Tema 06

La lógica en Inteligencia Artificial

Contenidos

- Introducción
- Lógica proposicional
- Lógica de predicados o de primer orden
- Conclusiones

Introducción

La lógica en IA sirve para representar el conocimiento y lo que es más importante, para estudiar si un determinado **razonamiento** seguido para llegar a una conclusión está bien realizado (**es válido**).

Si un razonamiento se expresa en base a las reglas descritas mediante un formalismo lógico podremos saber, de forma teórica (sin tener que ejecutar el sistema):

- Si los razonamientos realizados por ese sistema para llegar a cierta conclusión son siempre correctos (**soundness - solidez**),
- Si el sistema podrá encontrar toda la información que se pueda derivar de la situación inicial (**completeness - completitud**).
- Si existe un método en ese sistema que pueda ser usado para probar que cualquier expresión existente es verdadera o falsa (**decidability - decidibilidad**).

Lógica proposicional

Es la lógica más sencilla (y también "menos expresiva") que podemos utilizar.

La lógica proposicional trabaja con **enunciados** (**statements**) (también llamados **proposiciones**, también llamadas **expresiones lógicas** (**logical expression**).

Un enunciado es una expresión a la que se le puede asignar un valor de verdad: *verdadero (true, T)* o *falso (false, F)*.

Hoy es viernes

Ayer llovió

Hace frío

~~Cuidado!~~

La música era bonita

~~¿Tienes hambre?~~

(El valor de verdad dependerá muchas veces del contexto o del actor)

Lógica proposicional

Los enunciados en lógica proposicional se denotan por una letra, habitualmente p, q, r, …

Para trabajar en lógica proposicional habitualmente se usan esos símbolos sin asociarles un enunciado concreto, entonces los podemos llamar **variables**.

Con los enunciados se pueden formar expresiones compuestas utilizando **operadores lógicos** representados mediante símbolos especiales.

Symbol	Name	Example	English Equivalent
$\wedge$	conjunction	$\mathbf{p} \wedge \mathbf{q}$	**p** and **q**
$\vee$	disjunction	$\mathbf{p} \vee \mathbf{q}$	**p** or **q**
$\sim$	negation	$\sim\mathbf{p}$	not **p**
$\Rightarrow$	implication	$\mathbf{p} \Rightarrow \mathbf{q}$	if **p** then **q** *or* **p** implies **q**
$\Leftrightarrow$	biconditional	$\mathbf{p} \Leftrightarrow \mathbf{q}$	**p** if and only if **q** *or* **p** is equivalent to **q**

Lógica proposicional

Los operadores lógicos reciben estos nombres:

$\neg$ $(\sim)$	Operador "no"	Negación
$\wedge$	Operador "y"	Conjunción
$\vee$	Operador "o"	Disyunción
$\Rightarrow$	Operador "implica"	Implicación
$\Leftrightarrow$	Operador "sí y solo sí"	Bicondicional

Lógica proposicional

Las semántica, o valor de verdad de estas expresiones compuestas se representan mediante tablas de verdad en las que se representa el valor de verdad de la expresión compuesta para cada valor de verdad que puedan tomar las variables que la forman...

(a) AND function

p	q	p ∧ q
F	F	F
F	T	F
T	F	F
T	T	T

(b) OR function (c)

p	q	p ∨ q
F	F	F
F	T	T
T	F	T
T	T	T

(c) NOT function

p	~p
F	T
T	F

Cada combinación de valores de verdad para todas las variables de la expresión se denomina interpretación.

Si una interpretación hace que la expresión sea verdadera, a esa interpretación se la denomina modelo.

Para la expresión p^q:

I1={p=F, q=F}, I2={p=V, q=F}, etc, son interpretaciones

M1={p=V, q=V} es una interpretación y también un modelo

Lógica proposicional

Los operadores menos "intuitivos" son el **OR exclusivo** (XOR) la **implicación** y el **bicondicional**

OR exclusivo

p	q	$p \veebar q$
F	F	F
F	T	T
T	F	T
T	T	F

Implicación ($\Rightarrow$) y bicondicional ($\Leftrightarrow$)

p	q	$p \Rightarrow q$	$p \Leftrightarrow q$
F	F	T	T
F	T	T	F
T	F	F	F
T	T	T	T

Lógica proposicional

Se pueden encontrar expresiones compuestas que son equivalentes a otras, es decir, que tienen el mismo valor de verdad siempre. Por ejemplo:

$$(p \Rightarrow q) \equiv \sim p \vee q$$

La expresión anterior (si es que ocurre que las dos expresiones tienen el mismo valor de verdad siempre) se denomina tautología o teorema. Para demostrar que esta expresión es un teorema se pueden comprobar las tablas de verdad de ambos lados del símbolo $\equiv$

Dicho de otro modo: una expresión es una tautología si es verdadera para cualquier interpretación

p	q	$(p \Rightarrow q)$	$(\sim p \vee q)$	$(p \Rightarrow q) \equiv (\sim p \vee q)$
F	F	T	T	T
F	T	T	T	T
T	F	F	F	T
T	T	T	T	T

Lógica proposicional

Algunos teoremas tienen nombres específicos, como la ley de De Morgan:

$$(\sim p \vee \sim q) \equiv \sim(p \wedge q)$$

p	q	$(\sim p \vee \sim q)$	$\sim(p \wedge q)$	$(\sim p \vee \sim q) \equiv \sim(p \wedge q)$
F	F	T	T	T
F	T	T	T	T
T	F	T	T	T
T	T	F	F	T

Lógica proposicional

Algunos teoremas útiles

Theorem	Name
$p \vee q \equiv q \vee p$	Commutative Property 1
$p \wedge q \equiv q \wedge p$	Commutative Property 2
$p \vee p \equiv p$	Idempotency Law 1
$p \wedge p \equiv p$	Idempotency Law 2
$\sim\sim p \equiv p$	Double Negation (or Involution)
$(p \vee q) \vee r \equiv p \vee (q \vee r)$	Associative Law 1
$(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$	Associative Law 2
$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$	Distributive Law 1
$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$	Distributive Law 2
$p \vee T \equiv T$	Domination Law 1
$p \wedge F \equiv F$	Domination Law 2
$(p \equiv q) \equiv (p \Rightarrow q) \wedge (q \Rightarrow p)$	Law of Elimination 1
$(p \equiv q) \equiv (p \wedge q) \vee (\sim p \wedge \sim q)$	Law of Elimination 2
$p \vee \sim p \equiv T$	Law of Excluded Middle
$p \wedge \sim p \equiv F$	Contradiction

Lógica proposicional

Mediante el uso de unos teoremas se pueden ir deduciendo otros:

Que siempre es verdadera para cualquier interpretación

Prove that $[(\sim\mathbf{p} \vee \mathbf{q}) \wedge \sim\mathbf{q}] \Rightarrow \sim\mathbf{p}$	is a tautology
$[(\sim\mathbf{p} \wedge \sim\mathbf{q}) \vee (\mathbf{q} \wedge \sim\mathbf{q})] \Rightarrow \sim p$	distributive law 1
$[(\sim\mathbf{p} \wedge \sim\mathbf{q}) \vee F] \Rightarrow \sim\mathbf{p}$	non-contradiction
$(\sim\mathbf{p} \wedge \sim\mathbf{q}) \Rightarrow \sim\mathbf{p}$	domination law 2
$\sim(\sim\mathbf{p} \wedge \sim\mathbf{q}) \vee \sim\mathbf{p}$	alternate definition for implication
$(\sim\sim\mathbf{p} \vee \sim\sim\mathbf{q}) \vee \sim\mathbf{p}$	De Morgan's law
$(\mathbf{p} \vee \mathbf{q}) \vee \sim\mathbf{p}$	involution
$\mathbf{p} \vee (\mathbf{q} \vee \sim\mathbf{p})$	associative law 1
$\mathbf{p} \vee (\sim\mathbf{p} \vee \mathbf{q})$	commutative law1
$(\mathbf{p} \vee \sim\mathbf{p}) \vee \mathbf{q}$	associative law 1
$\mathbf{T} \vee \mathbf{q}$	law of excluded middle
$\mathbf{T}$	domination law 1

Mediante la aplicación de teoremas conocidos llegamos a que la expresión es siempre verdadera, luego es una tautología (un teorema)

Theorem	Name
$p \vee q \equiv q \vee p$	Commutative Property 1
$p \wedge q \equiv q \wedge p$	Commutative Property 2
$p \vee p \equiv p$	Idempotency Law 1
$p \wedge p \equiv p$	Idempotency Law 2
$\sim\sim p \equiv p$	Double Negation (or Involution)
$(p \vee q) \vee r \equiv p \vee (q \vee r)$	Associative Law 1
$(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$	Associative Law 2
$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$	Distributive Law 1
$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$	Distributive Law 2
$p \vee T \equiv T$	Domination Law 1
$p \wedge F \equiv F$	Domination Law 2
$(p \equiv q) \equiv (p \Rightarrow q) \wedge (q \Rightarrow p)$	Law of Elimination 1
$(p \equiv q) \equiv (p \wedge q) \vee (\sim p \wedge \sim q)$	Law of Elimination 2
$p \vee \sim p \equiv T$	Law of Excluded Middle
$p \wedge \sim p \equiv F$	Contradiction

Lógica proposicional

Hemos visto que una expresión que **siempre es verdadera** se denomina **tautología**.

De forma similar, a una expresión que **siempre es falsa** se la denomina una **contradicción**.

$$(p \wedge \sim p)$$

Una expresión se dice que es **satisfacible** (satisfiable) cuando existe al menos una asignación de valores de verdad a las proposiciones que la forman (una interpretación) que hace que la expresión entera sea verdadera.

Para **comprobar si una expresión es satisfacible** el único medio es comprobar la tabla de verdad de la misma asignando todas las combinaciones de valores de verdad posibles a cada una de las variables.

Esta comprobación tiene una **complejidad exponencial** al aumentar el número de sentencias: *con dos variables (p, q) tenemos 2^2 posibles combinaciones de valores de verdad, con tres variables (p, q, r) tenemos 2^3, y así sucesivamente*.

Lógica proposicional

La lógica proposicional puede utilizarse para comprobar la **validez de un argumento** utilizando el concepto de tautología.

Un argumento está formado por una serie de **premisas** y una **conclusión**:

P1
P2
.
.
Pr
____
Conclusión

El argumento se considera válido si la implicación formada por la conjunción de las premisas tomadas como antecedente y la conclusión tomada como consecuente es una tautología:

$(P1 \wedge P2 \wedge \ldots \wedge Pr) \Rightarrow$ Conclusión → ¿es una tautología?

Si lo es, el argumento es válido

Lógica proposicional

Por ejemplo, ¿es válido este argumento:?

1. $p \Rightarrow q$
2. $q \Rightarrow \sim r$
3. $\sim p \Rightarrow \sim r$

$\sim r$

Es lo mismo que comprobar si la siguiente expresión es una tautología:

$$[\,(p \Rightarrow q) \wedge (q \Rightarrow \sim r) \wedge (\sim p \Rightarrow \sim r)\,] \Rightarrow \sim r$$

Es una tautología, por tanto, el argumento es válido.

Podemos comprobarlo construyendo la tabla de verdad:

1	2	3	4	5	6	7	8
p	q	r	$p \Rightarrow q$	$q \Rightarrow \sim r$	$\sim p \Rightarrow \sim r$	$4 \wedge 5 \wedge 6$	$7 \Rightarrow \sim r$
F	F	F	T	T	T	T	T
F	F	T	T	T	F	F	T
F	T	F	T	T	T	T	T
F	T	T	T	F	F	F	T
T	F	F	F	T	T	F	T
T	F	T	F	T	T	F	T
T	T	F	T	T	T	T	T
T	T	T	T	F	T	F	T

Lógica proposicional

Si la expresión es una tautología se trata de un **razonamiento universalmente** válido, independientemente de la interpretación de las variables.

Si la expresión da valores verdaderos y falsos se trata de un **razonamiento contingente**, unas veces es válido y otras no dependiendo de lo que signifiquen las variables.

Si la expresión siempre da valores falsos se trata de una **contradicción**.

A la lógica únicamente le interesan los razonamientos del primer tipo, es decir, aquellos que siempre sean válidos. Un caso en el que una de las posibilidades fuese falsa sería suficiente para descartar esa "forma" de argumentar por ser inválida.

Lógica proposicional

La lógica proposicional propone esquemas de razonamiento válidos que permiten a un sistema avanzar en su conocimiento. Son las **reglas de inferencia**

La regla más utilizada se denomina **modus ponens**:

$$\frac{\alpha \Rightarrow \beta \qquad \alpha}{\beta}$$

cada vez que encontramos dos sentencias en la forma $\alpha \Rightarrow \beta$ y α, entonces la sentencia β puede ser inferida.

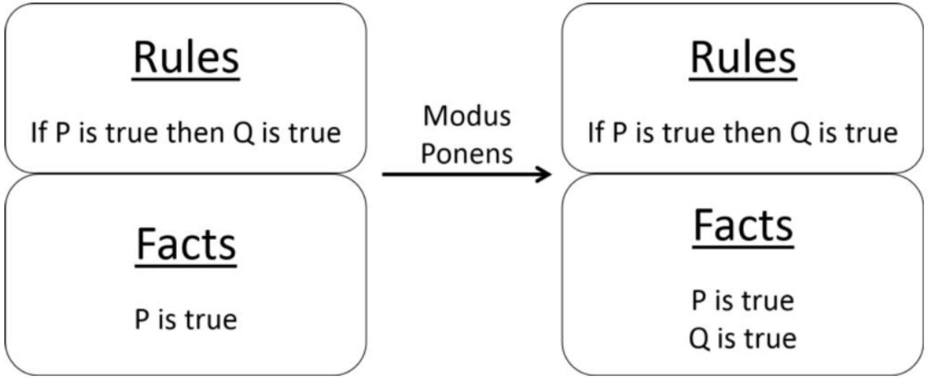
Esto supone que **el sistema ampliará su conocimiento**, al obtener una nueva información (β) lo que puede llevar a poder seguir aplicando más inferencias.

Lógica proposicional

La regla **modus ponens**: nos recuerda a la forma de funcionar de los sistemas expertos basados en reglas.

$$\frac{\alpha \Rightarrow \beta \qquad \alpha}{\beta}$$

La implicación ⇒ se puede leer como una cláusula if…then… (¡una **regla**!) solamente será utilizada si el antecedente α es verdadero. α sería el **hecho**.



Muchos problemas se pueden representar con expresiones basadas en la implicación donde α es una conjunción de proposiciones y β es una proposición afirmativa:

$$(p1 \land p2 \land \ldots \land pn) \Rightarrow Q$$

Cláusula de Horn

Lógica proposicional

La **validez** de un argumento tiene que ver con la **estructura** del mismo, no con su significado,.

Nos cuesta entender esto porque siempre buscamos conexiones causales y ejemplos intuitivos en vez de abstraernos de lo que representa cada proposición, tratándola simplemente como un símbolo.

La **lógica proposicional es demasiado sencilla** para representar el conocimiento en dominios a poco complejos que sean.

La **lógica de predicados toma como base a la lógica proposicoinal** pero trata de crear un formalismo para representar frases más cercanas al lenguaje natural, pero sin la ambiguedad de este ni la dependencia del contexto.

Lógica de predicados

La lógica de predicados (o lógica de primer orden) nos permite hablar del mundo de una forma cercana a como lo hace el ser humano, hablando de objetos y sus propiedades o relaciones con otros objetos.

El hombre_1 tiene un sombrero (es de color) azul

Los nombres denotan objetos

Los verbos denotan relaciones

Si son unitarias se denominan propiedades

También pueden ser relaciones n-arias

"Ana es hermana de Jose"

Algunas relaciones son funciones

"el padre de", "el mejor amigo de"

La lógica de primer orden también puede expresar hechos acerca de algunos o todos los objetos de un universo de discurso, por tanto nos permite representar leyes generales o reglas.

Lógica de predicados

La lógica de predicados tiene un poder expresivo mayor que la lógica de proposiciones.

Una expresión en lógica de predicados está compuesta de un **predicado** con una serie de **argumentos**, que van entre paréntesis.

En las expresiones en lógica de predicados se pueden utilizar los mismos operadores lógicos que en la lógica de proposiciones:

$$\sim, \wedge, \vee, \Rightarrow, \Leftrightarrow$$

Además, a las variables utilizadas en los predicados se les puede asociar un cuantificador

Cuantificador existencial: $\exists$

$\exists x$ "existe un x": se garantiza que uno o más valores de x existen

Cuantificador universal: $\forall$

$\forall x$ "para todo x": lo que dice el predicado se puede aplicar a todo x

Lógica de predicados

Ejemplo de predicados:

$$\forall(x)\{[\mathrm{Animal}(x) \wedge \mathrm{Has\_Hair}\ (x) \wedge \mathrm{Warm\_Blooded}\ (x)] \Rightarrow \mathrm{Mammal}\ (x)\}$$

"Para todo x, tal que x es un animal y x tiene pelo y x tiene sangre caliente, entonces x es un mamífero".

Un predicado dice algo de un sujeto (en este caso la variable x).

$$\{\mathrm{Brother}\ (x, \mathrm{Sam}) \Rightarrow (\exists \mathbf{y})\ [(\mathrm{Parent}\ (\mathbf{y}, \mathbf{x}) \wedge \mathrm{Parent}\ (\mathbf{y}, \mathrm{Sam}) \wedge \mathrm{Male}\ (\mathbf{x}) \wedge \sim \mathrm{Equal}\ (\mathbf{x}, \mathrm{Sam})]\}$$

"Si x es hermano de Sam, entonces existe al menos un y tal que y es el padre de x, e y es el padre de Sam, y x es macho y x no es Sam".

Lógica de predicados

La comprobación del valor de verdad de los predicados es más compleja que en el caso de la lógica proposicional.

La lógica de predicados utiliza el mismo vocabulario que la proposicional:(proposiciones, operadores booleanos y reglas de inferencia (como modus ponens), pero añade el uso de variables con cuantificadores.

Esto hace que la lógica de predicados sea mucho más potente, pero haciendo más complicada la ejecución de las inferencias por la necesidad de emparejar las variables.

Conclusiones

La lógica de proposiciones puede representar el mundo mediante una serie de sentencias de las que podemos comprobar su valor de verdad.

La lógica de predicados puede representar el mundo mediante una serie de objetos y sus propiedades, permitiendo hacer referencia a reglas genéricas que se pueden aplicar a multitud de casos particulares.

Lógica de proposiciones

Habría que hacer una expresión para cada combinación de seres que quisiésemos representar!

***Si** el gato está a la izquierda del perro y el perro está a la izquierda del humano **entonces** el gato está a la izquierda del humano.*

Lógica de predicados

Esta sentencia sirve para cualquier tipo de elemento que pueda ser la variable A, la B o la C

Para todo A, B y C: **Si** A está a la izquierda de B y B está a la izquierda de C, **entonces** A está a la izquierda de C

Fin