

Un Modelo de Desacumulación de Evidencia para Sistemas Expertos

Ramón García Martínez - Antonio Marsiglio



PROGRAMA DE INVESTIGACION Y DESARROLLO EN INTELIGENCIA ARTIFICIAL
UNIVERSIDAD NACIONAL DE LUJAN

Resumen : Se presenta una fórmula de desacumulación de evidencia que no requiere el archivo de todas las evidencias asociadas a la evidencia acumulada resultante. Se prueba estadísticamente para una cantidad fija de evidencias que el error al utilizar la fórmula propuesta respecto de la fórmula clásica está acotado.

Abstract : A formula for evidence desacumulation that doesn't need the record of the associated evidences to the resulting accumulated evidence is presented in this paper. It is statistically proved for a fixed number of evidences that the error between the proposed formula and the clasical formula is bounded.

1. Introducción :

Es frecuente en la tecnología de Sistemas Expertos (Necher,1985; Nevins,1985; Waterman,1983) ponderar reglas o hechos sobre el dominio de experticia del Sistema Experto mediante el uso de evidencias

Por ejemplo :

SI	Hecho 1
Y	Hecho 2
ENTONCES	Hecho 3 EV 0.9

La semántica asociada es la siguiente : Se sabe con evidencia (0.9) que de la ocurrencia del Hecho 1 y del Hecho 2 , puede concluirse el Hecho 3 . Cuando los Hecho 1 , Hecho 2 y Hecho 3 no son bivaluados (Admiten un valor en el conjunto {verdadero,falso}) sino que admiten evidencia (un valor numérico en el intervalo (0,1)) se recurre a un modelo de propagación de evidencias . Uno de los más utilizados es el plausibilístico que aplicado al ejemplo resulta :

$EV(\text{Hecho 3}) = \text{Mínimo}(EV(\text{Hecho 1}) , EV(\text{Hecho 2})) \times EV(\text{Regla})$

Si $EV(\text{Hecho } 1) = 0.6$ y $EV(\text{Hecho } 2) = 0.8$ se obtiene que $EV(\text{Hecho } 3) = 0.54$

Puede ocurrir que el Hecho 3 ya tenga una evidencia asociada, en este caso se recurre a un modelo de acumulación de evidencia. El plausiblestístico aplicado al ejemplo resulta:

$$EV_r(\text{Hecho } 3) = EV_v(\text{Hecho } 3) + (1 - EV_v(\text{Hecho } 3)) \times EV_n(\text{Hecho } 3)$$

Donde	$EV_r(\text{Hecho } 3)$	evidencia Resultante del Hecho 3
	$EV_v(\text{Hecho } 3)$	evidencia Vieja del Hecho 3
	$EV_n(\text{Hecho } 3)$	evidencia Nueva del Hecho 3

Por ejemplo si: $EV_v(\text{Hecho } 3) = 0.5$

$EV_n(\text{Hecho } 3) = 0.6$

Resulta que: $EV_r(\text{Hecho } 3) = 0.5 + (1 - 0.5) \times 0.6 = 0.8$

Puede definirse recursivamente una función F de acumulación general para N evidencias, la cual tendrá la siguiente forma:

$$EV_1 \quad \text{si } N = 1$$

$$F(EV_1, \dots, EV_N) = < \quad (1)$$

$$F(EV_1, \dots, EV_{N-1}) + (1 - F(EV_1, \dots, EV_{N-1})) \cdot EV_N$$

si $N > 1$

La limitación del modelo plausiblestístico radica en su monotonicidad (Quinlan, 1979), es decir, un hecho nunca va a tener menos evidencia de la que ya haya acumulado. En Sistemas Expertos que deben manejar Hechos que modelizan Hipótesis, el modelo plausiblestístico presentado no permite el crecimiento y decrecimiento dinámico de las evidencias asociadas a los Hechos. La forma clásica (Doyle, 1979; Shastri, 1985) de abordar este problema es la siguiente:

Dado un Hecho con las evidencias asociadas $EV_1, EV_2, \dots, EV_N$, podemos definir la evidencia resultante EV_r :

$$EV_r = F(EV_1, \dots, EV_N)$$

Si tenemos la evidencia que se opone EV_T se puede calcular en forma exacta la evidencia resultante de considerar que EV_T nunca existió (situación que idealiza el hecho que la evidencia primero se acumuló y luego se desacumuló), aplicando (Driankov, 1986; Prade, 1983; Yagger, 1986) la función de acumulación general F definida en (1) de la siguiente manera:

$$EV_{r1} = F(EV_1, \dots, EV_{T-1}, EV_{T+1}, \dots, EV_N) \quad (2)$$

Esta forma de recalcular la evidencia resultante de un hecho después que se desacumuló una evidencia previamente acumulada presenta las siguientes desventajas :

- [1] Por cada Hecho de la Base de Hechos debe guardarse cada una de las evidencias que se utilizaron para calcular su evidencia resultante .
- [2] Cada evidencia desacumulada obliga a recalcular el valor de la función de acumulación general para el grupo de evidencias asociadas al hecho al cual se le desacumuló la evidencia.

2. Objetivo del Trabajo :

En este trabajo se propone la siguiente fórmula de desacumulación (Jeffrey, 1983; Prade, 1986; Marsiglio, 1990) :

$$EV_{r2} = EV_v - (\text{signo}(EV_v) - EV_v) \times \text{signo}(EV_v) \times EV_T - K \quad (3)$$

Donde EV_v es la evidencia Vieja asociada al hecho.

Observar que esta forma de calcular la evidencia resultante tiene las siguientes ventajas respecto de (2) :

- [1] Se ahorra memoria ya que solo se necesita guardar el valor de EV_v asociada al hecho.
- [2] Se ahorra tiempo de cómputo ya que no se necesita recalcular la función de acumulación por cada evidencia que es desacumulada .

A sido un objetivo del trabajo verificar la siguiente hipótesis :

El error medio cometido al calcular la evidencia resultante EV_{r2} mediante la fórmula de desacumulación propuesta en (3) respecto de la forma exacta de calcular la evidencia resultante EV_{r1} dada por (2) , es menor que un cierto ϵ .

Donde ϵ es un error despreciable a los efectos del tratamiento de la evidencia , en el cálculo de evidencia se trabaja con dos decimales, por lo tanto se requiere que $\epsilon < 0.01$.

3. Supuestos Teóricos :

Para la experimentación (Cappelletti, 1983) se supuso :

- [1] Que EV_{r1} y EV_{r2} son variables aleatorias continuas apareadas.
- [2] Que EV_{r1} y EV_{r2} calculados para un grupo concreto de evidencias conforman un conjunto de pares homólogos de la forma (EV_{r1i}, EV_{r2i}) .
- [3] Que EV_{r1} tiene distribución $N(\mu_1, \sigma^2)$
- [4] Que EV_{r2} tiene distribución $N(\mu_2, \sigma^2)$
- [5] Que μ_1 , μ_2 y σ^2 son desconocidas.
- [6] Que se puede construir la variable aleatoria continua :
$$D_i = EV_{r1i} - EV_{r2i}$$

Que tiene media $\mu_D = \mu_1 - \mu_2$
- [7] Que la hipótesis puede reformularse de la siguiente manera :
El intervalo $[- \epsilon, \epsilon]$ contiene al μ_D

4. Metodología :

Para evaluar la hipótesis se siguieron los siguientes pasos :

- 1) Se fijó el número de evidencias a tratar por ensayo en 5 y el tamaño de la muestra en 500 ensayos.
- 2) Se construyó un programa que genera la cantidad de ensayos que especifica el experimentador. Cada ensayo consta de :
 - a) La generación aleatoria de 5 evidencias, sean EV_1, EV_2, EV_3, EV_4 y EV_5 .
 - b) La generación aleatoria de la posición de la evidencia que será desacumulada, sea T .
 - c) El cálculo de de la Evidencia Resultante EV_{r1} calculada por el método clásico dada por :
$$EV_{r1} = F(EV_1, \dots, EV_{T-1}, EV_{T+1}, \dots, EV_5)$$
 - d) El cálculo de de la Evidencia Resultante EV_{r2} calculada por el método propuesto dada por :
$$EV_{r2} = EV_v - (1 - EV_v) \times EV_T - K$$

Donde :

EV_V Evidencia Vieja

EV_T Evidencia a desacumular

A los fines de cada Experimento :

$$EV_T = F(EV_1, EV_2, EV_3, EV_4, EV_5)$$

- 3) Se construyó un intervalo de confianza para la media de $|EV_{r1} - EV_{r2}|$
- 4) Se verificó que el intervalo de confianza construido estuviera contenido en el intervalo $[-\epsilon, \epsilon]$.

5. Resultados :

Los parametros estadísticos obtenidos a partir del experimento son los siguientes :

Media Aritmética de la Diferencia : -0.013

Desviación Stándar de la Diferencia : .0509

Con un coeficiente de confianza 0.995 , μ_D es un valor en el intervalo
[-0.008 , 0.004]

Dado que el intervalo [-0.01 , 0.01] contiene al intervalo [-0.008 , 0.004] entonces puede afirmarse que el error promedio al calcular la evidencia resultante mediante la fórmula de desacumulación propuesta en (3), difiere de la forma exacta propuesta en (2) , en menos de 0.01 .

Complementariamente se calculó el coeficiente de correlación entre EV_{r1} y EV_{r2} para una muestra de 1000 ensayos, el valor obtenido fue :

Coeficiente de Correlación : .9133

De lo que se deduce que la fórmula propuesta (EV_{r2}) tiene un alto grado de asociación con la fórmula clásica (EV_{r1}).

6. Conclusiones :

En el presente trabajo se ha presentado una fórmula de desacumulación de evidencia que presenta la ventaja de no requerir el archivo de todas las evidencias asociadas a la evidencia acumulada resultante (Zadeh, 1979) . Se ha probado para una cantidad fija de evidencias que el error esperado es despreciable y que puede asumirse si ello involucra economía de memoria y disminución de los tiempos de cálculo . Es un problema abierto el verificar que el error de la fórmula sigue estando dentro de los límites para una cantidad variable de evidencias .

7. Bibliografía :

- CAPPELLETTI, C. 1983. Elementos de Estadística. 2ª edición Argentina. Cesarini Hnos Editores. 368 p.
- DOYLE, J. 1979. A Truth Maintenance System. Artificial Intelligence Journal. (E.U.A.). 12(3) : 231-272.
- DRIANKOV, D. 1986. Uncertainty calculus with verbally defined belief intervals. International Journal of Intelligent Systems. (E.U.A.). 1(4): 219-246.
- JEFFREY, R. 1983. The logic of decision. 1ª edición. E.U.A. University of Chicago Press 231 p.
- MARSIGLIO, A. 1990. Una Fórmula de Acumulación y Desacumulación de Evidencias. En Anales del Primer Simposio en Inteligencia Artificial y Robótica. Argentina. pp. 205-207.
- NECHER, R. 1985. Explainable and maintainable expert systems. En Proceedings of the Ninth International Joint Conference on Artificial Intelligence. (E.U.A.). pp. 382-389.
- NEVINS, A. J. 1985. An architecture for knowledge based deduction. En Proceedings of the Ninth International Joint Conference on Artificial Intelligence. (E.U.A.). pp. 338-339
- PRADE, H. 1983. A synthetic view of approximate reasoning techniques. En Proceedings of the Eighth International Joint Conference on Artificial Intelligence. (R.F.A.). pp. 130-136
- , 1986. On the unicity of Dempster rule of combination. International Journal of Intelligent Systems. (E.U.A.). 1(2) : 133-142.
- QUINLAN, J. 1979. Consistency and plausible reasoning. In Proceedings of the Eighth International Joint Conference on Artificial Intelligence. (R.F.A.). pp. 137-144
- SHASTRI, L. 1985. Evidential reasoning in semantic networks : a formal theory. En Proceedings of the Ninth International Joint Conference on Artificial Intelligence. (E.U.A.). pp. 465-474
- WATERMAN, F. 1983. Building Expert Systems. 1ª edición. E.U.A. A.Wesley. 345 p.
- YAGER, R. 1986. Toward a general theory of reasoning and uncertainty : nonspecificity and fuzziness. International Journal of Intelligent Systems.(E.U.A.).1(1): 45-68.
- ZADEH, L. 1979. Approximate reasoning based on fuzzy logic. En Proceedings of the Seventh International Joint Conference on Artificial Intelligence. (Japón). pp. 1004-1010.

8. Contacto :

Ramón García Martínez
Erézcano 1134
Adrogué (1846)
Buenos Aires
ARGENTINA