



**Universidad
Internacional
de Valencia**
**Máster Universitario
en Inteligencia Artificial**

UNIVERSIDAD INTERNACIONAL DE VALENCIA

MASTER EN INTELIGENCIA ARTIFICIAL

MEMORIA DEL CURSO MATEMÁTICAS PARA LA INTELIGENCIA ARTIFICIAL

PROGRAMACIÓN LÓGICA INDUCTIVA: MÉTODOS Y TEORÍAS

ALUMNOS:

Franklin Garcia Ruiz

Cristian Francisco Rojas Ledezma

Laura Rodriguez King

Mayra Pullupaxi Ataballo

BY STEPHEN MUGGLETON AND LUC DE RAEDT

La memoria de este artículo científico gira en torno a la programación lógica inductiva, y por ende su contenido detalla sobre su proceso, metodología, aplicaciones e interacciones, pero ¿qué es la ILP o programación lógica inductiva? Para dar respuesta a esta primera y central inquietud hagamos uso del método deductivo e iniciemos por entender su entorno de trabajo.

Globalizadamente la ILP se describe en cuatro niveles: nivel semántico, nivel algorítmico, nivel demostrativo-teórico y nivel semántico probabilístico de confianza.

En ese orden de ideas la Programación lógica inductiva primero detecta y define el problema, seguido a ello, define y estudia que reglas de la inferencia aplican a dicho problema, los algoritmos de ILP transforman esas reglas para que sean comprendidas con facilidad (computacionalmente costoso para ser de interés práctico) y finaliza con la justificación de las hipótesis de inducción.

Ahora bien, los niveles hablan del término "semántica" y de hecho es un término bastante común en la ILP que debemos comprender para abarcar el estudio de la programación lógica inductiva.

La semántica definida como elementos lógicos, hacen parte del modelo-teórico del ILP y podemos encontrar dos tipos: la semántica normal y no monotónica.

Iniciemos con la semántica normal:

Dentro del modelo-teórico el objetivo de la semántica normal es hallar una hipótesis tal que cumpla con la siguiente condición:

PROOF. We prove that, under these assumptions, $\mathcal{M}^+(B \wedge E^+) = \mathcal{M}^+(B \wedge H)$.
 Define \mathcal{B}_P as $\mathcal{B}(B \wedge H)$
 1) $\mathcal{M}^+(B \wedge E^+) \subset \mathcal{M}^+(B \wedge H)$ because E^+ is true in $\mathcal{M}^+(B \wedge H)$ (posterior sufficiency) and B is true in $\mathcal{M}^+(B \wedge H)$
 so $E^+ \wedge B$ is true in $\mathcal{M}^+(B \wedge H)$
 so $\mathcal{M}^+(E^+ \wedge B) \subset \mathcal{M}^+(B \wedge H)$
 2) $\mathcal{M}^+(B \wedge H) \subset \mathcal{M}^+(B \wedge E^+)$ because $B \wedge H \wedge E^- \not\models \square$ (posterior satisfiability)
 so $\mathcal{M}^+(B \wedge H) \cap \mathcal{M}^+(E^-) = \emptyset$
 so $\mathcal{M}^+(B \wedge H) \cap \mathcal{M}^-(B \wedge E^+) = \emptyset$
 so $\mathcal{M}^+(B \wedge H) \subset \mathcal{B}_P - \mathcal{M}^-(B \wedge E^+)$
 so $\mathcal{M}^+(B \wedge H) \subset \mathcal{M}^+(B \wedge E^+) \quad \square$

Donde B equivale al conocimiento previo, E como la evidencia y por último H como la hipótesis, cabe resaltar que la evidencia puede ser positiva (verdadera) o negativa (falsa).

El criterio de suficiencia en ocasiones actúa como la integridad con respecto a la evidencia positiva y el criterio de satisfabilidad posterior como consistencia con evidencia negativa.

La semántica normal posee un caso especial llamado "semántica definida" y su configuración es más sencilla puesto que trabaja con el modelo Herbrand y la evidencia se restringe a ser verdadera o falsa.

Prior Satisfiability. all $e \in E^-$ are false in $\mathcal{M}^+(B)$
Posterior Satisfiability. all $e \in E^-$ are false in $\mathcal{M}^+(B \wedge H)$
Prior Necessity. some $e \in E^+$ are false in $\mathcal{M}^+(B)$
Posterior Sufficiency. all $e \in E^+$ are true in $\mathcal{M}^+(B \wedge H)$

Semántica No-monotónica:

Introducido por Nicolas Helf y Flach, este escenario se compone de un grupo de cláusulas definidas (teoría de fondo), evidencia vacía e hipótesis usando el mismo alfabeto que la teoría de fondo.

Para este entorno, se deben cumplir las siguientes condiciones para H y B :

Validity: all $h \in H$ are true in $\mathcal{M}^+(B)$

Completeness: if general clause g is true in $\mathcal{M}^+(B)$, then $H \models g$

Minimality: there is no proper subset G of H which is valid and complete

La principal disimilitud entre ambos escenarios radica en que el entorno normal la hipótesis inducida puede ser utilizada como reemplazo ya que tanto la teoría como la hipótesis implican ejemplos observados, mientras que en el escenario no monotónico, la hipótesis es un conjunto de propiedades que contiene un conjunto de ejemplos.

Prior Satisfiability. $B \wedge E^- \not\models \square$

Posterior Satisfiability. $B \wedge H \wedge E^- \not\models \square$

Prior Necessity. $B \not\models E^+$

Posterior Sufficiency. $B \wedge H \models E^+$

Y aunque ambas parezcan muy diferentes, ambas aplican técnicas del ILP en común.

Para la formación y justificación de hipótesis basta con entender la siguiente fórmula:

Induction = Abduction + Justification

Detalladamente así:

Induction = Formación de hipótesis + Grado de creencia de la hipótesis reforzada con la cantidad de evidencias.

El sesgo es la influencia que tienen las inferencias inductivas con base a la experiencia, y se distinguen el sesgo declarativo que lanza la pregunta ¿Qué buscar? y el sesgo preferencial que determina ¿Cómo buscar? y con esto damos entrada a las probabilidades semánticas de la ILP.

Expón la interacción entre esta disciplina y la probabilidad, así como el rol que juega el teorema de Bayes y, en general, la inferencia bayesiana.

La interacción entre ambas disciplinas yace en que son aplicables incluso en la escasa posibilidad de tener solo ejemplos positivos disponibles. Cada consigna lógica inferida inductivamente tiene un valor de probabilidad. El rol que juega el teorema Bayesiano dentro del entorno ILP es la confirmación y corroboración de hipótesis, por ejemplo:

Supongamos que estamos intentando demostrar un grado de creencia y nos llega un ejemplo con negativos, es aquí donde debemos devolvemos a la sesión anteriormente explorada de la satisfactibilidad posterior que se encargará de reemplazar el elemento superior único por un conjunto de elementos superiores, y con ello hacer uso del teorema de Bayes:

$$p(T|E) = \frac{p(T) \cdot p(E|T)}{p(E)}$$

Resta asumir que la evidencia es correcta, esto conduce a la paradoja de si T tiene un conjunto finito de modelos es necesariamente falso y se aplica la inferencia bayesiana en la toma de decisión e inferir que la hipótesis es cierta.

Entonces la inferencia Bayesiana toma sentido cuando hay gran cantidad de circunstancias en donde debido a las posibilidades de controlar errores se puede aplicar enfoques de probabilidad inductivos en donde la probabilidad puede ser actualizada de acuerdo a las hipótesis planteadas y evidencias encontradas.

De la probabilidad pasamos a la estadística con sus técnicas de regresión en las redes neuronales. Ya anteriormente habíamos visto que los algoritmos generan reglas fáciles de comprender a diferencia de estos imitadores del aprendizaje humano. Durante años se han realizado diversos intentos para solventar los problemas de predicción general, El instituto de Turing recientemente logró un notable acercamiento a la predicción general haciendo uso de conocimientos biológicos previos y la capacidad de describir relaciones estructurales.

Para la conexión entre el aprendizaje automático y las redes neuronales se necesita proporcionar la suficiente cantidad de datos a las capas de las neuronas para que puedan reconocer patrones, clasificarlos y categorizarlos, para ello usa algoritmos para procesar datos, aprender de ellos y ser capaces de hacer predicciones, al trabajar con gran cantidad de datos permite que los algoritmos sean mejorados y obtiene mejores resultados.

A diferencia de la inferencia deductiva que trabaja a partir de fórmulas lógicas suministradas por el usuario y a partir de conocimientos previos logrando una predicción por hechos ya registrados y en las probabilidades la inferencia inductiva no garantiza que las conclusiones sigan mas allá de lo que se sabe y por ende el razonamiento inductivo sigue siendo menos conciso que las matemáticas deductivas.

Pongamos un par de ejemplos para hacer más clara la precisión de ambos conceptos:

Lógica inductiva:

Las manos de un humano generalmente tiene 5 dedos, lo que se podrían inducir que las manos tienen 5 dedos, sin embargo, en algún lugar del mundo podría nacer un humano con 6 dedos, o incluso sin dedos, a esto le llamamos lógica inductiva, donde la hipótesis puede ser revocada por evidencias.

Lógica deductiva:

Los animales herbívoros comen plantas, la vaca come planta, por lo tanto la vaca es herbívora.

Y para finalizar la lógica inductiva tiene una orientación semántica y varias técnicas bien establecidas, además de que está interesada en la convergencia de algoritmos, complejidad de los procesos, potencialmente utilizada para el análisis, aplicación de teorías incluso creación de las mismas. Casos de éxito como en la medicina muestra que la ILP es capaz de construir nuevas reglas que predicen la actividad de las medicinas aún no testeadas en humanos y permite encontrar mayor precisión que otros métodos. Así, pasando de lo general a lo particular, cerramos la memoria ILP como el proceso que emplea técnicas de aprendizaje automático encargado de expandirnos el horizonte.