

ПРИМЕНЕНИЕ ПОДХОДА К ПОСТРОЕНИЮ ПЛАНА ПОИСКА В БАЗЕ ЗНАНИЙ НА ОСНОВЕ СТОИМОСТИ ОПЕРАЦИЙ

Ивашенко В. П., Кальчевский Д. А.

Кафедра интеллектуальных информационных технологий,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: ivashenko@bsuir.by

Рассматривается задача поиска по образцу в базе знаний и подход к её решению, основанный на методах теории расписаний и ориентированный на составление оптимизированного плана поиска. Для представления знаний используются однородные семантические сети на основе модели унифицированного семантического представления знаний.

ВВЕДЕНИЕ

В интеллектуальных системах, к которым относятся системы, основанные на знаниях, и системы, управляемые знаниями [1], одной из ключевых задач является задача поиска в базе знаний [2]. Это задача актуальна и для систем, использующих унифицированное семантическое представление знаний [1,2]. Задача поиска описывается с помощью поискового запроса, который константами описывает данные, известные в исходном состоянии, и описывает переменными данные, которые могут быть найдены или не найдены в каком-либо из целевых состояний. Запросы бывают простыми и сложными [3,4]. Простые состоят из одного образца поиска, а сложные – из нескольких. Образец поиска задаётся структурой семантической сети, обратное отношение отношению интерпретации которой является функцией. Далее рассмотрим задачу построения плана поиска для заданного образца поиска.

I. ЗАДАЧА ПОИСКА

Пусть задан образец поиска фрагмента семантической сети модели унифицированного семантического представления знаний:

- V – множество sc-элементов;
- E – множество связей отношения инцидентности sc-элементов ($E \subseteq V \times V$);
- T – матрица типов sc-элементов (размерность $|V| \times d$);
- M – матрица стоимостей операций поиска по связкам отношения инцидентности (размерность $d \times d \times s \times s$);
- W – матрица весов операций поиска по связкам отношения инцидентности (размерность $d \times d \times s$);
- μ – метрическая форма (размерность $s \times s$);
- c – мера стоимости.

Требуется построить оптимальный план поиска для одного исполнителя, где оптимизируемая функция – мера стоимости c .

Пусть $\mathbf{1}_S^V$ обозначает вектор значений индикаторной функции элементов S на V , который

имеет размер $|V|$. Определим функцию $c_i(p)$

$$c_i(p) = \sum_{j=1}^i \sum_{q=1}^j f(\langle p(j)_1, p(j)_2, p(q)_1, p(q)_2 \rangle),$$

где план поиска:

$$p \in \{r \mid r^{-1} \in \mathbf{N}^E\} \cap E^{\{n \mid n \leq |E|\}} \cap \mathbf{N},$$

а функция f :

$$f(\langle \chi, \gamma, \omega, v \rangle) = h(\langle \mathbf{1}_{\{\chi\}}^V * \mathbf{1}_{\{\gamma\}}^V, \mathbf{1}_{\{\omega\}}^V * \mathbf{1}_{\{v\}}^V \rangle)$$

Также пусть

$$h(\langle \beta^{mn}, \lambda^{rt} \rangle) = g(\lambda^{rt} * T_r^i * T_t^j * \beta^{mn} * W_{mn}^k)$$

и

$$g(\alpha^{ijk}) = (\alpha^{ijk} * M_{ijk}^p) * (\alpha^{ijk} * M_{ijk}^q) * \mu_{pq}$$

Тогда задача может быть сведена к задаче теории расписаний [5] $1|tree| \sum_{i=1}^n w_i * C_i$, для решения которой на каждом шаге достаточно проверить условие

$$f(\langle x, y, u, v \rangle) \leq f(\langle u, v, x, y \rangle)$$

или неотрицательность выражения

$$\Delta(\langle x, y, u, v \rangle) = f(\langle u, v, x, y \rangle) - f(\langle x, y, u, v \rangle)$$

Оптимизация ориентирована на то, чтобы как можно раньше выполнить большее количество проверок, что увеличит вероятность раннего завершения поиска в случае отсутствия очередных искомых данных при условии, что это отсутствие для любых искомых данных равновероятно. Также для запросов, образцы которых имеют циклическую структуру связей отношения инцидентности, оптимизация ориентирована на сокращение временных издержек поиска, связанных с разбиением образца поиска на ациклический образец поиска и граф проверок при условии, что операции проверки имеют стоимость заведомо меньшую стоимости любой операции поиска, и распределением направлений поиска для ациклического образца поиска, когда стоимости поиска в прямом и обратном направлении различны.

II. АЛГОРИТМ СОСТАВЛЕНИЯ ПЛАНА ПОИСКА

Для решения рассмотренной задачи рассмотрим следующий алгоритм, который получает на вход исходные данные, а на выходе возвращает два множества S и C .

1. Добавить все константы во множество $F(0)$ вершин с найденными значениями.
2. Множество рёбер $E(0) = E^{-1} \cup E$.
3. Множество $S(0) = \emptyset, t = 0$.
4. Для каждого ребра из $E(t) \cap (F(t) \times F(t))$ выполнить нижеописанную подпрограмму, перенести все рёбра из множества $E(t) \cap (F(t) \times F(t))$ во множество $C(t)$.
5. Проверить является ли непустым множество $M(t)$, если оно пустое, то перейти на пункт 9.

$$M(t) = (E(t)/C(t)) \cap (F(t) \times (V/F(t)))$$

6. Из множества $M(t)$ выбрать ребро $\langle x, y \rangle$ такое, что

$$0 \leq \min_{\langle u, v \rangle \in M(t)/\{\langle x, y \rangle\}} \{\Delta(\langle x, y, u, v \rangle)\}$$
7. В $F(t+1)$ добавить все вершины из $F(t)$ вместе с y , перенести во множество $S(t+1)$ ребра из $S(t)$ вместе с $\langle x, y \rangle$, ребро $\langle y, x \rangle$ исключить из $E(t)$ в $E(t+1)$.
8. Увеличить $t \leftarrow t+1$ и перейти на пункт 4.
9. Вернуть $\langle S(t), C(t) \rangle$.

Подпрограмма рассматривает две встречные ветви графа поиска $S(t)$ и заменяет в нём поиск значения вершины образца по направлению в первой ветви на поиск значения вершины образца во встречном направлении в ветви последней операции поиска, пока не выполняется некоторое условие, например:

$$w_i * C_{i-1} + \sum_{k=i}^z w_k * m_i \leq w_j * C_{i-1} + \sum_{k=j}^z w_k * m_j,$$

где i, j – номера операции поиска в прямом и во встречном направлениях ($z \geq j$), w – веса, m, C – времена и времена завершения операций. Исключённые дуги из графа поиска добавляются во множество $E(t)$, дуги в графе поиска и встречные им исключаются из множества $E(t)$, t увеличивается на соответствующее количество шагов. Возвращаются обновлённые $E(t), S(t), t$.

Например, есть образец поиска (Рис. 1), который может быть представлен графом [6]. В результате применения алгоритма можно получить следующий результат (Рис. 2) при условии задания некоторых значений T, M, W и μ .

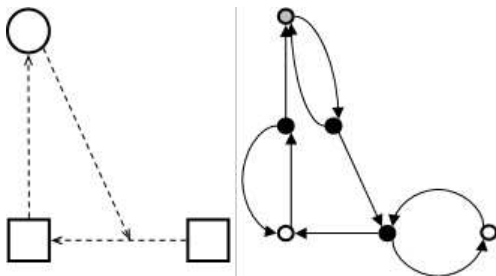


Рис. 1 – Образец поиска и графовое представление

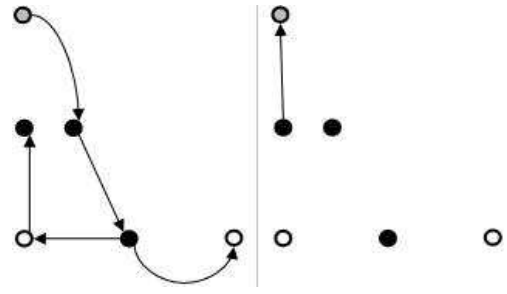


Рис. 2 – Дерево поиска и граф проверок (справа)

ЗАКЛЮЧЕНИЕ

Была проведена реализация разработанных алгоритмов средствами C/C++. Проведено тестирование и вычислительные эксперименты, подтверждающие работоспособность реализованных алгоритмов, разработан транслятор графов поиска на язык SCP [2]. Разработанные алгоритмы ориентированы на использование в рамках интеграционной платформы [1].

1. Ivashenko, V. Application of an integration platform for ontological model-based problem solving using an unified semantic knowledge representation / V. Ivashenko // Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2021) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол. : В. В. Голенков [и др.]. – Минск, 2021. – Вып. 5. – С. 179–186.
2. Программирование в графодинамических ассоциативных машинах / В. В. Голенков [и др.] – Минск : БГУИР, 2001, – 412 с.
3. Иващенко, В.П. Язык описания синтаксических правил для однородных семантических сетей / В.П. Иващенко // Дистанционное обучение – образовательная среда XXI века : материалы VI Междунар. науч. метод. конф., Минск, 22–23 нояб. 2007 г. / Белорус. гос. ун-т информатики и радиоэлектроники. – Минск, 2007. – С. 185–188.
4. Ivashenko, V. P. Application of optimization methods for search of semantic network constructions / V. P. Ivashenko // ECCO XVIII: Conference of the European chapter on combinatorial optimization: combinatorics for modern manufacturing, logistics and supply chains, Minsk, May 26–28, 2005 / Belarus. State Univ. ; ed. M.Y. Kovalyov. – Minsk, 2005. – P. 23–24.
5. Tanaev, V. S. Scheduling Theory: Single-Stage Systems / V. S. Tanaev, V. S. Gordon, Y. M. Shafransky. – Dordrecht; Boston; London: Kluwer Academic Publishers, 1994. – 380 p.
6. Иващенко, В. П. Представление семантических сетей и алгоритмы их организации и семантической обработки на вычислительных системах с массовым параллелизмом / В. П. Иващенко и другие // Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2015) : материалы V междунар. науч. техн. конф. (Минск, 19–21 февраля 2015 года) / редкол. : В. В. Голенков (отв. ред.) [и др.]. – Минск : БГУИР, 2015. – С. 133–140.