

ILP en Aprendizaje Grupal

Telma Delladio

`td@cs.uns.edu.ar`

Departamento de Ciencias e Ingeniería de Computación

Universidad Nacional del Sur

Avda. Alem 1253 - Bahía Blanca - Bs. As.

Tel (0291) 459-5135 - Fax (0291) 459-5136

1 Introducción

La Programación en Lógica Inductiva (ILP) es una propuesta de aprendizaje que utiliza a la Programación en Lógica (PL) como herramienta de representación de conocimiento. Se han desarrollado muchos mecanismos de aprendizaje a través de la generación de programas lógicos que sintetizan una definición para un determinado concepto.

El tipo de aprendizaje desarrollado con ILP es un aprendizaje a partir de ejemplos. La clase de los ejemplos considerados (positivos y/o negativos) y el modo de procesamiento de los mismos varía de acuerdo a la técnica utilizada. El principal problema en este tipo de sistemas es la complejidad. Toda tarea de aprendizaje es compleja por naturaleza y por eso cualquier sistema que implemente algún mecanismo de aprendizaje deberá tratar con esta complejidad.

Se han desarrollado sistemas "generales" de aprendizaje a partir de ejemplos [Bra90, BG96] que son capaces de aprender conceptos diversos. El principal problema de estos sistemas es que la calidad de los conceptos aprendidos depende en gran medida de los ejemplos provistos (calidad y orden de procesamiento). Además para lograr cierto grado de eficiencia se suelen establecer restricciones en la generación de cláusulas junto con declaraciones de modos y/o tipos sobre los argumentos de los predicados que se desean generar en el proceso de aprendizaje. La generalidad de estos sistemas deriva en la ineficiencia de los mismos. Esto es, para lograr aprender cualquier concepto, hay que hacer un gran esfuerzo por brindar muy buenos ejemplos y gran cantidad de información adicional para que se pueda generar la definición del concepto en un tiempo y forma adecuados. Por otro lado, están los sistemas de aprendizaje, por ejemplo algunos clasificadores, que se han desarrollado para algún problema específico y que hacen uso de heurísticas para poder llevar adelante el aprendizaje en forma aceptable. Estos sistemas son de aplicación específica lo que hace que se puedan desarrollar mecanismos de aprendizaje puntuales al dominio de aplicación [LD94].

Cualquier sistema de aprendizaje tiene que tratar con el problema de la eficiencia, pues cualquier tarea de aprendizaje es muy compleja. Los sistemas desarrollados en el ámbito ILP basan su tarea de aprendizaje en la búsqueda de una definición para uno o más conceptos. En ILP las definiciones buscadas son programas lógicos, es decir un conjunto de cláusulas. Esta búsqueda se lleva a cabo en el Espacio de Hipótesis formado por todas aquellas cláusulas que son potenciales definiciones del concepto, o conceptos, que se están aprendiendo. Cuanto mayor sea este espacio, más difícil es la búsqueda. Todas las heurísticas o restricciones que puedan establecer este tipo de sistemas, tienen como objetivo reducir este espacio de hipótesis lo que hace más eficiente la búsqueda.

Por todo esto, uno de los principales objetivos en ILP es encontrar propuestas que ayuden de alguna forma a hacer este problema más tratable. Una propuesta considerable es encarar la tarea de aprendizaje en forma cooperativa. En muchas ocasiones es más efectivo o más fácil aprender algo en forma grupal. Diferentes personas pueden tener distintos puntos de vista, distinto conocimiento, distintas experiencias. Todo esto puede facilitar la tarea de aprendizaje.

2 Aprendizaje grupal

En el ámbito de ILP también es posible encarar la tarea de aprendizaje de forma cooperativa. La idea general es la misma; aprender un concepto, pero esta vez en forma grupal. Esto es, un grupo de entidades o agentes con distintos conocimientos, capacidades y observaciones trabajando en conjunto para aprender un mismo concepto. Esta propuesta no solo tiene que verse como una forma de llevar a cabo la tarea de aprendizaje en forma más eficiente, sino que puede presentarse en muchas situaciones como la única forma de lograr dicho aprendizaje. Por ejemplo, consideremos cuatro entidades E_1, E_2, E_3 y E_4 . Cada una de ellas con habilidades diferentes. E_1 puede distinguir colores: rojo, amarillo, violeta, verde. E_2 puede distinguir sabores: dulce, amargo, ácido. E_3 puede distinguir formas: esférica, oblonga, achatada y E_4 puede distinguir tamaños: grande, chico. Si el objetivo es aprender a identificar frutas como peras, manzanas y ciruelas es difícil que cada una de estas entidades por separado pueda hacerlo correctamente, pues ninguna de esas características (color, sabor, forma, tamaño) por separado es suficiente para distinguir una de estas frutas, pues la variedades de peras, manzanas y ciruelas es muy grande. Las manzanas ácidas pueden ser verdes o rojas. No todas las peras son verdes o amarillas, hay ciruelas rojas, violetas o amarillas, etc.

2.1 Esquema general

Esta propuesta de aprendizaje grupal consiste en definir un *grupo de aprendizaje*. Este grupo será un conjunto de n entidades, E_1, \dots, E_n , que compartirán sus conocimientos y capacidades para aprender un concepto, digamos C . Todas estas entidades tienen el mismo objetivo: “aprender C ”.

Cada entidad E_i tendrá su propio conocimiento previo (o base de conocimiento) B_i y tendrá acceso a un conjunto de observaciones O_i . La forma de acceso y el tipo de las observaciones O_i

dependerá de las características de cada E_i . Las observaciones pueden ser positivas y/o negativas y procesarse en conjunto o incrementalmente, dependiendo del mecanismo de aprendizaje, M_i , que posea cada E_i . Cada entidad desarrollará una definición (subjetiva) parcial, $D_i(C)$ del concepto C a partir de su conocimiento previo y sus observaciones. Luego a partir de $D_1(C), \dots, D_n(C)$ se buscará sintetizar en forma conjunta una definición global $D(C)$. La calidad de esta definición, es decir; su correctitud y su completitud, dependerá de las definiciones parciales obtenidas y de la manera en que se lleve a cabo la composición de dicha definición.

La idea hasta aquí expuesta es muy general, un grupo de entidades con distintas características trabajando en conjunto para aprender un concepto determinado. Lo que se quiere estudiar es cómo llevar adelante esta propuesta en el marco de la Programación en Lógica Inductiva.

Una de las características de la propuesta ILP es la uniformidad de la representación. Esto es, en ILP *todo* son programas lógicos. Las observaciones pueden expresarse como hechos lógicos. El conocimiento previo se expresa a través de un programa lógico y los conceptos aprendidos también se expresan como programas lógicos que pasarán a formar parte del conocimiento previo. Por esto, el mapeo de este esquema general a la propuesta ILP es directo. Lo que no es directo en esta propuesta, ni en otras, es la solución de aquellos problemas propios de cualquier actividad grupal, como por ejemplo la comunicación, el consenso de soluciones, etc.

2.2 Esquema ILP

Como ya se dijo, todo en ILP son programas lógicos. Por lo tanto, el conocimiento previo B_i de cada entidad será un programa lógico que representará todo el conocimiento disponible que posee E_i . Este conocimiento puede ser conocimiento inicial o conocimiento adquirido a través del tiempo, por ejemplo como consecuencia de alguna actividad de aprendizaje. Las observaciones, O_i , de cada entidad son representadas por un conjunto de hechos lógicos. Por último las definiciones parciales $D_i(C)$ estarán formadas por un conjunto de cláusulas obtenidas por E_i y que sirven como definición subjetiva del concepto C .

Lo importante aquí, es que cada definición $D_i(C)$ es *parcial* desde el punto de vista de la definición general que se está buscando. Es de esperar que cada definición parcial sea lo más correcta y completa desde el punto de vista de la entidad que la genera. Esto es, posiblemente $D_i(C)$ sea una definición completa y correcta para E_i considerando su conocimiento previo B_i y el conjunto de observaciones disponibles O_i . Por lo general, cuando se consideran observaciones positivas y negativas, en el marco ILP, se dice que una definición $D_i(C)$ será completa si deriva o cubre todas las observaciones positivas y se dice que es correcta si no deriva o cubre ninguna de las observaciones negativas. Dependiendo del mecanismo de aprendizaje M_i que cada entidad implemente, es posible que cada una de las definiciones parciales $D_i(C)$ puedan llegar a ser completas y correctas con respecto a su conjunto de observaciones O_i , pero no necesariamente lo serán con respecto a todas las observaciones disponibles en el grupo. Entonces, una definición parcial puede ser correcta desde el punto de vista de una entidad y puede no serlo si se considera el conocimiento

y las observaciones de *todo* el grupo de aprendizaje.

Cada entidad es capaz de definir al concepto C solo a partir de los recursos que tiene disponibles (conocimiento y observaciones). Si una entidad E_i solo tiene acceso a su base de conocimiento y a sus observaciones, es posible que la definición $D_i(C)$ generada esté en contradicción, por ejemplo, con el conjunto de observaciones O_j de otra entidad E_j . Esto posiblemente, provoque una contradicción entre las definiciones parciales $D_i(C)$ y $D_j(C)$.

Teniendo en cuenta todo esto, el objetivo del *grupo de aprendizaje* es sintetizar una definición general para $D(C)$. Indudablemente, dicha definición se conformará a partir de todas las definiciones parciales $D_1(C), \dots, D_n(C)$. Una propuesta ingenua sería considerar la unión de estas definiciones parciales como definición general. No obstante, un simple análisis muestra que esto no es factible. El argumento principal es que cada definición parcial cubre solo un aspecto del concepto que se quiere aprender y la definición general buscada deberá cubrir de la mejor manera posible todos los aspectos del concepto, es decir, ser lo más completa posible. Por otro lado, como ya se dijo, es posible que la simple unión entre dos definiciones parciales cree una contradicción.

Una de las cuestiones que hay que tener presente es que cada definición parcial es generada por una entidad en particular. Por lo tanto puede suceder que la definición parcial $D_i(C)$ no sea “comprendida” por la entidad E_j ($i \neq j$). Cada definición parcial $D_i(C)$ es un conjunto de cláusulas lógicas, esto es, la definición de un predicado que modela el concepto C desde el *punto de vista* de E_i . En la definición de ese predicado es posible que E_i utilice conceptos (predicados) ya definidos en su base de conocimiento los cuales no necesariamente estarán disponibles para las demás entidades.

3 Trabajo futuro

Lo que se ha presentado es un esquema simple de aprendizaje grupal desde el punto de vista de la Programación en Lógica Inductiva y se señalan los principales inconvenientes que se deben sortear al momento de llevar adelante una propuesta de este estilo. Se presentaron tan solo algunas de las cuestiones que hay que analizar y solucionar al momento de *consensuar* una definición general para poder sacar el mayor provecho de las características de cada una de las entidades. Cualquier definición general debe tener un cierto *grado de aceptación* de cada una de las entidades pertenecientes al grupo. Se debe encontrar una manera equilibrada de valorar los aportes que realiza cada entidad, de acuerdo a su conocimiento y sus capacidades. Así como también, la manera más efectiva de transmitir o compartir entre las entidades ciertas piezas de conocimiento o información que sirvan para “comprender” la definición general.

Los *grupos de aprendizaje* pueden tener características muy variadas. Se puede considerar una amplia gama que va desde los grupos totalmente homogéneos hasta aquellos totalmente heterogéneos. Un grupo totalmente homogéneo sería aquel compuesto por un conjunto de entidades de iguales características (observaciones, bases de conocimiento y mecanismos de aprendizajes). El principal interés está en el estudio de grupos que no sean totalmente homogéneos, puesto que

esos son los grupos que ofrecen un marco de trabajo diferente al del aprendizaje individual.

En el futuro se definirán distintos *tipos* de grupos de aprendizaje y en base a sus características se estudiarán las mejores formas de solucionar los problemas que se presentan principalmente al momento de *consensuar* una definición general para el concepto que se está aprendiendo.

Referencias

- [BG96] BERGADANO, F., AND GUNETTI, D. *Inductive Logic Programming, From Machine Learning to Software Engineering*. The MIT Press, 1996.
- [Bra90] BRATKO, I. *Prolog Programming for Artificial Intelligent*, second ed. Addison Wesley, 1990.
- [LD94] LAVRAC, N., AND DZEROSKI, S. *Inductive logic programming: Techniques and applications*, 1994.