Рассмотрим теперь пределы нижних и верхних границ индексов устойчивости при $k \to \infty$. Нижние границы поуровневых индексов ведут себя по-разному: для индексов верхних уровней они стремятся в пределе к 1, а для нижних уровней — к 0. В самом деле, в силу теоремы 1, имеет место:

$$J^k_{n+k-1} \geq \frac{1}{\binom{n+k}{n+k-1}} \left(\gamma_{n-1} + \binom{k}{k-1}\right)$$

$$\text{и} \lim_{k\to\infty} J^k_{n+k-1} = 1.$$

С другой стороны,

$$\underline{J}^k_2 = \frac{1}{\binom{n+k}{2}} \cdot \gamma_2 \text{ и } \lim_{k\to\infty} \underline{J}^k_2 = 0.$$

Открытым остается вопрос поведения нижних границ индексов средних уровней и усредненного индекса J_a.

Предел нижней границы интегрального индекса устойчивости строго больше нуля и меньше единицы, более точно

$$\lim_{k\to\infty} \underline{J}^k_\Sigma(k) = \frac{\gamma_\Sigma + 1}{2^n} \geq 0.$$

Верхние границы индексов устойчивости ведут себя единообразно: они монотонно возрастают и в пределе стремятся к 1.

Анализ асимптотики позволяет делать предположения о поведении лишь индекса J_Σ: скорее всего, при поступлении новых примеров, значение этого индекса будет возрастать, так как уменьшаться ему практически "некуда". Теперь, если мы представим множество объектов, подтверждающих какую-либо зависимость $X \to W$, $|X'| = n$, как возникшее в результате пополнения некоторого исходного множества размера $r < n$, мы можем сделать вывод о том, что у зависимостей с большим значением n скорее всего, J_Σ будет больше чем у зависимостей с малыми. Такая "мягкая" зависимость J_Σ от числа примеров позволяет отдать предпочтение интегральному индексу устойчивости как наиболее емкому из индексов с известной асимптотикой нижних границ: с одной стороны, явно, J_Σ отражает устойчивость зависимостей, а с другой — неявно: количество объектов, "подтверждающих" зависимость, т. е. $|X'|$.

Идея устойчивости может быть воплощена в различных индексах устойчивости. При различных обстоятельствах некоторые из них могут оказаться предпочтительней других. Примером может служить следующая "устойчивость прогноза". Пусть $K_+ = (G_+, M, I_+)$, $K_- = (G_-, M, I_-)$, $K_\tau = (G_\tau, M, I_\tau)$ — положительный, отрицательный и недоопределенный контексты относительно некоторого свойства W, не входящего в M. P_+, P_- — множества всех положительных и отрицательных прогнозов, полученных на основе всех зависимостей вида $X \to W$, порожденных по K_+ и K_-. Тогда I_p есть доля всех подмножеств множества $G + \cup G_-$, для которых множества всех порожденных прогнозов совпадают с прогнозами, полученными по всему множеству объектов $G_+ \cup G_-$.

СПИСОК ЛИТЕРАТУРЫ

1. Wille R. Restructuring lattice theory: an approach based on hierarchies of concepts // Ordered sets / Ed. I. Rival. — Reidel, Dordrecht-Boston, 1982.— P. 445–470.

2. Wille R. Dependencies between many-valued attributes // Classification and Related Methods of Data Analysis / Ed. H. H. Bock.— North-Holland, Amsterdam, 1998. — P. 581–586.

3. Финн В. К. Правдоподобные рассуждения в интеллектуальных системах типа ДСМ // Итоги науки и техники. Сер. Информатика. — М.: ВИНИТИ, 1991. — Т. 15.— С. 54–101.

4. Кузнецов С. О. Сложность алгоритмов обучения и классификации, основанных на поиске пересечения множеств // НТИ. Сер. 2.— 1991. — № 9. — С. 8–15.

УДК 164:510.65

Д. В. Виноградов

Логические программы для квазиаксиоматических теорий*

Как альтернатива многозначной слабой логике второго порядка квазиаксиоматическая теория (КАТ) правдоподобных рассуждений типа ДСМ может быть формализована в языке стратифицированных логических программ. Новая формализация допускает прямое исследование немонотонного поведения ДСМ-рассуждений, основанных на критерии достаточного основания индуктивного вывода.

Можно надеяться, что этот подход более приемлем для первоначального изучения теории правдоподобных рассуждений специалистами по искусственному интеллекту со знаниями в области логического программирования.

1. ВВЕДЕНИЕ

КАТ имеют дело со слабо формализованными и хорошо структурированными предметными областями: фармакологией, социологией, лингвистикой и т. д.

Для таких областей теории имеют следующие особенности:

- открытость (неполноту),
- наличие правил (правдоподобного) вывода, не сохраняющих тождественную истинность,
- немонотонность.

* Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант № 97-06-80191.

Наиболее развитой квазиаксиоматической теорией является теория правдоподобных рассуждений типа ДСМ [1]. Правила правдоподобного вывода в ней являются развитием методов сходства, различия и "сопутствующих изменений", предложенных первоначально Дж. С. Миллем [2] в XIX в.

Хорошая структурированность предметной области означает, что имеется операция "локального сходства", удовлетворяющая аксиомам нижней полурешетки на "замыкании" любого конечного множества, соответствующего конечной обучающей выборке. Эта операция порождает операцию "глобального сходства", соответствующую группировке всех объектов, обладающих общим сходством.

Совершенно аналогично случаю транзитивного замыкания бинарного отношения на конечных моделях, операция "глобального сходства" не может быть выражена формулой первого порядка. Поэтому для этой КАТ Д. П. Скворцовым [3] был предложен формализм слабой логики второго порядка, так называемое исчисление предикатов с кванторами по конечным множествам (и более выразительное исчисление предикатов с кванторами по кортежам переменной длины).

В работах [3; 4] разными методами было установлено, что множество формул, истинных в стандартной модели, не перечислимо.

Другим недостатком (с нашей точки зрения) является невозможность прямого исследования немонотонного поведения ДСМ-рассуждения. С помощью многозначных логик, основанных на логике предикатов с кванторами по кортежам переменной длины, в [5] описана дедуктивная имитация правдоподобного вывода типа ДСМ на фиксированной обучающей выборке. Однако ДСМ-рассуждение включает в себя механизм расширения обучающей выборки — критерий достаточного основания индуктивного вывода, обеспечивающий немонотонность правдоподобного вывода.

Мы предложим новый подход к формализации КАТ, допускающий исследование вопросов немонотонности.

2. ФОРМАЛИЗМ СТРАТИФИЦИРОВАННЫХ ЛОГИЧЕСКИХ ПРОГРАММ

Стратифицированные логические программы были введены и впервые изучены Чандрой и Харелом [6].

Общая логическая программа — это множество правил вида

$$t_0 \leftarrow t_1, t_2, \ldots, t_k;$$

где t_i — литералы. Литерал t_0 — *голова* правила, остальные литералы формируют *тело* правила. Каждый литерал в теле — это атомарная формула $Q(x_1, \ldots, x_n)$ или отрицание атомарной формулы $\neg\, Q(x_1, \ldots, x_n)$, где Q — один из реляционных символов R сигнатуры σ, или некоторая реляционная переменная (реляционный символ, не содержащийся в сигнатуре σ). Голова правила — атомарная формула $Q(x_1, \ldots, x_n)$, где Q — реляционная переменная. Одна из реляционных переменных программы выделена и называется *целевым предикатом*.

Общая логическая программа π, имеющая реляционными переменными символы $S_1, \ldots, S_m$, называется *стратифицированной*, если найдется такое разбиение $P = \bigcup_{i \leqslant l} P_i$ правил программы π, что выполняются следующие два условия:

- если реляционная переменная S_t встречается без отрицания в теле правила из P_i, то каждое правило с S_t в голове содержится в $P = \bigcup_{j \leqslant i} P_j$;
- если реляционная переменная S_t встречается под отрицанием в теле правила из P_i, то каждое правило с S_t в голове содержится в $P = \bigcup_{j < i} P_j$ (в частности, ни одна переменная не встречается под отрицанием в теле правила из первого страта P_1).

Пусть $D = (A; R_1, \ldots, R_s)$ — реляционная структура для сигнатуры σ. Рассмотрим программу π_1, образованную правилами первого страта P_1. Пусть она содержит реляционные переменные $S_1, \ldots, S_m$, арностей $n_1, \ldots, n_m$, соответственно. Программа π_1 задает оператор Θ_1, определенный на последовательностях $\underline{S} = (S_1, \ldots, S_m)$ из n_1-местного, ..., n_m-местного отношения на носителе A структуры D. Образ $\Theta_1(\underline{S})$ представляет собой такую же последовательность $\Theta_1(\underline{S}) = ({}^*S_1, \ldots, {}^*S_m)$ из n_1-местного, ..., n_m-местного отношения на носителе A. Эти отношения получаются "применением" всех правил π_1 к отношениям $R_1, \ldots, R_s$-структуры D и отношениям $S_1, \ldots, S_m$.

Точнее, $(a_1, \ldots, a_r) \in {}^*S_j$ если найдется такое правило $S_j(x_1, \ldots, x_r) \leftarrow t_1, t_2, \ldots, t_k$ в π_1, и для переменных $z_1, z_2, \ldots, z_q$ в $t_1, t_2, \ldots, t_k$, отличных от $x_1, \ldots, x_r$ найдутся такие элементы $c_1, c_2, \ldots, c_q$ носителя A, что одновременно:

- Для t_i, имеющего вид $R_p(y_1, \ldots, y_n)$, верно $(b_1, \ldots, b_n) \in R_p$; где $b_u = a_v$, если $y_u = x_v$, и $b_u = c_w$, если $y_u = z_w$.
- Для t_i, имеющего вид $\neg R_p(y_1, \ldots, y_n)$, верно $(b_1, \ldots, b_n) \notin R_p$; где $b_u = a_v$, если $y_u = x_v$, и $b_u = c_w$, если $y_u = z_w$.
- Для t_i, имеющего вид $S_p(y_1, \ldots, y_n)$, верно $(b_1, \ldots, b_n) \in S_p$; где $b_u = a_v$, если $y_u = x_v$, и $b_u = c_w$, если $y_u = z_w$.

Получающийся оператор Θ_1 монотонен, и можно рассмотреть его наименьшую неподвижную точку Θ_1^∞ (в упорядоченном множестве последовательностей $\underline{S} = (S_1, \ldots, S_m)$, состоящих из n_1-местного отношения $S_1, \ldots, n_m$-местного отношения S_m на носителе A, с покомпонентным порядком, на каждой компоненте являющимся теоретико-множественным включением подмножеств из соответствующей декартовой степени A). Определим, $\Theta_1^1 = \Theta_1(\underline{\varnothing})$ и $\Theta_1^{n+1} = \Theta_1(\Theta_1^n)$, тогда $\Theta_1^\infty = \bigcup_{n \in \omega} \Theta_1^n$, где объединение берется покомпонентно.

После этого рассмотрим расширенную сигнатуру $\sigma_1 = \sigma \cup \{S_1, \ldots, S_m\}$. Полученная наименьшая неподвижная точка Θ_1^∞ вместе со структурой D образует реляционную структуру $D_1 = (A; R_1, \ldots, R_s, S_1, \ldots, S_m)$ для сигнатуры σ_1. Теперь описанную выше процедуру можно применить к программе π_2, содержащей правила из множества P_2. Повторяя эту технику l раз, получим *каноническую модель* для всей стратифицированной программы π.

Легко доказывается

Предложение. Каноническая модель стратифицированной логической программы не зависит от ее стратификации.

3. КВАЗИАКСИОМАТИЧЕСКИЕ ТЕОРИИ ДЛЯ ПРАВДОПОДОБНЫХ РАССУЖДЕНИЙ ТИПА ДСМ

Применим указанную технику к простейшей квазиаксиоматической теории — теории простого ДСМ-метода правдоподобных рассуждений с одной активностью v.

Основным исходным понятием для ДСМ-рассуждения является понятие локального сходства. Чаще всего это понятие описывается через бинарную операцию, результаты которой вычисляются в ходе ДСМ-рассуждения. Однако так как модель конечна, можно предполагать, что она вычислена заранее, и мы будем ее описывать тернарным предикатом $z = x\Pi y$. Аналогично, определим унарный предикат "быть тривиальным объектом" $x = \Lambda$. Через $J_{<+1,0>}(X \Rightarrow_1 \{v\})$ обозначается унарный предикат: "Объект X проявляет активность v", через $J_{<-1,0>}(X \Rightarrow_1 \{v\})$ — "Объект X не проявляет активности v", а через $J_{<t,0>}(X \Rightarrow_1 \{v\})$ — "Проявляет ли объект X активность v, неизвестно". Этими предикатными символами образуется сигнатура.

Теперь простой ДСМ-метод с одной активностью v и без запрета на контрпримеры может быть задан следующей стратифицированной программой:

$$I_+(U, \{v\}) \leftarrow J_{<+1,0>}(X_1 \Rightarrow_1 \{v\}),$$
$$J_{<+1,0>}(X_2 \Rightarrow_1 \{v\}),\ U = X_1\Pi X_2,\ \neg U = \Lambda;$$

$$I_+(U, \{v\}) \leftarrow J_{<+1,0>}(X_1 \Rightarrow_1 \{v\}),$$
$$I_+(U_1, \{v\}), U = X_1\Pi U_1, \neg U = \Lambda;$$

$$I_-(U, \{v\}) \leftarrow J_{<-1,0>}(X_1 \Rightarrow_1 \{v\}),$$
$$J_{<-1,0>}(X_2 \Rightarrow_1 \{v\}), U = X_1\Pi X_2, \neg U = \Lambda;$$

$$I_-(U, \{v\}) \leftarrow J_{<-1,0>}(X_1 \Rightarrow_1 \{v\}),$$
$$I_-(U_1, \{v\}),\ U = X_1\Pi U_1, \neg U = \Lambda,$$

$$J_{<0,1>}(U \Rightarrow_2 \{v\}) \leftarrow I_+(U, \{v\}), I_-(U, \{v\});$$

$$J_{<+1,1>}(U \Rightarrow_2 \{v\}) \leftarrow I_+(U, \{v\}),$$
$$\neg J_{<0,1>}(U \Rightarrow_2 \{v\});$$

$$J_{<-1,1>}(U \Rightarrow_2 \{v\}) \leftarrow I_-(U, \{v\}),$$
$$\neg J_{<0,1>}(U \Rightarrow_2 \{v\});$$

$$J_{<0,1>}(X \Rightarrow_1 \{v\}) \leftarrow J_{<t,0>}(X \Rightarrow_1 \{v\}),$$
$$J_{<+1,1>}(U_1 \Rightarrow_2 \{v\}),\ J_{<-1,1>}(U_2 \Rightarrow_2 \{v\}),$$
$$U_1 = X\Pi U_1,\ U_2 = X\Pi U_2;$$

$$E_{<+1,1>}(X \Rightarrow_1 \{v\}) \leftarrow J_{<+1,0>}(X \Rightarrow_1 \{v\}),$$
$$J_{<+1,1>}(U \Rightarrow_2 \{v\},\ U = X\Pi U;$$

$$E_{<-1,1>}(X \Rightarrow_1 \{v\}) \leftarrow J_{<-1,0>}(X \Rightarrow_1 \{v\}),$$
$$J_{<-1,1>}(U \Rightarrow_2 \{v\}),\ U = X\Pi U;$$

$$J_{<+1,1>}(X \Rightarrow_1 \{v\}) \leftarrow J_{<t,0>}(X \Rightarrow_1 \{v\}),$$
$$J_{<+1,1>}(U \Rightarrow_2 \{v\}), U = X\Pi U,$$
$$\neg J_{<0,1>}(X \Rightarrow_1 \{v\});$$

$$J_{<-1,1>}(X \Rightarrow_1 \{v\}), \leftarrow J_{<t,0>}(X \Rightarrow_1 \{v\}),$$
$$J_{<-1,1>}(U \Rightarrow_2 \{v\}), U = X\Pi U,$$
$$\neg J_{<0,1>}(X \Rightarrow_1 \{v\});$$

$$J_{<t,1>}(X \Rightarrow_1 \{v\}) \leftarrow J_{<t,0>}(X \Rightarrow_1 \{v\}),$$
$$\neg J_{<+1,1>}(X \Rightarrow_1 \{v\}), \neg J_{<-1,1>}(X \Rightarrow_1 \{v\}),$$
$$\neg J_{<0,1>}(X \Rightarrow_1 \{v\}).$$

Унарные реляционные переменные $I_+(U, \{v\})$ $I_-(U, \{v\})$ служат для нахождения всех пересечений. Унарные реляционные переменные $J_{<+1,1>}(X \Rightarrow_2\{v\})$, $J_{<-1,1>}(X \Rightarrow_2\{v\})$ и $J_{<0,1>}(X \Rightarrow_2\{v\})$ представляют гипотезы о причинах наличия активности, причинах отсутствия активности и противоречивые гипотезы. $E_{<+1,1>}(X\Rightarrow_1\{v\})$ и $E_{<-1,1>}(X\Rightarrow_1\{v\})$ — унарные реляционные переменные, соответствующие проверке критерия достаточного основания правдоподобного вывода. Наконец, $J_{<+1,1>}(X\Rightarrow_2\{v\})$, $J_{<-1,1>}(X\Rightarrow_2\{v\})$, $J_{<0,1>}(X\Rightarrow_2\{v\})$ и $J_{<t,1>}(X\Rightarrow_2\{v\})$ — унарные реляционные переменные, описывающие положительное и отрицательное предсказание, противоречивый прогноз и отказ от прогноза. Несколько необычная символика была выбрана специально для лучшего сравнения с классическим изложением, основывающемся на многозначных логиках с J-операторами.

Аналогичные стратифицированные программы могут быть написаны и для всех остальных известных в настоящее время квазиаксиоматических теорий.

4. ЗАКЛЮЧЕНИЕ

Формализм стратифицированных логических программ обладает хорошей семантикой. Все известные в настоящее время квазиаксиоматические теории могут описываться изложенным формализмом. Существует значительное число результатов о выразительных возможностях языка стратифицированных логических программ на конечных реляционных моделях. Эти результаты могут