Lógica probabilística y la deducción *anytime*

Dalia Camacho García-Formentí

November 2019

1. Introducción

El concepto de probabilidad es muy utilizado, por lo tanto se ha buscado abordarlo desde la lógica. Existen distintas formas de definir una lógica que contemple la idea de probabilidad. Una aproximación puede ser considerar una lógica en que los valores de verdad se encuentren en el intervalo $[0, 1]$ y que sean tratados como probabilidades, esto es lo que propone Lukasiewicz [1].

Por otro lado Nilsson [2] considera que la probabilidad de un enunciado está dada a partir de la suma de la probabilidad de los mundos posibles donde el enunciado es verdadero. Además, Nilsson propone un método para encontrar la probabilidad de un nuevo enunciado a partir de enunciados sobre los que se conoce previamente su probabilidad.

Sin embargo, el método propuesto por Nilsson es computacionalmente ineficiente, de hecho es un problema NP-completo [3]. Ante esta situación Frisch y Haddawy proponen el método de deducción *anytime* [3]. Ellos definen el lenguaje $\mathcal{L}_{PL}$ con el cual es posible expresar intervalos de probabilidad para enunciados de lógica proposicional o de lógica de primer orden. Con el método *anytime* es posible acotar la probabilidad de un enunciado nuevo dado un conjunto de enunciados sobre los que se conoce su intervalo de probabilidad. Además, es posible obtener resultados intermedios en cualquier momento.

En este trabajo abordaremos de forma general la propuesta de Nilsson y más a detalle la deducción *anytime* definida por Frisch y Haddawy.

2. La probabilidad en términos de mundos posibles

Nilsson propone considerar la probabilidad de un enunciado en términos de los mundos posibles en que el enunciado es verdadero [2]. Para un solo enunciado se tienen dos mundos posibles uno en que el enunciado es verdadero y otro en que el enunciado es falso. En este caso si la probabilidad de cada uno de los dos mundos es idéntica la probabilidad del enunciado será $\frac{1}{2}$. Dentro de cada mundo posible se siguen las reglas de lógica proposicional y no puede haber enunciados contradictorios. Para n enunciados hay a lo más 2^n mundos posibles, sin embargo es usual que sean menos, ya que en algunos casos una combinación de enunciados puede generar inconsitencias y los mundos en que se producen inconsistencias no son considerados.

Para encontrar los mundos consistentes a partir de un conjunto de enunciados se puede utilizar un árbol binario en que al agregar un nuevo enunciado todas las ramas se bifurcan y se añade el enunciado como verdadero en una hoja y como falso en la otra. Si en un mundo hay inconsistencias, entonces la rama se cierra y no puede formar un mundo posible. Las ramas que permanecen abiertas después de haber añadido todos los enunciados corresponden a los mundos posibles. Nilsson hace un ejemplo con los enunciados P, $P \to Q$ y Q con lo que queda el árbol de la Figura 1.

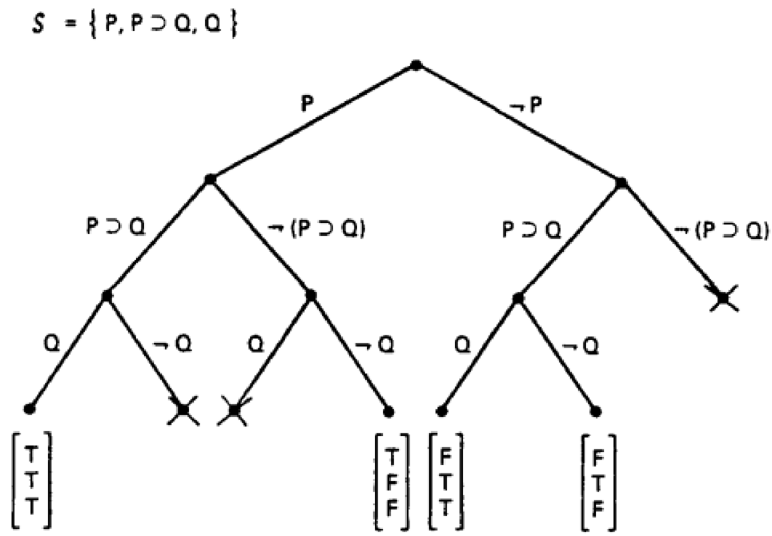

Figura 1: Árbol binario para P, $P \to Q$ y Q tomado de [2].

Dados los mundos posibles, si la probabilidad de cada mundo es conocida, entonces la probabilidad de un enunciado es simplemente la suma de las probabilidades de los mundos donde el enunciado es verdadero. Para el ejemplo de la Figura 1 si los mundos posibles son equiprobables la probabilidad de $\mathbf{P}(P) = \mathbf{P}(Q) = 0{,}5$ y la de $\mathbf{P}(P \to Q) = 0{,}75$.

En la siguiente sección se explica el método propuesto por Nilsson para asignarle probabilidad a un nuevo enunciado, al que le llama *entailment* probabilístico.

3. *Entailment* probabilístico

En esta sección se describe el método de *entailment* probabilístico. Para esto es necesario definir una forma de calcular las probabilidades de los enunciados.

Dado un ordenamiento arbitrario de los enunciados $S_1, S_2, \ldots, S_n$ y un ordenamiento de los mundos $W_1, W_2 \ldots W_l$, es posible calcular la probabilidad de los distintos enunciados de la siguiente manera:

$$\Pi = VP, \tag{1}$$

donde $\Pi \in \mathbb{R}^{n\times 1}$ corresponde al vector que contiene la probabilidad de cada enunciado de acuerdo al ordenamiento definido. $V \in \{0,1\}^{n\times l}$ corresponde a una matriz cuyas filas representan a los enunciados y las columnas a los mundos, si la entrada $V_{ij} = 0$ entonces el enunciado i es falso en el mundo j; por el contrario si $V_{ij} = 1$ entonces el enunciado i es verdadero en el mundo j. Finalmente, $P \in \mathbb{R}^{l\times 1}$ indica la probabilidad de cada uno de los mundos en el orden correspondiente.

Los valores de probabilidad que pueden tomar los enunciados deben estar restringidos para mantener la consistencia de los mundos posibles. Estas restricciones se pueden encontrar a partir de la envoltura convexa que se obtiene de los puntos en que la probabilidad de un mundo es 1 y la probabilidad del resto de los mundos es 0. Los valores de probabilidad que pueden tomar los enunciados de forma conjunta son únicamente aquellos que se encuentran en la envoltura convexa.

En el ejemplo con P, $P \to Q$ y Q se tienen 4 mundos posibles. Cuando cada uno de estos mundos tienen probabilidad 1 se obtienen los siguientes puntos en $\mathbb{R}^3$: $(1,1,1)$, $(1,0,0)$, $(0,1,1)$ y $(0,1,0)$, donde la primer coordenada indica la probabilidad de P, la segunda la de $P \to Q$ y la tercera la de Q. Dados estos puntos se obtiene su envoltura convexa que se muestra en la Figura 2.

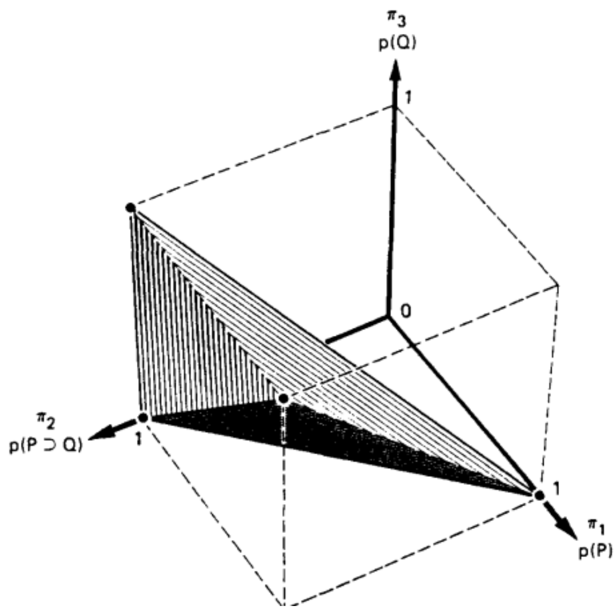

Figura 2: Envoltura convexa de $(1, 1, 1)$, $(1, 0, 0)$, $(0, 1, 1)$ y $(0, 1, 0)$ tomada de [2].

Conocer la envoltura convexa hace posible a grupos de expertos en un área asignar probabilidades de forma consistente. Haciendo uso de la envoltura convexa se vuelve más fácil que los expertos den un estimado de las probabilidades de los enunciados que a las probabilidades de los mundos. Esto se contempla al momento de definir el problema de *entailment* probabilístico.

El *entailment* probabilístico consiste en asignar una probabilidad a un nuevo enunciado dadas las probabilidades de un conjunto de enunciados conocidas previamente.

Para resolver el problema se hace lo siguiente:

1. Encontrar los mundos posibles con asignaciones de verdad consistentes para el conjunto que contiene los n enunciados cuya probabilidad se conoce y el nuevo enunciado.

2. Encontrar la probabilidad de cada mundo posible a partir del sistema $\Pi' = V'P$. Con $V' \in \{0, 1\}^{n+1 \times l}$ la matriz cuya primer fila representa la restricción de que la suma de la probabilidad de los mundos debe ser uno. Para las n filas restantes $V'_{ij} = 0$ si el enunciado $i - 1$ es falso en el mundo j y si $V'_{ij} = 1$ el enunciado $i - 1$ es verdadero en el mundo j, esto para los enunciados conocidos previamente. $\Pi' \in \mathbb{R}^{n+1}$ contiene en la primera entrada un uno que de igual manera corresponde a la restricción sobre la suma de las probabilidades de los mundos; y las demás entradas contienen la probabilidad ya conocida.

3. Se define un vector $S \in \{0,1\}^{1\times l}$ donde $S_i = 0$ si el nuevo enunciado es falso en el mundo i y $S_i = 1$ si el nuevo enunciado es verdadero en el mundo i.

4. Para encontrar la probabilidad del nuevo enunciado se realiza la multiplicación SP.

Generalmente $n < l$, entonces no hay una solución única de P para el sistema $\Pi' = V'P$, por lo que se puede obtener el vector P con máxima entropía que cumpla con las restricciones y se puede resolver por ejemplo con multiplicadores de Lagrange. La entropía de P se define como $H = -P^t log(P)$.

Nilsson también considera el caso de estimar la probabilidad condicional de dos enunciados a partir de la regla de Bayes $\mathbf{P}(A|B) = \frac{\mathbf{P}(A\wedge B)}{\mathbf{P}(B)}$. A partir del método de *entailment* probabilístico obtiene por un lado $\mathbf{P}(A \wedge B)$ y por otro $\mathbf{P}(B)$ y simplemente realiza la división correspondiente.

4. Problemas principales del método de Nilsson

El método de *entailment* probabilístico de Nilsson no es computacionalmente eficiente. Para realizar el primer paso, que es encontrar todos los mundos posibles consistentes, en el peor de los casos se necesita evaluar la consistencia de 2^n mundos posibles con n enunciados en cada mundo. Más aún, es un problema NP-completo [3], por lo cual ni siquiera el primer paso se puede realizar en tiempo razonable. Otro problema que Frisch y Haddawy encuentran es que si el proceso se detiene en un paso intermedio no es posible tener mayor conocimiento sobre la probabilidad del nuevo enunciado de la que se tenía en un principio. Para resolver estos problemas Frisch y Haddawy plantean el método de deducción *anytime*. Aunado a ello, permiten cierta flexibilidad sobre las probabilidades de los enunciados al definirlas dentro de un intervalo y no de manera puntual. En la siguiente sección se define el lenguaje $\mathcal{L}_{PL}$ propuesto por Frisch y Haddawy para lógica probabilística y posteriormente se describe el método *anytime* y sus reglas de deducción.

5. El lenguaje $\mathcal{L}_{PL}$

El lenguaje $\mathcal{L}_{PL}$ a diferencia de lo realizado por Nilsson permite definir la probabilidad de enunciados de lógica proposicional dentro de un intervalo (cerrado) y considera la probabilidad condicional de forma más sencilla y directa.

La lógica definida por el lenguaje $\mathcal{L}_{PL}$ es una lógica bivalente, la probabilidad subyace en los enunciados que pertenecen a $\mathcal{L}_{PL}$. Estos son de la forma $\mathbf{P}(\phi|\epsilon) \in I$, donde ϕ y ϵ son elementos de $\mathcal{L}_{pro}$ e I es un intervalo cerrado cuyos elementos están entre 0 y 1. En general los enunciados de $\mathcal{L}_{PL}$ son condicionales y para definir un enunciado no condicional es posible hacerlo como $\mathbf{P}(\phi|T) \in I$, donde T es una tautología. Para facilitar su escritura simplemente se puede escribir como $\mathbf{P}(\phi) \in I$.

Las proposiciones atómicas de $\mathcal{L}_{PL}$ se denotan con letras mayúsculas, aunque en particular las letras I y J están reservadas para los intervalos de probabilidad; las letras minúsculas se utilizan para denotar los límites de los intervalos de probabilidad; las letras griegas minúsculas hacen referencia a enunciados de $\mathcal{L}_{pro}$. El símbolo ψ corresponde a un enunciado de $\mathcal{L}_{PL}$ y Ψ a un conjunto de enunciados de $\mathcal{L}_{PL}$.

Dada $\mathbf{P}(\phi|\epsilon) \in I$ a I se le conoce como **componente probabilístico** y a $\mathbf{P}(\phi|\epsilon)$ como **componente lógico**.

También se define una relación con la idea de Nilsson de mundos posibles. Para un conjunto M de mundos posibles, se considera que M es un modelo y $\mathbf{P}(\phi)$ bajo ese modelo está igualmente dada por la suma de las probabilidades de los mundos en que ϕ es verdadero. A la probabilidad de ϕ dado ϵ bajo algún modelo M se le denota como $[\mathbf{P}(\phi|\epsilon)]^M$ y los valores de verdad esta expresión se definen como sigue:

$$[H][\mathbf{P}(\phi|\epsilon)]^M = \begin{cases} Verdadero, & si \quad [\mathbf{P}(\epsilon)]^M = 0 \\ Verdadero, & si \quad [\mathbf{P}(\epsilon)]^M > 0y[\mathbf{P}(\phi \wedge \epsilon)]^M/[\mathbf{P}(\epsilon)]^M \in I \\ Falso, & e.o.c. \end{cases} \quad (2)$$

El primer caso es para tratar la indefinición del condicional cuando se divide entre cero. No se utiliza el valor indefinido para que $\mathcal{L}_{PL}$ sea una lógica bivalente. Además, si $\mathbf{P}(\epsilon) = 0$ se tiene que $\mathbf{P}(\phi|\epsilon) \in I$ es verdad para cualquier intervalo, en particular $\mathbf{P}(\phi|\epsilon) \in \emptyset$ es verdadero. Por lo tanto $\mathbf{P}(\phi|\epsilon) \in \emptyset$ y $\mathbf{P}(\epsilon) = 0$ son enunciados equivalentes.

En el caso de $\mathcal{L}_{PL}$ el concepto de infinito juega un papel importante. Por ejemplo, a partir de $\{\mathbf{P}(A) \in \left[0{,}5 - \frac{1}{n},\ 0{,}5 + \frac{1}{n}\right],\ \forall n \geq 2\}$ se puede deducir que $\mathbf{P}(A) = [0{,}5, 0{,}5]$, sin embargo si se considera un subconjunto finito de $\{\mathbf{P}(A) \in \left[0{,}5 - \frac{1}{n},\ 0{,}5 + \frac{1}{n}\right],\ \forall\, n \geq 2\}$ solamente se puede decir que $\mathbf{P}(A) = \left[0{,}5 - \frac{1}{k}, 0{,}5 + \frac{1}{k}\right]$, donde $k = \text{máx}\{n_i | \mathbf{P}(A) \in \left[0{,}5 - \frac{1}{n_i},\ 0{,}5 + \frac{1}{n_i}\right] \text{fue seleccionado}\}$.

Habiendo definido $\mathcal{L}_{PL}$ es posible definir el método de deducción *anytime* con el cuál se va acotando el intervalo de probabilidad de un nuevo enunciado a partir de elementos que se saben verdaderos en $\mathcal{L}_{PL}$.

6. Deducción *anytime* y sus reglas de inferencia

El problema de *entailment* probabilístico puede ser denotado por $(\Psi, (\phi|\epsilon))$, donde Ψ es un conjunto de enunciados de $\mathcal{L}_{PL}$ cuyos valores de verdad son verdaderos y $\mathbf{P}(\phi|\epsilon)$ es la probabilidad objetivo para la cual se busca encontrar el intervalo más pequeño en donde $\mathbf{P}(\phi|\epsilon) \in I$ sea verdadero, a esto se le conoce como *entailment* estrecho. Si se deduce que $\mathbf{P}(\epsilon) = [0,\ 0]$, entonces $\mathbf{P}(\phi|\epsilon) \in \emptyset$.

Además, se tiene que para $(\phi|\epsilon)$ el *entailment* estrecho que se deduce de Ψ es un intervalo cerrado. Sin embargo, no necesariamente se deduce un intervalo puntual a partir de enunciados con intervalos de probabilidad puntual.

Se define como completo para *entailment* estrecho a un sistema tal que para cualquier problema de *entailment* probabilístico es posible deducir el intervalo I de *entailment* estrecho.

El método *anytime* da una solución aproximada al problema de *entailment* estrecho y tiene dos propiedades importantes las cuales son **parcialidad** y **exactitud**. La parcialidad se refiere a que la información que se tiene sobre el problema incrementa de forma monotónica creciente y existe un momento en que se tiene información parcial del resultado. La exactitud indica que la información que se tiene en cualquier momento es correcta.

Se dice que un método de deducción *anytime* es **convergente** si para un intervalo I y para cualquier intervalo J que contenga a I se cumple que a partir de un momento en el tiempo cualquier conjunto que sea derivado será subconjunto de todo conjunto J. El proceso es **convergente en tiempo finito** si para algún momento en el tiempo el intervalo encontrado ya no se modifica.

Para llevar a cabo la deducción *anytime* los autores proponen 38 reglas. En las primeras cuatro los enunciados que se utilizan están condicionadas por el mismo enunciado de $\mathcal{L}_{pro}$, por lo que se pueden reducir al caso no condicional, donde las primeras tres se deducen de $\mathbf{P}(\alpha \vee \beta) = \mathbf{P}(\alpha) + \mathbf{P}(\beta) - \mathbf{P}(\alpha \wedge \beta)$ y la cuarta a la negación. Estas se encuentran en la Figura 3.

(i)

$$\frac{\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \quad \boldsymbol{P}(\alpha \vee \beta \mid \delta) \in [u\ v] \quad \boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [w\ z]}{\boldsymbol{P}(\beta \mid \delta) \in [\max(w, u-y+w)\ \ \min(v, v-x+z))]}$$

provided $w \leqslant y,\ x \leqslant v,\ w \leqslant v$

(ii)

$$\frac{\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \quad \boldsymbol{P}(\beta \mid \delta) \in [u\ v] \quad \boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [w\ z]}{\boldsymbol{P}(\alpha \vee \beta \mid \delta) \in [\max(x, u, w, x+u-z)\ \ \lfloor y+v-w \rfloor_1]}$$

provided $w \leqslant y,\ w \leqslant v,\ z \geqslant x+u-1$

(iii)

$$\frac{\boldsymbol{P}(\beta \mid \delta) \in [x\ y] \quad \boldsymbol{P}(\alpha \mid \delta) \in [u\ v] \quad \boldsymbol{P}(\alpha \vee \beta \mid \delta) \in [w\ z]}{\boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [\lceil x+u-z \rceil^0\ \ \min(y, v, z, y+v-w)]}$$

provided $z \geqslant x,\ z \geqslant u$

(iv)

$$\frac{\boldsymbol{P}(\alpha \mid \delta) \in [x\ y]}{\boldsymbol{P}(\neg\alpha \mid \delta) \in [1-y\ \ 1-x]}$$

Figura 3: Reglas i a iv, Figura de [3].

Las reglas v a la ix se obtienen a partir de la definición de probabilidad condicional $\mathbf{P}(\alpha|\beta) = \frac{\mathbf{P}(\alpha \wedge \beta)}{\mathbf{P}(\beta)}$ y se encuentran en la Figura 4.

(v)

$$\frac{\begin{array}{l} \boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\beta \mid \delta) \in [u\ v] \end{array}}{\boldsymbol{P}(\alpha \mid \beta \wedge \delta) \in [x/v\ z]}$$

where $z = \begin{cases} 1, & \text{if } y > u \\ 0, & \text{if } y = u = 0 \\ y/u, & \text{otherwise} \end{cases}$

provided $x \leqslant v,\ v > 0$.

(vi)

$$\frac{\begin{array}{l} \boldsymbol{P}(\beta \mid \alpha \wedge \delta) \in [x\ y] \\ \boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [u\ v] \end{array}}{\boldsymbol{P}(\alpha \mid \delta) \in [u/y\ \lfloor v/x \rfloor_1]}$$

provided $y \geqslant u,\ x > 0,\ y > 0$

(vii)

$$\frac{\begin{array}{l} \boldsymbol{P}(\beta \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\alpha \mid \beta \wedge \delta) \in [u\ v] \end{array}}{\boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [x \cdot u\ y \cdot v]}$$

(viii)

$$\frac{\begin{array}{l} \boldsymbol{P}(\alpha_k \mid \alpha_1 \wedge \delta) \in [w\ z] \\ \boldsymbol{P}(\alpha_i \mid \alpha_{i+1} \wedge \delta) \in [x_i\ y_i],\ 1 \leqslant i \leqslant k-1 \\ \boldsymbol{P}(\alpha_{i+1} \mid \alpha_i \wedge \delta) \in [u_i\ v_i],\ 1 \leqslant i \leqslant k-1 \end{array}}{\boldsymbol{P}(\alpha_1 | \alpha_k \wedge \delta) \in [(w \cdot \prod_{i=1}^{k-1} \frac{x_i}{v_i})\ (z \cdot \prod_{i=1}^{k-1} \frac{y_i}{u_i})]}$$

(ix)

$$\frac{\begin{array}{l} \boldsymbol{P}(\alpha \mid \beta \wedge \delta) \in [x_1\ y_1] \\ \boldsymbol{P}(\beta \mid \alpha \wedge \delta) \in [x_2\ y_2] \\ \boldsymbol{P}(\gamma \mid \beta \wedge \delta) \in [u_1\ v_1] \\ \boldsymbol{P}(\beta \mid \gamma \wedge \delta) \in [u_2\ v_2] \end{array}}{\boldsymbol{P}(\gamma \mid \alpha \wedge \delta) \in [x_2 \cdot \lceil \frac{1-u_1}{x_1} \rceil^0\ \min(1,\ 1-x_2+\frac{x_2 \cdot v_1}{x_1}, z)]}$$

where $z = \begin{cases} 1, & \text{if } u_2 = 0 \\ \frac{y_2}{x_1} \cdot (\frac{v_1}{u_2} + \min(0, x_1 - v_1)), & \text{otherwise} \end{cases}$

Figura 4: Reglas v a ix, Figura de [3].

Las reglas x, xi y xviii se derivan de $\mathbf{P}(\epsilon) = [0,\ 0]$, entonces $\mathbf{P}(\phi|\epsilon) \in \emptyset$. De la xiii a la xv se considera la relación "α *entails* β" la cual indica que para que α sea verdadero β debe ser verdadero. La regla xvi representa la ausencia de conocimiento sobre una probabilidad. La regla xvii, también conocida como *multiple derivation rule*, es una de las más importantes ya que indica que si se tienen dos intervalos posibles para una probabilidad, ésta debe encontrarse en la intersección. A partir de esta regla se pueden ir acotando los intervalos de probabilidad para acercarse al intervalo más pequeño. La xix indica que si una probabilidad se encuentra en un intervalo I también se encuentra en cualquier intervalo J que contenga a I. Estas se encuentran en la Figura 5.

(x)

$$\frac{\boldsymbol{P}(\beta \mid \delta) \in [0\ 0]}{\boldsymbol{P}(\alpha \mid \beta \wedge \delta) \in \emptyset}$$

(xi)

$$\frac{\boldsymbol{P}(\alpha \mid \beta) \in \emptyset}{\boldsymbol{P}(\beta \mid T) \in [0\ 0]}$$

(xii)

$$\frac{\boldsymbol{P}(\alpha \mid \beta) \in [x\ y]}{\boldsymbol{P}(\alpha \mid \delta) \in [x\ y]}$$

provided β logically equivalent to δ

(xiii)

$$\frac{}{\boldsymbol{P}(\alpha \mid \delta) \in [1\ 1]}$$

provided δ entails α

(xiv)

$$\frac{\boldsymbol{P}(\alpha \mid \delta) \in [x\ y]}{\boldsymbol{P}(\beta \mid \delta) \in [x\ 1]}$$

provided α entails β

(xv)

$$\frac{\boldsymbol{P}(\beta \mid \delta) \in [x\ y]}{\boldsymbol{P}(\alpha \mid \delta) \in [0\ y]}$$

provided α entails β

(xvi)

$$\frac{}{\boldsymbol{P}(\alpha \mid \delta) \in [0\ 1]}$$

(xvii)

$$\frac{\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \qquad \boldsymbol{P}(\alpha \mid \delta) \in [u\ v]}{\boldsymbol{P}(\alpha \mid \delta) \in [\max(x, u)\ \ \min(y, v)]}$$

(xviii)

$$\frac{\boldsymbol{P}(\alpha \mid T) \in \emptyset}{\boldsymbol{P}(\beta \mid \gamma) \in \emptyset}$$

(xix)

$$\frac{\boldsymbol{P}(\alpha \mid \delta) \in I}{\boldsymbol{P}(\alpha \mid \delta) \in J}$$

provided $I \subset J$

Figura 5: Reglas x a xix, Figura de [3].

Las reglas xx a xxxii pueden derivarse a partir de las primeras haciendo uso de equivalencias lógicas, como $\beta \to \alpha \equiv \neg\beta \vee \alpha$. Estas se encuentran en la Figura 6.

(xx)

$$\frac{\begin{array}{l}\boldsymbol{P}(\beta \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\beta \to \alpha \mid \delta) \in [u\ v]\end{array}}{\boldsymbol{P}(\alpha \mid \delta) \in [\lceil x+u-1 \rceil^{0}\ v]}$$

provided $y \geqslant 1-v$

(xxi)

$$\frac{\begin{array}{l}\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\beta \to \alpha \mid \delta) \in [u\ v]\end{array}}{\boldsymbol{P}(\beta \mid \delta) \in [1-v\ \lfloor y+1-u \rfloor_{1}]}$$

provided $x \leqslant v$

(xxii)

$$\frac{\begin{array}{l}\boldsymbol{P}(\beta \wedge \gamma \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\beta \to \alpha \mid \delta) \in [u\ v]\end{array}}{\boldsymbol{P}(\alpha \wedge \beta \wedge \gamma \mid \delta) \in [\lceil x+u-1 \rceil^{0}\ \min(y,v)]}$$

(xxiii)

$$\frac{\begin{array}{l}\boldsymbol{P}(\alpha \to \beta \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\alpha \to \gamma \mid \delta) \in [u\ v]\end{array}}{\boldsymbol{P}(\alpha \to \beta \wedge \gamma \mid \delta) \in [\lceil x+u-1 \rceil^{0}\ \min(y,v)]}$$

(xxiv)

$$\frac{\begin{array}{l}\boldsymbol{P}(\beta \to \alpha \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\gamma \to \alpha \mid \delta) \in [u\ v]\end{array}}{\boldsymbol{P}(\beta \wedge \gamma \to \alpha \mid \delta) \in [\max(x,u)\ \lfloor y+v \rfloor_{1}]}$$

(xxv)

$$\frac{\begin{array}{l}\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\beta \mid \delta) \in [u\ v]\end{array}}{\boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [\lceil x+u-1 \rceil^{0}\ \min(y,v)]}$$

(xxvi)

$$\frac{\boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [x\ y]}{\boldsymbol{P}(\alpha \mid \delta) \in [x\ 1]}$$

(xxvii)

$$\frac{\begin{array}{l}\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\beta \mid \delta) \in [u\ v]\end{array}}{\boldsymbol{P}(\alpha \vee \beta \mid \delta) \in [\max(x,u)\ \lfloor y+v \rfloor_{1}]}$$

(xxviii)

$$\frac{\boldsymbol{P}(\alpha \vee \beta \mid \delta) \in [x\ y]}{\boldsymbol{P}(\alpha \mid \delta) \in [0\ y]}$$

(xxix)

$$\frac{\begin{array}{l}\boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\alpha \wedge \neg\beta \mid \delta) \in [u\ v]\end{array}}{\boldsymbol{P}(\alpha \mid \delta) \in [x+u\ \lfloor y+v \rfloor_{1}]}$$

provided $v \leqslant 1-x,\ y \leqslant 1-u$

(xxx)

$$\frac{\boldsymbol{P}(\alpha \mid \delta) \in [x\ y]}{\boldsymbol{P}(\beta \mid \delta) \in [x\ y]}$$

provided β **logically equivalent to** α

(xxxi)

$$\frac{}{\boldsymbol{P}(\alpha \mid \delta) \in [0\ 0]}$$

provided α **is unsatisfiable**

(xxxii)

$$\frac{}{\boldsymbol{P}(T \mid \delta) \in [1\ 1]}$$

Figura 6: Reglas xx a xxxii, Figura de [3].

Las últimas 6 reglas se definieron para tratar con probabilidades independientes, para lograrlo se añade a $\mathcal{L}_{PL}$ el operador de independencia $\left[Indep(\phi, \delta, \epsilon)\right]^M$ el cual es verdadero si y sólo si $\left[\mathbf{P}(\phi \wedge \epsilon \wedge \delta)\right] = \left[\mathbf{P}(\phi \wedge \delta)\right]^M \cdot \left[\mathbf{P}(\epsilon \wedge \delta)\right]^M$. Las reglas relativas a eventos independientes se encuentran en la Figura 7.

(xxxiii)

$$\frac{\begin{array}{l}\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\beta \mid \delta) \in [u\ v] \\ \mathsf{Indep}(\alpha, \delta, \beta)\end{array}}{\boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [x \cdot u\ \ y \cdot v]}$$

(xxxiv)

$$\frac{\begin{array}{l}\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [u\ v] \\ \mathsf{Indep}(\alpha, \delta, \beta)\end{array}}{\boldsymbol{P}(\beta \mid \delta) \in [u/y\ \ \lfloor v/x \rfloor_1]}$$

(xxxv)

$$\frac{\begin{array}{l}\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\beta \mid \delta) \in [u\ v] \\ \mathsf{Indep}(\alpha, \delta, \beta)\end{array}}{\boldsymbol{P}(\alpha \wedge \beta \mid \delta) \in [x + u - x \cdot u\ \ \ y + v - y \cdot v]}$$

(xxxvi)

$$\frac{\begin{array}{l}\boldsymbol{P}(\alpha \mid \delta) \in [x\ y] \\ \boldsymbol{P}(\alpha \vee \beta \mid \delta) \in [u\ v] \\ \mathsf{Indep}(\alpha, \delta, \beta)\end{array}}{\begin{array}{l}\boldsymbol{P}(\beta \mid \delta) \in [\lceil (u-y)/(1-y) \rceil^0 \\ \qquad\qquad (v-x)/(1-x)]\end{array}}$$

(xxxvii)

$$\frac{\mathsf{Indep}(\alpha, \delta, \beta)}{\mathsf{Indep}(\beta, \delta, \alpha)}$$

(xxxviii)

$$\frac{\mathsf{Indep}(\alpha, \delta, \gamma)}{\mathsf{Indep}(\beta, \delta \gamma)}$$

provided α and β are equivalent

Figura 7: Reglas i a iv, Figura de [3].

Se pueden definir casos particulares de problemas de *entailment* probabilístico a los que Frisch y Haddawy denominan problemas tipo A y tipo B. Estos problemas tienen únicamente probabilidades no condicionales.

En los problemas tipo A las implicaciones tienen probabilidad 1, es decir $\mathbf{P}(\alpha \rightarrow \beta) \in [1, 1]$ y la conclusión es la probabilidad de una proposición atómica de $\mathcal{L}_{pro}$. Para los problemas tipo A con las reglas xx, xxi, xvi y xvii es posible formar un sistema de deducción que cumple con la propiedad de completitud, con lo cual se tiene que es posible deducir el *entailment* estrecho.

Los problemas de tipo B consideran $\mathbf{P}(\alpha \rightarrow \beta) \in I$ y la conclusión puede ser la conjunción de distintas proposiciones atómicas de $\mathcal{L}_{pro}$. Los problemas tipo A son un caso particular de los problemas tipo B. De los problemas tipo B conjeturan que también se puede definir un sistema completo haciendo uso de las reglas xx a xxvi, xvi y xvii. Sin embargo no lo demostraron.

7. Conclusiones

La propuesta de Nilsson fue una de las primeras aproximaciones al tratamiento de probabilidad en lógica y, a pesar de definir un método de deducción, no es un método que pueda utilizarse en la práctica [2].

La propuesta de Frisch y Haddawy resuelve el problema del método de Nilsson para que sea posible encontrar una solución parcial al problema de *entailment* probabilístico en un tiempo razonable [3]. Más aún, proveen una serie de reglas con las que se puede acotar la probabilidad y para problemas tipo A el método de deducción que proponen es completo.

Referencias

[1] J. Lukasiewicz, "Logical foundations of probability theory," *Jan Lukasiewicz, Selected Works*, pp. 16–63, 1970.

[2] N. J. Nilsson, "Probabilistic logic," *Artificial Intelligence*, vol. 28, no. 1, pp. 71 – 87, 1986.

[3] A. M. Frisch and P. Haddawy, "Anytime deduction for probabilistic logic," *Artificial Intelligence*, vol. 69, no. 1, pp. 93 – 122, 1994.