УДК 004.832.38 DOI: 10.17587/prin.7.330-336

А. Н. Шмарин, аспирант, e-mail: tim-shr@mail.ru, Воронежский государственный университет

О реализации приближения числа слоев без циклов в задаче нечеткого LP-вывода

Представлен способ вычисления числа слоев без циклов в LP-структуре на основе мощностей классов эквивалентности специального фактор-множества. Исследованы вопросы алгоритмической реализации соответствующих компонентов. Полученные результаты могут быть использованы для разработки интеллектуальных программных систем, минимизирующих число трудоемких запросов о состоянии предметной области при работе с базами знаний большого объема.

Ключевые слова: *LP-вывод, бинарное отношение, фактор-множество, базис циклов, NP-трудность, алгоритмы, Python*

Введение

С увеличением объемов и сложности обрабатываемой информации все более актуальным становится использование систем искусственного интеллекта. При этом одной из наиболее распространенных моделей представления знаний является продукционная модель [1]. Алгоритмы поиска решения в таких системах основаны на механизме логического вывода. На практике существует большое число задач, для которых получение информации о состоянии предметной области является ресурсоемкой операцией. Таковыми, например, могут быть запросы к СУБД, анализ медицинских данных или другая сложная последовательность вычислений. Зачастую в подобных задачах решение (либо множество возможных решений) ищут, зная лишь часть информации о состоянии предметной области. В силу данного обстоятельства требуется оптимально использовать ограниченные ресурсы.

Существуют задачи управления сложными объектами со случайными процессами (управление рисками, социально-экономическими системами), в которых управление основано на нечеткой модели представления знаний. Для них результат управляющего воздействия можно предсказать лишь с некоторой вероятностью. Кроме того, процесс получения информации, необходимой для принятия решения о том или ином управляющем воздействии, может быть сложным. В таких ситуациях необходимо поддерживать или приводить систему в заданное состояние, затрачивая как можно меньше ресурсов на трудоемкие операции получения необходимой информации, учитывая при этом нечеткость логических связей между знаниями.

С фактором ресурсоемкости операций получения данных, необходимых для принятия решения, также связана задача минимизации числа запросов о состоянии предметной области в ходе логического вывода. Данная задача является NP-трудной [2, 3]. Для достижения глобальной минимизации существует общий метод LP-вывода [4], к сожалению, обладающий экспоненциальной вычислительной сложностью относительно числа атомарных фактов в базе знаний. Идея кластерно-релевантного LP-вывода [5, 6] основана на вычислении специфических оценок (показателей релевантности) для ограниченного подмножества продукций и на обобщении полученных оценок на все множество продукций.

Ранее [7] были рассмотрены структурные свойства разбиения бинарного отношения, на основании которых можно выбирать для анализа такое подмножество продукций, которое наиболее существенно влияет на показатель релевантности. В настоящей работе продолжено исследование структурных свойств разбиения бинарного отношения на слои, а также рассмотрены вопросы алгоритмической и программной реализации метода вычисления приближенной оценки числа слоев без циклов в задаче нечеткого LP-вывода [8].

Основные определения

Пусть задано конечное множество $F = \{f\}$. На его основе определим атомно-порожденную ограниченную алгебраическую решетку $L = \lambda(F)$, где λ — функция-булеан [9]. На решетке зададим бинарное отношение R, являющееся каноническим [6], т. е. $R = \{r = (\tau, f)\colon \tau \in L, f \in F\}$.

Применительно к продукционным системам такое отношение R соответствует множеству продукций, F (множество атомов решетки) соответствует элементарным фактам продукционной системы, а сама решетка L соответствует множеству предпосылок и заключений продукций.

Обозначим символом A фактор-множество — разбиение бинарного отношения R на классы эквивалентности относительно уникальных правых частей его элементов, т. е. $A = \{\alpha_f = \{r = (\tau, f): r \in R\}: f \in F\}$.

Рассматриваемое фактор-множество служит основой для структурного расслоения [4] исходного отношения R на виртуальные слои (частичные отношения). Рассмотрим ряд вспомогательных функций.

Отображения $in(\alpha, f) = \{(\tau, g) \in \alpha : f \in \tau\}$ и $\overline{in}(\alpha, f) = \{(\tau, g) \in \alpha : f \notin \tau\}$ переводят класс эквивалентности α в множество пар, которые соответственно содержат либо не содержат в своих левых частях атом f. Следует заметить, что $in(\alpha, f) = \alpha \setminus \overline{in}(\alpha, f)$ и $\overline{in}(\alpha, f) = \alpha \setminus in(\alpha, f)$.

Функции in и $\overline{in}$ можно обобщить на все фактор-множество A:

$$in(A, f) = \left\{ \bigcup_{a \in A} in(\alpha, f) \right\}, \ \overline{in}(A, f) = \left\{ \bigcup_{a \in A} \overline{in}(\alpha, f) \right\}.$$

Отображение

$$\lambda(X, k, i): X \to \lambda(X), \ k = \overline{1, |X|}, \ i = \overline{1, \binom{|X|}{k}}$$

переводит множество X в i-й элемент множества всех его k-элементных подмножеств.

Определение 1. $A_f = \{\alpha \in A | \exists (\tau, g) \in \alpha : f \in \tau\} = \{\alpha \in A : in(\alpha, f) \neq \varnothing\}$ — отображение, переводящее фактор-множество A в такое подмножество его классов эквивалентности, что каждый элемент подмножества включает в себя хотя бы одну пару, содержащую f в левой части.

Используя введенные обозначения, можно переписать определенную в работе [7] функцию вычисления общего числа слоев для заданного атома f в виде

$$W(A, f, i) = \\ = \sum_{i}^{\binom{|A_f|}{i}} \left(\prod_{\beta \in in(\lambda(A_f, i, s), f)} |\beta| \prod_{\gamma \in \overline{in}(A \setminus \lambda(A_f, i, s), f)} |\gamma| \right).$$

В данном выражении левое произведение выполняется по мощностям элементов s-й выборки по i элементов из фактор-множества A. Они были прорежены до множеств, содержащих только пары, включающие атом f в своей левой части. При этом A_f исключает из произведения пустые множества. Правое же произведение выполняется по мощностям таких элементов A, которые не задействованы для вычисления левого произведения. Эти элементы прорежены до множеств, не содержащих пары, включающие атом f в левой части.

В работе [7] наглядно показано, что значение $W(A, f, i)$ тем больше, чем больше элементов в $\lambda(A\_f, i, s)$, т. е. чем больше значение биномиального коэффициента $\binom{|A\_f|}{i}$.

О программной реализации структур, соответствующих основным определениям

Для оптимизации последующей работы с бинарным отношением R целесообразно иметь возможность доступа к классам эквивалентности, входящим в A, по атому решетки или по его индексу: $A[f] = \alpha_f, \ f \in F, \quad A[i] = \alpha_{f_i}, \ f_i \in F, \ i = \overline{1, |F|}$. Сам же класс эквивалентности $\alpha \in A$ можно определить в виде вектора, содержащего индексы соответствующих пар в отношении. Если размер отношения позволяет сохранить его многократно в оперативной памяти как битовый вектор, то целесообразно классы эквивалентности определять в виде битовых векторов. Однако при таком определении необходимо сохранять возможность быстрого получения мощности соответствующего множества.

Результатами выполнения функций in и $\overline{in}$ являются классы эквивалентности, поэтому они должны представляться так же, как и классы эквивалентности $\alpha \in A$. Отображения A_f можно найти один раз для всех f в виде вектора списков индексов классов эквивалентности и хранить их в памяти в процессе работы метода.

Для генерации перестановок необходимо реализовать (или задействовать существующий) итерируемый генератор всех элементных подмножеств заданного множества. Также должна быть возможность использования сгенерированного множества несколькими объектами.

Процесс вычисления общего количества слоев можно попытаться оптимизировать путем кэширования компонентов произведений, состоящих из общих множителей.

Вычисление количества слоев без циклов

Ранее [7] было показано, что количество слоев без циклов для заданного атома решетки имеет смысл вычислять как разность количества всех слоев и слоев, содержащих циклы.

Общее количество слоев для атома f вычисляется как сумма $\sum_{i=1}^{|A_f|} W(A, f, i)$. Число слоев с циклами также целесообразно вычислять на основе суммы значений функции W, которой в качестве аргумента передаются фактор-множества, построенные специальным образом на основе исходного фактор-множества A и дополнительных данных о структурных свойствах бинарного отношения R.

Как показано в работе [7], для построения таких фактор-множеств можно использовать специальным образом сгенерированную фундаментальную систему (базис) циклов [10], на основе которой определяется новая структура — фундаментальная система циклов канонического бинарного отношения. Ее определение дано в работе [7]. Далее приведен алгоритм ее построения.

Алгоритм построения фундаментальной системы циклов канонического бинарного отношения

Шаг 1. На основе бинарного отношения R построить граф $G = \{v\}$:

$v \in G, v = (f,\ g), f, g \in F \Leftrightarrow \exists(\tau,\ g) \in R{:} f \in \tau.$

input: R
output: G
$G \leftarrow \varnothing$
foreach $r = (\tau,\ g) \in R$ **do**
  foreach $f \in \tau$ **do**
    if $v \notin (f,\ g)$ **then**
      $G \leftarrow G \cup \{v\}$
    end
  end
end

Шаг 2. По графу G, используя один из известных математических методов, построить фундаментальную систему циклов $S = \{s : s \subseteq G\}$.

Шаг 3. На основе отношения R и полученной фундаментальной системы циклов S построить E — *фундаментальную систему циклов канонического бинарного отношения.*

Построение заключается в том, что на основе дуг $v = (f,\ g),\ f,\ g \in F$, образующих элементы фундаментальной системы циклов $S = \{s\}$, строится множество $E = \left\{e_1 \ldots e_{|S|}\right\}$, состоящее из подмножеств отношения R, таких, что $\forall S_i = \{v = (f,\ g) \in G\} : e_i = \{(\tau,\ g) \in R{:} f \in \tau\}$.

input: A, S
output: E
$E \leftarrow \varnothing$
foreach $s \in S$ **do**
  foreach $v = (f,\ g) \in s$ **do**
    $e \leftarrow \varnothing$
    # найти все такие r, левая часть которых
    # содержит f,
    # а правая равна g: $r = (\tau,\ g) : f \in \tau$
    foreach $r = (\tau,\ g) \in A[g]$ **do**
      if $f \in \tau$ **then**
        $e \leftarrow e \cup \{r\}$
      end
    end
    $E \leftarrow E \cup \{e\}$
  end
end

Для единообразия операций хранение элементов E в машинной памяти целесообразно выполнять тем же способом, который используется для хранения классов эквивалентности фактор-множества A.

Вычисление количества всех слоев с циклами

Построения, приведенные далее, основаны на следующей идее. Из основного фактор-множества A можно выделить подмножество, описывающее все слои, содержащие все циклы из некоторого заданного множества фундаментальных циклов. Число таких слоев можно посчитать за линейное время как произведение мощностей его классов эквивалентности. На основе этого вычисления можно построить математическую структуру, описывающую число слоев, которые содержат циклы из заданного множества.

Пусть $\varepsilon \subseteq E$, $\varepsilon = \{e \in E\}$ — подмножество фундаментальных циклов канонического бинарного отношения R, $\dot{\varepsilon}(\varepsilon) = \bigcup\limits_{e \in \varepsilon} \bigcup\limits_{r \in e} r$ — объединение всех пар бинарного отношения, формирующих элементы данного подмножества.

Определение 2. Сужение фактор-множества A, описывающее всевозможные слои, которые содержат только фундаментальные циклы из множества ε, определяется следующим отображением:

$$\theta(\alpha, \varepsilon) = \begin{cases} \alpha \cap \dot{\varepsilon}(\varepsilon), & \text{если } \alpha \cap \dot{\varepsilon}(\varepsilon) \neq \varnothing \\ \alpha & \text{иначе} \end{cases},$$

$$\Theta(A, \varepsilon) = \left\{\theta(\alpha_1 \varepsilon),\ \ldots,\ \theta\left(\alpha_{|A|},\ \varepsilon\right)\right\}.$$

При таком построении имеем:

- $\Theta(A, \{e\})$, $e \in E$ — описывает все слои, включающие в себя цикл e;
- $\Theta(A, \{e,\ g\}), e,\ g \in E$ — описывает все слои, включающие в себя одновременно цикл e и цикл g; такой эффект достигается за счет неравенства $|e \cup g| > \max(|e|,\ |g|)$, благодаря чему α_i сужаются сильнее.

Для исследования структурных свойств данного отображения необходимо ввести дополнительные обозначения.

Определение 3. $S(A)$: $A \to \lambda(R)$ — отображение, переводящее фактор-множество $A = \{\alpha_1, ..., \alpha_n\}$ в множество всех слоев, которые A описывает:

$$S(A) = \left\{\left\{r_1, \ldots, r_{|A|}\right\} \middle| r_i \in \alpha_i, i = \overline{1, |A|}\right\} = \bigcup_{(r_1, \ldots, r_{|A|}) \in \alpha_1 \times \ldots \times \alpha_{|A|}} \left\{\left\{r_1, \ldots, r_{|A|}\right\}\right\}.$$

Если фактор-множество A состоит из классов эквивалентности $\alpha_1, ..., \alpha_n$, то число всевозможных слоев, которые можно построить на основе A, равно числу всевозможных выборок из каждого класса эквивалентности по одному элементу, т. е. $|S(A)| = |\alpha_1| \cdot \ldots \cdot |\alpha_n|$.

Определение 4. $S_{all}(A, \Theta, \varepsilon)$ — множество всевозможных слоев, включающих в себя все циклы из множества $\varepsilon \subseteq E$. При этом выполняется:

$$S_{all}(A,\ \Theta,\ \varepsilon) = S(\Theta(A,\ \varepsilon)), \varepsilon \subseteq E;$$

$$\left|S_{all}(A,\ \Theta,\ \varepsilon)\right| = |\theta(\alpha_1,\ \varepsilon)| \cdot \ldots \cdot \left|\theta\left(\alpha_{|A|},\ \varepsilon\right)\right|, A = \left\{\alpha_1,\ \ldots,\ \alpha_{|A|}\right\}.$$

Замечание 1. Следует обратить внимание, что если $\varepsilon \neq E$, то слои из $S_{all}(A, \Theta, \varepsilon)$ могут включать в себя не только циклы из множества ε, но и циклы из множества $E \backslash \varepsilon$. То есть из того, что слой включает в себя некоторое известное подмножество циклов, не следует, что слой не включает в себя какие-либо еще циклы.

Есть единственный случай, когда исходя из определения можно утверждать, что слой не включает в себя иных циклов, кроме заданных. Это случай, когда слой включает в себя все циклы (из $\varepsilon = E$ следует, что $E \backslash \varepsilon = \varnothing$).

Определение 5. $S_{subsets}(A, \Theta, \varepsilon)$ — множество всевозможных слоев, включающих в себя произвольное подмножество циклов из множества $\varepsilon \subseteq E$.

Замечание 2. Из определения S_{all} следует, что

$$\forall \varepsilon \subseteq E, \forall e \in \varepsilon : S_{all}(A, \Theta, \varepsilon) \subseteq S_{all}(A, \Theta, \{e\}).$$

Таким образом, $S_{subsets}$ можно выразить через S_{all}:

$$S_{subsets}(A, \Theta, \{x_1, \ldots, x_l\}) = \bigcup_{k=1}^{l} S_{all}(A, \Theta, \{x_k\}).$$

Определение 6. $S_{only}(A, \Theta, \varepsilon, E)$ — множество всевозможных слоев, которые включают в себя все циклы из ε и не включают циклов из $E \backslash \varepsilon$, т. е.

$$S_{only}(A, \Theta, \varepsilon, E) = \{l \in S(A) \mid (\forall e \in \varepsilon : e \subseteq l) \wedge$$
$$\wedge (\nexists g \in E \setminus \varepsilon : g \subseteq l \wedge g \neq \varnothing)\}.$$

Рассмотрим свойства данного определения.

Утверждение 1. $\forall e_1, e_2, E : e_1 \subseteq E, e_2 \subseteq E, e_1 \neq e_2$ выполняется следующее: $S_{only}(A, \Theta, \{e_1\}, E) \cap S_{only}(A, \Theta, \{e_2\}, E) = \varnothing$.

Доказательство. Для $e_1 \cap e_2 = \varnothing$ доказательство вытекает из определения. Рассмотрим случай $e_1 \cap e_2 = \varnothing$.

Предположим противное:

$S_{only}(A, \Theta, \{e_1\}, E) \cap S_{only}(A, \Theta, \{e_2\}, E) \neq \varnothing$.

Тогда

$\exists l : l \in S_{only}(A, \Theta, \{e_1\}, E) \wedge l \in S_{only}(A, \Theta, \{e_2\}, E)$.

Следовательно, $e_1 \subseteq l \wedge e_2 \subseteq l$. Но из того, что $e_1 \neq e_2$, следует $(E \setminus \{e_1\}) \cap \{e_2\} \neq \varnothing \wedge (E \setminus \{e_1\}) \cap \{e_2\} \subseteq l$, а это невозможно по определению.

Утверждение 2.

$\forall \varepsilon \subseteq E, \varepsilon \neq \varnothing : S_{only}(A, \Theta, \varepsilon, \varepsilon) = S_{all}(A, \Theta, \varepsilon)$. Очевидно, что $\varepsilon \backslash \varepsilon = \varnothing$. Кроме того, по определению слои из множества $S_{only}(A, \Theta, \varepsilon, \varepsilon)$ не включают в себя никаких иных циклов, кроме тех, что принадлежат ε. А это соответствует определению $S_{all}(A, \Theta, \varepsilon)$.

Замечание 3. Если $\varepsilon \subseteq E$, $\varepsilon \neq \varnothing$, то, чтобы получить множество всех слоев, которые включают в себя все циклы из ε и не включают в себя никаких циклов из $E \backslash \varepsilon$, нужно из $S_{all}(A, \Theta, \varepsilon)$ — множества всех слоев, включающих в себя все циклы из множества ε, удалить все слои, включающие в себя циклы из множества $E \backslash \varepsilon$. Таким образом, имеет место следующее соотношение:

$$S_{only}(A, \Theta, \varepsilon, E) = S_{all}(A, \Theta, \varepsilon) \setminus$$
$$\setminus \left(\bigcup_{k=1}^{|E \backslash \varepsilon|} \bigcup_{j=1}^{\binom{|E \backslash \varepsilon|}{k}} S_{only}(A, \Theta, \varepsilon \cup \lambda(E \backslash \varepsilon, k, j), E) \right).$$

Теперь, используя введенные обозначения, можно выразить значение общего числа слоев с циклами $|S_{subsets}(A, \Theta, E)|$ в терминах мощностей множеств S_{all} и S_{only}:

$$S_{subsets}(A, \Theta, E) = \bigcup_{\varepsilon \in E} S_{only}(A, \Theta, \varepsilon, E) =$$
$$= \bigcup_{k=1}^{|E|} \bigcup_{i=1}^{\binom{|E|}{k}} S_{only}(A, \Theta, \lambda(E, k, i), E);$$
$$|S_{subsets}(A, \Theta, E)| = \sum_{\varepsilon \in E} |S_{only}(A, \Theta, \varepsilon, E)| =$$
$$= \sum_{k=1}^{|E|} \sum_{i=1}^{\binom{|E|}{k}} |S_{only}(A, \Theta, \lambda(E, k, i), E)|.$$

Для практического исследования полученных соотношений будет использован скрипт на языке Python. На основе рекуррентного соотношения, определяемого утверждением 2 и замечанием 3, по заданному множеству циклов генерируется сокращенное выражение $|S_{subsets}(A, \Theta, E)|$ в виде сумм $|S_{all}(A, \Theta, E)|$, $\varepsilon \subseteq E$.

```
#!/usr/bin/env python3.4

from sympy import Symbol
from itertools import combinations

def s_all(eps): return Symbol("".join(sorted(eps)))

def s_only(eps, E):
  if len(eps) = = len(E):
    return s_all(eps)
  else:
    l, sdiff = [], sorted(set(E) — eps)
   for i in range(1, len(sdiff) + 1):
   l + = [s_only(eps | set(perm), E)
         for perm in combinations(sdiff, i)]
  return s_all(eps) — (sum(l))

  def s_subsets(E):
  E, l = sorted(E), []
  for i in range(1, len(E) + 1):
   l + = [s_only(set(perm), E)
    for perm in combinations(E, i)]
  return sum(l)

if __name__=="__main__":
 for i in ["ab", "abc", "abcd"]:
  print("{0:5}: {1}".format(i, str(s_subsets(i))))
```

Далее приведены соответствующие результаты генерации для систем из двух, трех и четырех циклов.

- $|S_{subsets}(A, \Theta, \{a, b\})| = |S_{all}(A, \Theta, \{a\})| +$
  $+ |S_{all}(A, \Theta, \{b\})| - |S_{all}(A, \Theta, \{a, b\})|;$

- $\left|S_{subsets}\left(A,\Theta,\{a,b,c\}\right)\right| = \left|S_{all}\left(A,\Theta,\{a\}\right)\right| + {}$
$+ \left|S_{all}\left(A,\Theta,\{b\}\right)\right| + \left|S_{all}\left(A,\Theta,\{c\}\right)\right| - {}$
$- \left|S_{all}\left(A,\Theta,\{a,b\}\right)\right| - \left|S_{all}\left(A,\Theta,\{a,c\}\right)\right| - {}$
$- \left|S_{all}\left(A,\Theta,\{b,c\}\right)\right| + \left|S_{all}\left(A,\Theta,\{a,b,c\}\right)\right|;$
- $\left|S_{subsets}\left(A,\Theta,\{a,b,c,d\}\right)\right| = \left|S_{all}\left(A,\Theta,\{a\}\right)\right| + {}$
$+ \left|S_{all}\left(A,\Theta,\{b\}\right)\right| + \left|S_{all}\left(A,\Theta,\{c\}\right)\right| + {}$
$+ \left|S_{all}\left(A,\Theta,\{d\}\right)\right| - \left|S_{all}\left(A,\Theta,\{a,b\}\right)\right| - {}$
$- \left|S_{all}\left(A,\Theta,\{a,c\}\right)\right| - \left|S_{all}\left(A,\Theta,\{a,d\}\right)\right| - {}$
$- \left|S_{all}\left(A,\Theta,\{b,c\}\right)\right| - \left|S_{all}\left(A,\Theta,\{b,d\}\right)\right| - {}$
$- \left|S_{all}\left(A,\Theta,\{c,d\}\right)\right| + \left|S_{all}\left(A,\Theta,\{a,b,c\}\right)\right| + {}$
$+ \left|S_{all}\left(A,\Theta,\{a,b,d\}\right)\right| + \left|S_{all}\left(A,\Theta,\{a,c,d\}\right)\right| + {}$
$+ \left|S_{all}\left(A,\Theta,\{b,c,d\}\right)\right| - \left|S_{all}\left(A,\Theta,\{a,b,c,d\}\right)\right|.$

В полученных выражениях видна закономерность, которую можно обобщить на произвольное число циклов.

Утверждение 3. Для произвольной фундаментальной системы циклов $E = \{x_1, ..., x_l\}$ канонического бинарного отношения число всевозможных слоев с циклами вычисляется по формуле

$$\left|S_{subsets}\left(A,\Theta,E\right)\right| = \sum_{x_1\in E}\left|S_{all}\left(A,\Theta,\{x_1\}\right)\right| - \sum_{\substack{\{x_1,x_2\}\subset E\\ x_1\neq x_2}}\left|S_{all}\left(A,\Theta,\{x_1,x_2\}\right)\right| + \ldots$$
$$\ldots + (-1)^{l+1}\left|S_{all}\left(A,\Theta,\{x_1,\ldots,x_l\}\right)\right| = \sum_{k=1}^{l}(-1)^{k+1}\sum_{i=1}^{\binom{l}{k}}\left|S_{all}\left(A,\Theta,\{\lambda(E,k,i)\}\right)\right|.$$

Корректность утверждения на небольших размерах данных (ограниченных производительностью компьютера) можно проверить с помощью приведенного скрипта. Формальное доказательство корректности утверждения для произвольных объемов является пока открытым вопросом и не рассматривается в данной работе.

Далее выполним построение, описывающее число всех слоев с циклами, которые i раз содержат заданный атом f решетки L.

Определение 7. Будем говорить, что сужение θ корректно для A_1 и ε, если выполняются следующие условия:

$$S\left(\Theta\left(A_1,\,\varepsilon\right)\right) \subseteq S\left(A\right); \tag{1}$$

$$\forall l \in S(\Theta(A_1,\,\varepsilon))\,\exists e \in \varepsilon : e \subseteq l; \tag{2}$$

$$\forall e \in \varepsilon\,\exists l \in S\left(\Theta\left(A_1,\,\varepsilon\right)\right) : e \subseteq l. \tag{3}$$

Пусть задано некоторое фактор-множество A_1, для которого выполняется $S(A_1) \subseteq S(A)$, и некоторое ε такое, что $\varepsilon \subseteq E$, причем сужение Θ не обязательно корректно для A_1 и ε.

Определение 8. $F(A_1, \Theta, \varepsilon) : \lambda(E) \to \lambda(E)$ — отображение, переводящее множество циклов ε в такое его подмножество, что сужение Θ корректно для A_1 и $F(A_1, \Theta, \varepsilon)$, т. е. $F(A_1, \Theta, \varepsilon) = \left\{e \in \varepsilon \,|\, \exists l \in S\left(\Theta\left(A_1,\varepsilon\right)\right) : e \subseteq l\right\}$.

При таком построении условия (1) и (2) выполняются по определению отображения Θ, а условие (3) выполняется по определению отображения F.

Для практического применения необходимо переписать данное определение в терминах фактор-множеств:

$$F\left(A_1,\,\Theta,\,\varepsilon\right) = \left\{e \in \varepsilon \,\middle|\, \forall \alpha_e \in e/\sim\ \exists \alpha_{A_1} \in A_1/\sim : \alpha_e \subseteq \alpha_{A_1}\right\}.$$

Теперь можно определить формулу для вычисления точного количества слоев без циклов. Для этого введем вспомогательное отображение

$$\tilde{A}\left(i,\,s,\,f\right) = in\left(\lambda\left(A_f,\,i,\,s\right),f\right) \cup \overline{in}\left(A \backslash \lambda\left(A_f,\,i,\,s\right),\,f\right).$$

Кроме того, целесообразно переписать в более кратком виде определение функции для нахождения количества всех слоев:

$$W\left(A,\,f,\,i\right) = \sum_{s=1}^{\binom{|A_f|}{i}}\ \prod_{\beta\in\tilde{A}(i,s,f)}|\beta|.$$

Таким образом, используя введенные обозначения, можно построить функцию, возвращающую количество таких слоев с циклами, что каждый слой содержит ровно i пар с атомом f, включенным в левую часть:

$$W_{loops}(A,\,f,\,i) = \sum_{s=1}^{\binom{|A_f|}{i}}\left|S_{subsets}\left(\tilde{A}\left(i,\,s,\,f\right),\Theta,F(\tilde{A}\left(i,\,s,\,f\right),\,\Theta,\,\varepsilon)\right)\right|.$$

По аналогии с предыдущими рассуждениями количество всех слоев, которые содержат циклы для заданного атома f, можно найти в виде суммы

$$W_{loops}\left(A,\,f\right) = \sum_{i=1}^{|A_f|}W_{loops}\left(A,\,f,\,i\right).$$

Для нахождения числа слоев без циклов достаточно из общего числа слоев вычесть число слоев с циклами:

$$W_{noloops}\left(A,\,f\right) = W\left(A,\,f\right) - W_{loops}\left(A,\,f\right);$$

$$W_{noloops}\left(A,\,f\right) = \sum_{i=1}^{|A_f|}\sum_{s=1}^{\binom{|A_f|}{i}}\left(\prod_{\beta\in\tilde{A}(i,s,f)}|\beta| - \left|S_{subsets}\left(\tilde{A}\left(i,\,s,\,f\right),\Theta,F\left(\tilde{A}\left(i,\,s,\,f\right),\,\Theta,\,\varepsilon\right)\right)\right|\right).$$

О реализации вычисления приближенного значения числа слоев без циклов

Можно заметить наличие некоторых закономерностей, полезных для вычисления приближенного значения количества слоев без циклов.

Значения $W(A, f, i)$ и $W_{loops}(A, f, i)$ тем больше, чем больше значение биномиального коэффициента $\binom{|A_f|}{i}$. Известно, что наибольшее значение всегда имеет коэффициент $\binom{|A_f|}{\lceil |A_f|/2 \rceil}$ и, если $|A_f|$ — нечетное число, коэффициент $\binom{|A_f|}{\lceil |A_f|/2 \rceil + 1}$. Однако данные значения растут экспоненциально относительно $|A_f|$. Этот факт означает, что для достаточно больших значений $|A_f|$ не получится найти точные значения функций при i, близких к $\lceil |A_f|/2 \rceil$. Однако данный факт можно использовать для ранжирования атомов по степени их релевантности, а приближенные значения W и W_{loops} находить для небольших биномиальных коэффициентов.

Значение $|S_{all}(A, \Theta, \varepsilon)|$ тем больше, чем меньше пар с уникальными правыми частями содержится в ε (т. е. чем меньше $\left|\bigcup_{e \in \varepsilon} e/\sim\right|$) и чем больше элементов отношения R присутствует в элементах ε. Такая закономерность имеет место в силу особенностей отображения Θ, которое прореживает элементы фактор-множества A до пересечения с множеством $\bigcup_{e \in \varepsilon} e$, если такое пересечение не пусто.

Для приближенного подсчета $|S_{subsets}(A, \Theta, \varepsilon)|$ целесообразно в первую очередь выбирать те ε, что содержат меньшее число элементов, и среди таких ε в первую очередь при подсчете учитывать те, которые содержат большее число пар отношения R, т. е. имеют большее значение $\left|\bigcup_{e \in \varepsilon} e\right|$.

Кроме того, при реализации приближения функции $W_{noloops}$ целесообразно считать компоненты суммы, дающие наибольшие значения при наименьшем числе суммируемых членов, и выполнять за шаг приближение по всем f.

Заключение

Исследованы вопросы приближенного вычисления первого показателя релевантности LP-вывода — числа прообразов, содержащих заданный атом решетки. Формализован подход, позволяющий при нахождении приближенной оценки учитывать в первую очередь наиболее количественно значимые структурные особенности бинарного отношения на решетке.

Данный подход может быть применен для ускорения обратного вывода в интеллектуальных системах продукционного типа. Он может быть распространен и на более сложные логические системы в информатике [11].

Работа поддержана грантом РФФИ № 15-07-05341.

Список литературы

1. **Davis R., King J.** An overview of production systems // Machine Intelligence. 1977. Vol 8. P. 300—332.
2. **Махортов С. Д., Шмарин А. Н.** Нечеткий LP-вывод и его программная реализация // Программная инженерия. 2013. № 12. С. 34—38.
3. **Махортов С. Д., Шмарин А. Н.** Оптимизация метода LP-вывода // Нейрокомпьютеры. Разработка, применение. 2013. № 9. С. 59—63.
4. **Махортов С. Д.** Основанный на решетках подход к исследованию и оптимизации множества правил условной системы переписывания термов // Интеллектуальные системы. Теория и приложения. 2009. Т. 13. С. 51—68.
5. **Махортов С. Д.** Интегрированная среда логического программирования LPExpert // Информационные технологии. 2009. № . 12. С. 65—66.
6. **Болотова С. Ю., Махортов С. Д**. Алгоритмы релевантного обратного вывода, основанные на решении продукционно-логических уравнений // Искусственный интеллект и принятие решений. 2011. № 2. С. 40—50.
7. **Шмарин А. Н., Махортов С. Д**. О приближенных оценках количества слоев без циклов в задаче нечеткого LP-вывода // Нейрокомпьютеры. Разработка, применение. 2015. № 12. С. 44—51.
8. **Shmarin A. N.** Using Confidence Coefficients in the LP Inference // Modern informatization problems in simulation and social technologies: Proceedings of the XX-th International Open Science Conference. Yelm, WA, USA, January 1—10, 2015 / Editor in Chief Dr. Sci., Prof. O. Ja. Kravets. Yelm: Science Book Publishing House LLC, 2015. P. 196—201.
9. **Салий В. Н., Богомолов В. А.** Алгебраические основы теории дискретных систем. М.: Физматлит, 1997. 368 с.
10. **Харари Ф.** Теория графов: пер. с англ. СПб.: БХВ-Петербург., 2003. 296 с.
11. **Чечкин А. В.** Нейрокомпьютерная парадигма информатики // Нейрокомпьютеры: разработка, применение. 2011. № 7. С. 3—9.

About Implementation of Approximation the Number of Layers without Cycles in the Fuzzy LP-Inference Problem

A. N. Shmarin, e-mail: tim-shr@mail.ru, Voronezh State University, Voronezh, 394006, Russian Federation,

Corresponding author: **Shmarin Artem N.**, Postgraduate Student, Voronezh State University, Voronezh, 394006, Russian Federation, e-mail: tim-shr@mail.ru

Received on April 08, 2016
Accepted on April 19, 2016

The increase in the volumes and complexity of processed information makes the use of artificial intelligence systems more and more actual. In such systems, one of the most widespread knowledge representation models is the production model and search algorithms are based on the inference engine. In practice, obtaining information about the subject area status is time-consuming operation for many problems. The examples are the requests to the DBMS, the analysis of medical data or other complex sequences of calculations can be those. Often in such problems, the search for solution (or set of possible solutions) is performed in the absence of part of information about the subject area status. This circumstance requires optimal use of limited resources.

There are tasks of control of complex objects with random processes (risk management, socio-economic systems), in which the control process is based on fuzzy knowledge representation model. For them, the result of the controlling influence can be predicted only with some probability. In addition, process of obtaining information that is necessary for making decision about the control action can be time-consuming operation. In such situations it is necessary to maintain or to bring the system to the predetermined state, spending as little as possible resources for time-consuming operations of obtaining the necessary information, taking into account at the same time the fuzziness of the relationship between knowledge.

The complexity of the operations of data acquisition necessary for decision-making, leads to the problem of minimizing the number of requests for information about the subject area status during inference. This problem is NP-hard. To achieve global minimization there is the general method of LP-inference with exponential computational complexity relative to the number of atomic facts in the knowledge base. However, some of its heuristic modifications have polynomial complexity. The goal of the provided research is development of the approximating method to better minimize the number of requests performed in the inference.

Earlier, the structural properties of partition of the binary relation were considered, based on which it is possible to select for the analysis such subset of production, that most significantly changes the rate of relevancy.

This paper presents the method for calculating the value of the exact number of layers without cycles based on capacities of quotient set equivalence classes that builds as the partition of binary relations by unique right parts. In addition, this paper considers questions of the algorithmic implementation of components that are necessary to calculate the approximate value of the number of layers without cycles.

The presented approach can be applied to accelerate the reverse inference in production type systems of artificial intelligence. It can be extended also to more complex logical systems in computer science.

Keywords: *LP-inference, quotient set, binary relation, cycle basis, NP-hardness, algorithms, python*

Acknowledgements: *This work was supported by the Russian Foundation for Basic Research, project nos. 15-07-05341.*

For citation:

Shmarin A. N. About Implementation of Approximation the Number of Layers without Cycles in the Fuzzy LP-Inference Problem, *Programmnaya Ingeneria*, 2016, vol. 7, no. 7, pp. 330—336.

DOI: 10.17587/prin.7.330-336

References

1. **Davis R., King J.** An overview of production systems, *Machine Intelligence*, 1977, vol. 8, pp. 300—332.
2. **Makhortov S. D., Shmarin A. N.** Nechetkij LP-vyvod i ego programmnaja realizacija (Fuzzy LP-inference and its software implementation), *Programmnaya Ingeneria*, 2013, no. 12. pp. 34—38 (in Russian).
3. **Makhortov S. D., Shmarin A. N.** Optimizacija metoda LP-vyvoda (Optimizing LP-inference method), *Nejrokomp'jutery. Razrabotka, primenenie,* 2013, no. 9, pp. 59—63 (in Russian).
4. **Makhortov S. D.** Osnovannyj na reshetkah podhod k issledovaniju i optimizacii mnozhestva pravil uslovnoj sistemy perepisyvanija termov (Research and optimization of a set of rules in a conditional term rewriting system using approach based on the lattices), *Intellektualnyie Sistemyi,* 2009, no. 13, pp. 51—68 (in Russian).
5. **Makhortov S. D.** Integrirovannaja sreda logicheskogo programmirovanija LPExpert (LPExpert: Integrated development environment for logical programming), *Informaczionny'e Texnologii,* 2009, no. 12, pp. 65—66 (in Russian).
6. **Bolotova S.Yu., Makhortov, S. D.** Algoritmy relevantnogo obratnogo vyvoda, osnovannye na reshenii produkcionno-logicheskih uravnenij (Algorithms of the relevant backward inference that is based on production-logic equations solving), *Iskusstvennyiy Intellekt i Prinyatie Resheniy*, 2011, no. 2, pp. 40—50 (in Russian).
7. **Shmarin A. N., Makhortov S. D.** O priblizhennyh ocenkah kolichestva sloev bez ciklov v zadache nechetkogo LP-vyvoda (About approximate estimates of the number of layers without cycles in the fuzzy LP-inference), *Nejrokomp'jutery. Razrabotka, primenenie*, 2015, no. 12, pp. 44—51 (in Russian).
8. **Shmarin A. N.** Using confidence coefficients in the LP-inference, *Modern Informatization Problems in the Technological and Telecommunication Systems Analysis and Synthesis: Proceedings of the XX-th International Open Science Conference*, Yelm, WA, USA, January 1—10, 2015/ Editor in Chief Dr. Sci., Prof. O. Ja. Kravets, 2015, pp. 196—202.
9. **Salij V. N. Bogomolov V. A.** *Algebraicheskie osnovy teorii diskretnyh sistem* (Algebraic foundations of the theory of discrete systems), Moscow, Fizmatlit, 1997, 368 p. (in Russian).
10. **Harari F.** *Teorija grafov:* Per. s angl. (Graph theory: Trans. from engl.) SPb.: BHV-Peterburg, 2003, 296 p. (in Russian).
11. **Chechkin A. V.** Nejrokomp'juternaja paradigma informatiki (Neurocomputing paradigm of informatics), *Nejrokomp'jutery: Razrabotka, Primenenie*, 2011, no. 7, pp. 3—9 (in Russian).

ООО "Издательство "Новые технологии". 107076, Москва, Стромынский пер., 4
Технический редактор *Е. М. Патрушева.* Корректор *З. В. Наумова*

Сдано в набор 10.05.2016 г. Подписано в печать 20.06.2016 г. Формат 60×88 1/8. Заказ PI716
Цена свободная.

Оригинал-макет ООО "Адвансед солюшнз". Отпечатано в ООО "Адвансед солюшнз".
119071, г. Москва, Ленинский пр-т, д. 19, стр. 1. Сайт: **www.aov.ru**