

Errores Sistemáticos

Como su nombre indica, no son debidos al azar o a causas no controlables. Pueden surgir de emplear un método inadecuado, un instrumento defectuoso o bien por usarlo en condiciones para las que no estaba previsto su uso. Por ejemplo, emplear una regla metálica a una temperatura muy alta, puede introducir un error sistemático si la dilatación del material hace que su longitud sea mayor que la nominal. En este caso, todas las medidas pecarán (sistemáticamente) por defecto. El error podría evitarse eligiendo un material de coeficiente de dilatación bajo o controlando la temperatura a la que se mide.
Medir temperaturas con un termómetro graduado en grados Farenhait, suponiendo por equivocación que está graduado en grados Celsius, introduce también un error sistemático en la medida. El error se evita en este caso recabando información sobre la escala del termómetro.
Los errores sistemáticos no son objeto de la teoría de errores. Realmente son equivocaciones que pueden y deben evitarse, empleando métodos e instrumentos de medida correctos y adecuados a los fines que se deseen obtener.

Errores Accidentales

Estos son los que llamaremos simplemente errores en el sentido técnico de la palabra. Son incertidumbres debidas a numerosas causas incontrolables e imprevisibles que dan lugar a resultados distintos cuando se repite la medida en condiciones idénticas.

Los errores accidentales, o errores propiamente dichos, parecen fruto del azar, y por ello reciben el nombre de errores aleatorios. Pueden ser debidos a la acumulación de muchas incertidumbres sistemáticas incontrolables o bien pueden provenir de variaciones intrínsecamente aleatorias a nivel microscópico. En ambos casos el resultado es que las medidas de una magnitud siguen una distribución de probabilidad, que puede analizarse por medios estadísticos. Aunque la presencia de los errores accidentales no pueda evitarse, sí puede estimarse su magnitud por medio de estos métodos estadísticos.

Error Absoluto

Por motivos obvios, y por su propia naturaleza, no es posible determinar exactamente un error. En el mejor de los casos, puede llegarse a una estimación de ese error. Cuando el resultado de una medida se expresa por:

$$valor\ medido = x \pm \delta x (unidad)$$

lo que se quiere decir es que la magnitud medida se encuentran en el intervalo $(x - \delta x, x + \delta x)$ con una determinada probabilidad. Con una medida logramos acotar el intervalo de valores en los que se encuentra la magnitud que pretendemos medir, pero siempre con una determinada probabilidad. Es evidente que el error expresado por δx es una magnitud de la misma clase que la medida y se expresa por tanto con la misma unidad. También es claro que en las medidas de calidad normal el error δx debe ser mucho menor que el valor nominal, *x*. Por definición δx es siempre positivo.

Error relativo

El error definido arriba se llama error absoluto. Tiene también interés el error relativo, que se define como el cociente del error absoluto, δx dividido por $|x|$.

$$error\ relativo = \frac{\delta x}{|x|}$$

En medidas de una cierta calidad el error relativo debe ser mucho menor que la unidad. Frecuentemente se expresa multiplicado por 100, con lo que aparece en tanto por ciento del valor medido:

$$error\ relativo(\%) = \frac{\delta x}{|x|}.100$$

2.2 Error e incertidumbre

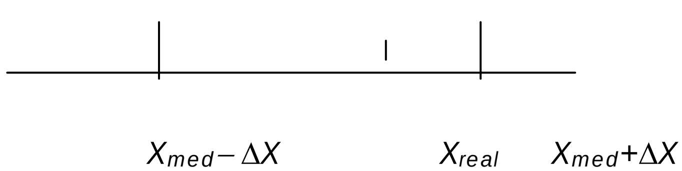
En un procedimiento experimental que nos proporciona el valor de una magnitud *X*, el resultado no coincide exactamente con el valor real de dicha magnitud. La diferencia entre el valor real y el valor medido se llama **error de la medida**:

$$\zeta = X_{med} - X_{real}$$

El error es siempre desconocido, pero puede estimarse una cota superior para su valor absoluto. Esta cota se denomina **incertidumbre de la medida** y se denota por ΔX. De la definición de error y de incertidumbre deducimos que el valor real de la medida se encuentra en el intervalo:

$$X_{real} \in [X_{med} - \Delta X, X_{med} + \Delta X]$$

Gráficamente podemos representar esta situación de la siguiente forma:



X_{med} se encuentra en el punto medio del intervalo. Por ello, el resultado de una medida se escribe siempre en la forma:

$$X = X_{med} \pm \Delta X$$

A veces es útil comparar el error de una medida con el valor de la misma. Se define para ello la **incertidumbre relativa** de una medida como el cociente:

$$\delta = \frac{\Delta X}{X_{med}}$$

Para distinguirla de la incertidumbre relativa, la incertidumbre ΔX se denomina **incertidumbre absoluta**. La incertidumbre relativa es útil para los comentarios de las prácticas. Sin embargo, para expresar el resultado de una medida hay que utilizar siempre las incertidumbres absolutas. Obsérvese que la incertidumbre relativa es adimensional (puede también expresarse en tanto por ciento) mientras que la absoluta tiene las mismas unidades que la magnitud medida.

Hemos visto la diferencia entre los conceptos *error* e *incertidumbre*. Distinguirlos facilita la comprensión de la teoría de errores. Sin embargo, por comodidad, es muy frecuente utilizar la palabra error para referirse a la incertidumbre de una medida.

3. PROBABILIDAD

Probabilidad de un suceso es el número al que tiende la frecuencia relativa asociada al suceso a medida que el número de veces que se realiza el experimento crece.

Un experimento aleatorio se caracteriza porque repetido muchas veces y en idénticas condiciones el cociente entre el número de veces que aparece un resultado (suceso) y el número total de veces que se realiza el experimento tiende a un número fijo. Esta propiedad es conocida como **ley de los grandes**

números, establecida por *Jakob Bernouilli*. Tiene el inconveniente de variar la sucesión de las frecuencias relativas de unas series de realizaciones a otras, si bien el valor al que se aproximan a medida que el número de realizaciones aumenta se mantiene estable.

La frecuencia relativa del suceso *A*:

$$f_r(A) = \frac{\text{número de veces que aparece } A}{\text{número de veces que se realiza el experimento}}$$

Propiedades de la frecuencia relativa:

1. $0 \leq f_r(A) \leq 1$ cualquiera que sea el suceso *A*.
2. $f_r(A \cup B) = f_r(A) + f_r(B)$ si $A \cap B = \varnothing$.
3. $f_r(E) = 1$ $f_r(\varnothing) = 0$.

Esta definición presenta el inconveniente de tener que realizar el experimento un gran número de veces y además siempre obtendremos un valor aproximado de la probabilidad.

3.1 DEFINICIÓN AXIOMÁTICA

La definición axiomática de probabilidad se debe a *Kolmogorov*, quien consideró la relación entre la frecuencia relativa de un suceso y su probabilidad cuando el número de veces que se realiza el experimento es muy grande.

Sea *E* el espacio muestral de cierto experimento aleatorio. La ***Probabilidad*** de cada suceso es un número que verifica:

4. Cualquiera que sea el suceso *A*, $P(A) \geq 0$.
5. Si dos sucesos son incompatibles, la probabilidad de su unión es igual a la suma de sus probabilidades.

   $A \cap B = \varnothing \Rightarrow P(A \cup B) = P(A) + P(B)$.

La probabilidad total es 1. $P(E) = 1$.

3.2 DEFINICIÓN DE LAPLACE.

En el caso de que todos los sucesos elementales del espacio muestral *E* sean equiprobables, *Laplace* define la probabilidad del suceso *A* como el cociente entre el número de resultados favorables a que ocurra el suceso *A* en el experimento y el número de resultados posibles del experimento.

Si $E = \{x_1, x_2, ..., x_k\}$ y $P(x_1) = P(x_2) = ... = P(x_k)$, entonces :

$$P(A) = \frac{\text{número de casos favorables al suceso } A}{\text{número de casos posibles}}$$

PROPIEDADES

1. $P(\overline{A}) = 1 - P(A)$
2. $P(\varnothing) = 0$
3. Si $A \subset B \Rightarrow P(B) = P(A) + P(A - B)$
4. Si $A \subset B \Rightarrow P(A) \leq P(B)$
5. Si $A_1, A_2, ..., A_k$, son incompatibles dos a dos, entonces:
   $P(A_1 \cup A_2 \cup ... \cup A_k) = P(A_1) + P(A_2) + ... + P(A_k)$
6. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
7. Si el espacio muestral *E* es finito y un sucesos es $A=\{x_1, x_2, ..., x_K\}$, entonces:
   $P(A) = P(x_1) + P(x_2) + ... + P(x_K)$

3.3 PROBABILIDAD CONDICIONADA

El conocimiento de que ha ocurrido el suceso *A* modifica, en algunas ocasiones, la probabilidad del suceso *B*, pero en otras no. Los sucesos en los que, conociendo que uno ha ocurrido, no se modifica la probabilidad del otro, decimos que son **independientes** y, si se modifica, decimos que son **dependientes** entre sí.

Decimos que dos sucesos *A* y *B* son ***independientes*** entre sí si la ocurrencia de uno de ellos no modifica la probabilidad del otro, es decir, si

$P(B/A) = P(B)$ ó $P(A/B) = P(A)$

Decimos que dos sucesos *A* y *B* son ***dependientes*** entre sí si la ocurrencia de uno de ellos modifica la probabilidad del otro, es decir, si

$P(B/A) \neq P(B)$ ó $P(A/B) \neq P(A)$

Como consecuencia inmediata de la definición se tiene:

- Dos sucesos *A* y *B* son independientes si se cumple:

  $P(A \cap B) = P(A) \cdot P(B)$

- Tres sucesos *A*, *B* y *C* son independientes si se cumplen a la vez:

  $P(A \cap B) = P(A) \cdot P(B)$

  $P(A \cap C) = P(A) \cdot P(C)$

  $P(B \cap C) = P(B) \cdot P(C)$

  $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

En el cálculo de las probabilidades de algunos sucesos, el valor de dicha probabilidad vará en función del conocimiento de determinadas informaciones relativas a estos sucesos. Veamos un ejemplo.

Si disponemos de una urna que contiene cuatro bolas numeradas del 1 al 4, extraemos una bola y seguidamente la volvemos a introducir para realizar una segunda extracción, la probabilidad de extraer, por ejemplo, la bola número 3 en la segunda extracción es la misma que en la primera.

Si realizamos el mismo proceso sin reemplazar la bola extraída la probabilidad de extraer, por ejemplo, la bola número 3 en la segunda extracción dependerá de la bola extraída en primer lugar.

Sean *A* y *B* dos sucesos tal que $P(A) \neq 0$, se llama ***probabilidad de B condicionada a A***, *P(B/A)*, a la probabilidad de *B* tomando como espacio muestral *A*, es decir, la probabilidad de que ocurra *B* dado que ha sucedido *A*.

$$P(B/A) = \frac{P(B \cap A)}{P(A)}$$

De esta igualdad se deduce: $P(B \cap A) = P(B/A) \cdot P(A)$

La fórmula anterior adopta la forma para tres sucesos, *A, B* y *C*:

$P(A \cap B \cap C) = P(A) \cdot P(B/A) \cdot P(C/A \cap B)$

Esta fórmula admite una generalización para un número cualquiera de sucesos.

3.4 Probabilidad Total

Llamamos **sistema completo de sucesos** a una familia de sucesos $A_1, A_2, \ldots, A_n$ que cumplen:

1. Son incompatibles dos a dos, $A_i \cap A_j = \varnothing$
2. La unión de todos ellos es el suceso seguro, $\bigcup_{i=1}^{n} A_i = E$

TEOREMA DE LA PROBABILIDAD TOTAL

Sea $A_1, A_2, \ldots, A_n$ un sistema completo de sucesos tales que la probabilidad de cada uno de ellos es distinta de cero, y sea B un suceso cualquier del que se conocen las probabilidades condicionales $P(B/A_i)$, entonces la probabilidad del suceso B viene dada por la expresión:

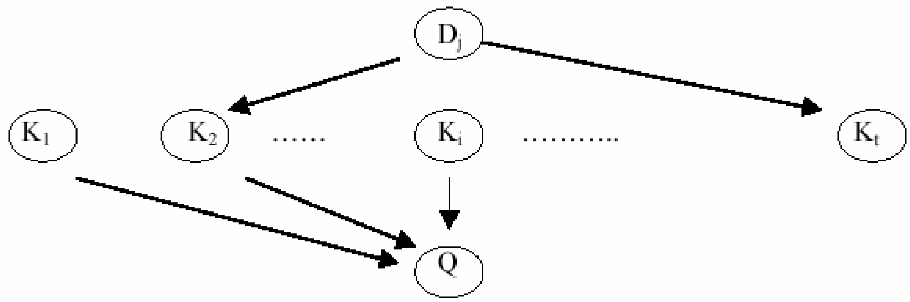
$$P(B) = P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + \ldots + P(A_n) \cdot P(B/A_n)$$

4. REDES DE INFERENCIA

4.1 Redes de Inferencia

Las redes de inferencia consideran la probabilidad como un grafo de creencia en la recuperación de información. Este modelo asocia variables aleatorias con palabras claves, documentos, y consultas. Una variable aleatoria asociada a un documento j representa el evento de observar ese documento. La observación del documento genera una creencia sobre la variable aleatoria de las palabras índices asociadas al documento.

La variable aleatoria de la consulta representa que la información requerida por la consulta ha sido satisfecha.



Todas las variables aleatorias son binarias.

Definición:

Sea $k = \{k1, k2, ..., kt\}$ con k_i siendo las variables aleatorias con valores 0 o 1. Estas variables definen los 2^t posibles estados de k. Además, d_j y q son también binarias variables aleatorias asociadas con el documento y consulta.

Que un documento sea relevante es determinado por la cantidad de apoyo evidencial de la observación d_j da la consulta q. Esto esta dado en redes de inferencia por $P(q \wedge d_j)$.

$$\begin{aligned} P(q \wedge d_j) &= \sum_{\forall \bar{k}} P(q \wedge d_j \mid \bar{k}) \times P(\bar{k}) \\ &= \sum_{\forall \bar{k}} P(q \wedge d_j \wedge \bar{k}) \\ &= \sum_{\forall \bar{k}} P(q \mid d_j \times \bar{k}) \times P(d_j \times \bar{k}) \\ &= \sum_{\forall \bar{k}} P(q \mid \bar{k}) \times P(\bar{k} \mid d_j) \times P(d_j) \end{aligned}$$

⇨ Redes de inferencia para modelo Booleano

$$P(d_j) = 1/N$$
$$\neg P(d_j) = 1 - P(d_j)$$
$$P(k_i \mid d_j) = \begin{cases} 1 \text{ if } g_i(d_j) = 1 \\ 0 \text{ otherwise} \end{cases}$$
$$P(\neg k_i \mid d_j) = 1 - P(k_i \mid d_j)$$
$$P(q \mid \overline{k}) = \begin{cases} 1 \text{ if } \exists \overline{q}_{cc} \mid (\overline{q}_{cc} \in \overline{q}_{dnf}) \wedge (\forall_{k_i}, g_i(\overline{k}) = g_i(\overline{q}_{cc}) \\ 0 \text{ otherwise} \end{cases}$$
$$P(\neg q \mid \overline{k}) = 1 - P(q \mid \overline{k})$$

➪ Redes de inferencia para modelo Vector

$$P(d_j) = 1/\mid \overline{d}_j \mid$$
$$\neg P(d_j) = 1 - P(d_j)$$
$$P(k_i \mid d_j) = f_{i,j} = \frac{freq_{i,j}}{\max_l freq_{l,q}}$$
$$P(\neg k_i \mid d_j) = 1 - P(k_i \mid d_j)$$
$$P(q \mid \overline{k}) = \begin{cases} \log \dfrac{N}{n_i} \text{ if } \overline{k} = \overline{k}_i \wedge g_i(\overline{q}) = 1 \\ 0 \quad \text{if } \overline{k} \neq \overline{k}_i \vee g_i(\overline{q}) = 0 \end{cases}$$
$$P(\neg q \mid \overline{k}) = 1 - P(q \mid \overline{k})$$

Una red de inferencia puede ser representada como un gráfico en el que los nodos representan parámetros que son los hechos obtenidos como datos o derivados de otros datos. Cada parámetro es una declaración acerca de algún aspecto del problema bajo análisis y puede servir como un antecedente o consecuente de una regla. Estas declaraciones pueden copiar un rango que va desde la conclusión final de un sistema, hasta hechos simples, observados o derivados. Cada uno de estos parámetros puede tener uno o más valores asociados, donde cada valor tiene una medida correspondiente de incertidumbre que representa cuan creíble es el valor particular de un parámetro.

Las reglas en el sistema están representadas dentro del gráfico por las interconexiones entre los varios nodos. Este conocimiento es utilizado por el proceso de inferencia para propagar resultados a través de la red.
Nótese que todas las interconexiones entre los varios nodos de la red de inferencia son conocidas previa a la ejecución del sistema. Esto trae como consecuencia la minimización del proceso de búsqueda de hechos que se identifiquen con las premisas. Adicionalmente, simplifican la implementación del mecanismo de inferencia y el manejo de las facilidades de explicación.
Las redes de inferencia son muy útiles para dominios donde el número de diferentes soluciones alternativas es limitado. Por ejemplo, la clasificación de elementos en las ciencias naturales y problemas de diagnóstico. Una red de inferencia es fácil de implementar, pero es menos poderosa ya que se debe conocer de antemano todas las relaciones entre reglas y hechos.
Sistemas comerciales de desarrollo, basados en esta arquitectura son los siguientes: Personal Consultant, EXSYS, y VP-Expert.

4.2 Inferencia Basada en Reglas

Una declaración de que algo es verdadero o es un hecho conocido, es una **afirmación** (*fact*). El conjunto de afirmaciones se conoce a menudo con el nombre de **memoria de trabajo o base de afirmaciones**. De igual forma, al conjunto de reglas se lo denomina **base de reglas**.

Un sistema basado en reglas utiliza el *modus ponens* para manipular las afirmaciones y las reglas durante el proceso de inferencia. Mediante técnicas de búsqueda y procesos de unificación, los sistemas basados en reglas automatizan sus métodos de razonamiento y proporcionan una progresión lógica desde los datos iniciales, hasta las conclusiones deseadas. Esta progresión hace que se vayan conociendo nuevos hechos o descubriendo nuevas afirmaciones, a medida que va guiando hacia la solución del problema.

En consecuencia, el proceso de solución de un problema en los sistemas basados en reglas va realizando una serie de inferencias que crean un sendero entre la definición del problema y su solución. Las inferencias están concatenadas y se las realiza en forma progresiva, por lo que se lo que se dice que el proceso de solución origina una **cadena de inferencias**.

Los sistemas basados en reglas difieren de la representación basada en lógica en las siguientes características principales:

- Son en general no-monotónicos, es decir hechos o afirmaciones derivadas, pueden ser retractados, en el momento en que dejen de ser verdaderos.
- Pueden aceptar incertidumbre en el proceso de razonamiento.

5. RAZONAMIENTO

El razonamiento determina un encadenamiento posible para las piezas de conocimiento.

5.1 Tipos de Razonamiento.

- ***Razonamiento Formal:*** Se deducen nuevas P.C siguiendo reglas de inferencia preespecificadas.
- ***Razonamiento Procedural:*** Utiliza la simulación para responder preguntas y resolver problemas.
- ***Razonamiento por analogía:*** Involucra asociar el problema a resolver con uno ya resuelto y utiliza el espacio de direcciones del segundo para resolver el primero.
- ***Meta Razonamiento:*** Involucra razonar sobre cuál es la mejor manera de razonar para un problema específico.

5.2 El Proceso de Razonamiento

El proceso de razonamiento en un sistema basado en reglas es una progresión desde un conjunto inicial de afirmaciones y reglas hacia una solución, respuesta o conclusión. Como se llega a obtener el resultado, sin embargo, puede variar significativamente:

- Se puede partir considerando todos los datos conocidos y luego ir progresivamente avanzando hacia la solución. Este proceso se lo denomina *guiado por los datos* o de **encadenamiento progresivo** (*forward chainning*).
- Se puede seleccionar una posible solución y tratar de probar su validez buscando evidencia que la apoye. Este proceso se denomina *guiado por el objetivo* o de **encadenamiento regresivo** (*backward chainning*).

5.3 Razonamiento Progresivo

En el caso del razonamiento progresivo, se empieza a partir de un conjunto de datos colectados a través de observación y se evoluciona hacia una conclusión. Se chequea cada una de las reglas para ver si los datos observados satisfacen las premisas de alguna de las reglas. Si una regla es satisfecha, es ejecutada derivando nuevos hechos que pueden ser utilizados por otras reglas para derivar hechos adicionales. Este proceso de chequear reglas para ver si pueden ser satisfechas se denomina *interpretación de reglas*.

La interpretación de reglas es realizada por una máquina de inferencia en un sistema basado en conocimiento. La interpretación de reglas, o inferencia, en el razonamiento progresivo involucra la repetición de los pasos que se indican en la siguiente figura.

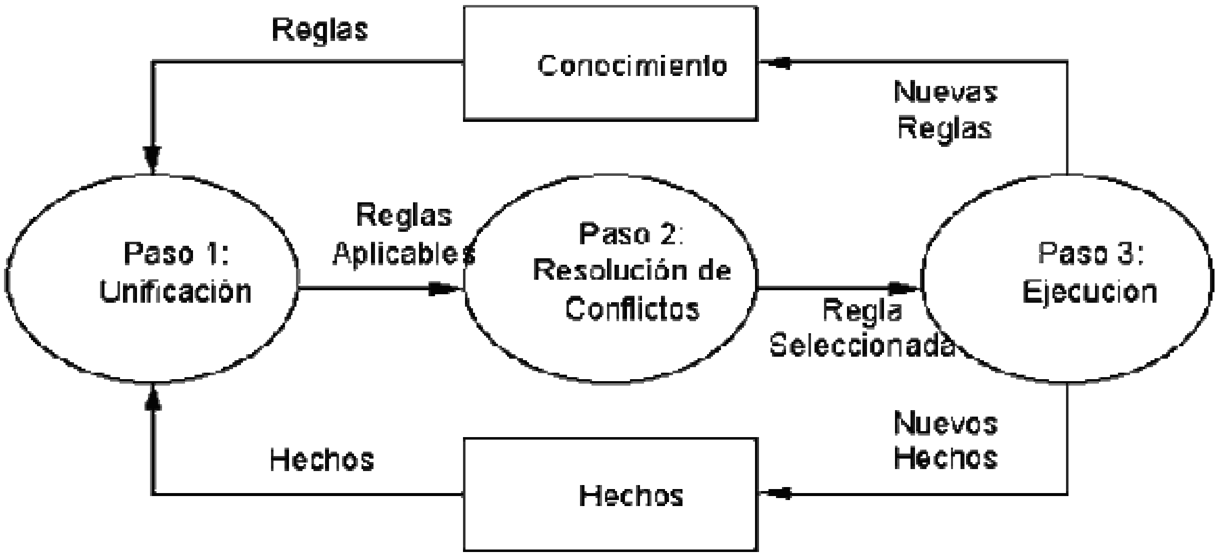


Figura 4.1 *Proceso de Razonamiento Progresivo*

Proceso de Razonamiento Progresivo

1. **Unificación (*Matching*).-** En este paso, en las reglas en la base de conocimientos se prueban los hechos conocidos al momento para ver cuáles son las que resulten satisfechas. Para decir que una regla ha sido satisfecha, se requiere que todas las premisas o antecedentes de la regla resuelvan a verdadero.
2. **Resolución de Conflictos.-** Es posible que en la fase de unificación resulten satisfechas varias reglas. La resolución de conflictos involucra la selección de la regla que tenga la más alta prioridad de entre el conjunto de reglas que han sido satisfechas.
3. **Ejecución.-** El último paso en la interpretación de reglas es la ejecución de la regla. La ejecución puede dar lugar a uno o dos resultados posibles: nuevo hecho (o hechos) pueden ser derivados y añadidos a la base de hechos, o una nueva regla (o reglas) pueden ser añadidas al conjunto de reglas (base de conocimiento) que el sistema considera para ejecución.

En esta forma, la ejecución de las reglas procede de una manera progresiva (hacia adelante) hacia los objetivos finales.

Un conjunto de aplicaciones adecuadas al razonamiento progresivo incluye supervisión y diagnóstico en sistemas de control de procesos en tiempo real, donde los datos están continuamente siendo adquiridos, modificados y actualizados. Estas aplicaciones tienen 2 importantes características:

1. Necesidad de respuesta rápida a los cambios en los datos de entrada.
2. Existencia de pocas relaciones predeterminadas entre los datos de entrada y las conclusiones derivadas.

Otro conjunto de aplicaciones adecuadas para el razonamiento progresivo está formado por: diseño, planeamiento y calendarización, donde ocurre la síntesis de nuevos hechos basados en las conclusiones de las reglas. En estas aplicaciones hay potencialmente muchas soluciones que pueden ser derivadas de los datos de entrada. Debido a que estas soluciones no pueden ser enumeradas, las reglas expresan conocimiento como patrones generales y las conexiones precisas entre estas reglas no pueden ser predeterminadas.

5.4 Razonamiento Regresivo

El mecanismo de inferencia, o interprete de reglas para el razonamiento regresivo, difiere significativamente del mecanismo de razonamiento progresivo. Si bien es cierto ambos procesos involucran el examen y aplicación de reglas, el razonamiento regresivo empieza con la conclusión deseada y decide si los hechos que existen pueden dar lugar a la obtención de un valor para esta conclusión. El razonamiento regresivo sigue un proceso muy similar a la búsqueda primero en profundidad.

El sistema empieza con un conjunto de hechos conocidos que típicamente está vacío. Se proporciona una lista ordenada de objetivos (o conclusiones), para las cuales el sistema trata de derivar valores. El proceso de razonamiento regresivo utiliza esta lista de objetivos para coordinar su búsqueda a través de las reglas de la base de conocimientos. Esta búsqueda consiste de los siguientes pasos:

1. Conformar una pila inicialmente compuesta por todos los objetivos prioritarios definidos en el sistema.
2. Considerar el primer objetivo de la pila. Determinar todas las reglas capaces de satisfacer este objetivo, es decir aquellas que mencionen al objetivo en su conclusión.
3. Para cada una de estas reglas examinar en turno sus antecedentes:
    a. Si todos los antecedentes de la regla son satisfechos (esto es, cada parámetro de la premisa tiene su valor especificado dentro de la base de datos), entonces ejecutar esta regla para derivar sus conclusiones. Debido a que se ha asignado un valor al objetivo actual, removerlo de la pila y retornar al paso (2).
    b. Si alguna premisa de la regla no puede ser satisfecha, buscar reglas que permitan derivar el valor especificado para el parámetro utilizado en esta premisa.
    c. Si en el paso (b) no se puede encontrar una regla para derivar el valor especificado para el parámetro actual, entonces preguntar al usuario por dicho valor y añadirlo a la base de datos. Si este valor satisface la premisa actual entonces continuar con la siguiente premisa de la regla. Si la premisa no es satisfecha, considerar la siguiente regla.

Si todas las reglas que pueden satisfacer el objetivo actual se han probado y todas no han podido derivar un valor, entonces este objetivo quedará indeterminado. Removerlo de la pila y retornar al paso (2). Si la pila está vacía parar y anunciar que se ha terminado el proceso.

El razonamiento regresivo es mucho más adecuado para aplicaciones que tienen mucho mayor número de entradas, que de soluciones posibles. La habilidad de la lógica regresiva para trazar desde las pocas conclusiones hacia las múltiples entradas la hace más eficiente que el encadenamiento progresivo. Una excelente aplicación para el razonamiento regresivo es el diagnóstico, donde el usuario dialoga directamente con el sistema basado en conocimiento y proporciona los datos a través del teclado. Problemas de clasificación también son adecuados para ser resuelto mediante el razonamiento regresivo.

6. Ejemplo de un sistema de razonamiento probabilista.

DAACS – un sistema de razonamiento probabilista

- *Dump (descarte) Analysis And Consulting System* (Daacs) analiza en tiempo real y resuelve errores en la reserva de pasajes del paquete Sabre de American Airlines.
    - "Dump Analysis" es el análisis sobre procesos de cómputo con errores fatales.
- El tema general de esta especialidad de la IA consiste en reconocer patrones de conducta que permiten el desbichado de programas.
- Daacs integra
    - el diseño tradicional del compilador,
    - las técnicas basadas en heurísticas y lógicas para la comprensión del programa y
    - el razonamiento probabilista para inferencias de diagnóstico.
- Toma una instantánea de los errores fatales del programa y coteja la información del caso fracasado con una red de creencia. El razonador controla y "desbicha" las inferencias resultantes y trabaja con datos incompletos e inciertos para producir una lista ordenada de diagnósticos. Cuando no puede establecer la causa verdadera con certeza, enumera una serie de hipótesis acerca de la fuente de la falla y proporciona un diagnóstico plausible y la evidencia que utiliza . Cuando le resulta posible,

Daacs determina la causa exacta de un error. Así el uso de redes de creencia permite que el diagnóstico continúe más allá de lo que es posible probar con certeza.