

АЛГОРИТМИЧЕСКИЕ ОСНОВЫ ПОСТРОЕНИЯ ФРАКТАЛА ПРОДУКЦИОННОГО ЗНАНИЯ

Мухамедиева Дилноз Тулкуновна 1,

Рустамов Ербол Насимович 2,

Васиева Дилфузадилшод кизи 3

1Профессор кафедры Цифровых технологий и искусственного интеллекта,
Национальный исследовательский университет “Ташкентский институт инженеров
иригации и механизации сельского хозяйства”

Email: dilnoz123@rambler.ru , Orcid ID: 0000-0002-4561-4474

2Исследователь кафедры Цифровых технологий и искусственного интеллекта,
Национальный исследовательский университет “Ташкентский институт инженеров
иригации и механизации сельского хозяйства”

Email: rustamoverbol24@gmail.com , Orcid ID: 0009-0000-1976-0059

3Докторант кафедры Цифровых технологий и искусственного интеллекта,
Национальный исследовательский университет “Ташкентский институт инженеров
иригации и механизации сельского хозяйства”

Email: vasiyevadilfuza1997@gmail.com

Аннотация

В статье рассматриваются алгоритмические основы построения продукционного знания и базы знаний в интеллектуальных информационных системах. Основное внимание уделяется формализации предметной области и алгоритмам преобразования первичной информации в продукционные знания. Показано, что база знаний является семантической моделью предметной области, формируемой на основе выделения глобальных и локальных свойств признаков. Описывается общий принцип построения продукционного знания с использованием допустимых таблиц, корреляционного анализа и фрактальной организации знаний. Предложенный подход обеспечивает объективность получения знаний и создает теоретическую и алгоритмическую основу для разработки интеллектуальных систем поддержки принятия решений.

Ключевые слова: интеллектуальные системы, база знаний, продукционное знание, предметная область, формализация знаний, алгоритмы обработки знаний, семантика знаний.

Introduction

Введение

Современный этап развития информационных технологий характеризуется возрастанием роли интеллектуальных информационных систем, ориентированных на обработку не просто данных, а знаний. В условиях цифровизации научной и производственной деятельности особую актуальность приобретает задача построения эффективных баз знаний, способных отражать семантику предметной области и

поддерживать процессы анализа, вывода и принятия решений. При этом ключевой проблемой остается алгоритмизация процедуры получения знаний из исходной информации [1-3].

Работа посвящена алгоритмам построения производственного знания и базы знаний как формализованной модели предметной области. В рамках данного подхода предметная область рассматривается как совокупность объектов, обладающих набором свойств, которые интерпретируются в различных контекстах. Первичные сведения о свойствах предметной области представляются в виде данных, организованных в допустимую таблицу, которая служит информационной моделью объекта исследования [4-5].

Переход от информационной модели к модели знаний требует выявления скрытых закономерностей и устойчивых связей между признаками. Для этого используются алгоритмы статистической обработки, в частности корреляционный анализ, позволяющий оценить степень взаимосвязи свойств объекта. В результате формируются глобальные и локальные характеристики признаков, на основе которых строится фрактал производственного знания. Такой фрактал отражает иерархическую и импликативную структуру знаний о предметной области [6-8].

Таким образом, алгоритмы, рассматриваемые в работе, направлены на формализацию и объективное порождение знаний, что является необходимым условием для создания интеллектуальных информационных систем нового поколения [9-11].

2. Методы. Термин “производство” принадлежит американскому логику Э. Поту. В понимании Пота, в качестве производств выступала только та ее часть, которую теперь называют ядром. Запись “ЕСЛИ А, ТО В” трактовалась как оператор замены источника А цепочкой В в некотором выходном слове, т.е. производств были тем, что позже стали называть подстановками и использовать при описании различных уточнений понятия алгоритма.

В общем виде, под *производством* понимается выражение следующего вида [1]:

$$\Phi_i^p : A \rightarrow B; N.$$

Здесь i – имя *производства*, с помощью которого данная производств выделяется из всего множества производств, p – контекст где интерпретируется $\tilde{I}^\Sigma(\delta_i(x_i))$ и в качестве имени может выступать некоторая лексема, отражающая суть данной производств (например, “покупка книги” или “набор видов замков”) и порядковый номер производств в их множестве, хранящийся в памяти системы.

Элемент Φ_i^p характеризует *сферу применения производств*, т.е. контекст, где интерпретируется $I^p(x)$. Такие сферы легко выделяются в когнитивных структурах человека. Наши знания как бы “разложены по полочкам”. На одной “полочке” хранятся знания о том, как надо готовить пищу, на другой – как добраться до работы и т.п. Разделение знаний на отдельные сферы в БЗ целесообразно и при представлении знаний производственных моделей. Процедура проектирования производственной БЗ начинается созданием фрактала производственного знания. Для этого нужно будет, в первую очередь,

как уже отметили, выбрать из $\delta_j \dot{\delta} X$, имеющиеся свойства, чтобы классифицировать знание. Алгоритмизация этой процедуры требует разработки принципов выбора таких $\delta_j \dot{\delta} X$. Уже после этого можно будет проектировать производственное знание. На основе проектирования фрактала производственного знания строится производственная БЗ для выбранной ПО. С другой стороны, сам алгоритм представления в реляционном (табличном) виде требует разработки методов такого представления. При этом надо обязательно разработать алгоритм обработки знаний, находящихся в БЗ.

В этой работе будем более практически описывать одну из возможных алгоритмических технологий создания производственной БЗ для систем искусственного интеллекта. Опишем алгоритм, с помощью которого можно решить многочисленные практические задачи, связанные с выделением информативных подсистем $\delta_j \dot{\delta} X$ признаков. Главное его назначение связано с выделением $\delta_j \dot{\delta} X$, имеющих глобальные и локальные свойства (алгоритм A), участвующих в построении фрактала производственного знания

$$Q^m : \{\Phi_k\} = (B_u^i \rightarrow \sum_{c=1}^{g(i)} b_c^{u,j})_k \quad (1)$$

Алгоритм отбора $\delta_j \dot{\delta} X$ с целью построения фрактала производственного знания, когда при этом использовались самые информативные по рангу признаки в группе входных и выходных (конечно, если мы можем исследуемый объект или явление представить в виде кибернетического черного ящика) параметров процесса. B_u^i строился в виде отношения выделенных признаков. Эту функцию назовем глобальным порождающим базисом ГПБ фрактала производственного знания Q^m .

Опишем алгоритм выбора признаков для построения $\sum_{c=1}^{g(i)} b_c^{u,j}$ локального порождающего базиса. В свою очередь эту функцию назовем локальным порождающим базисом (ЛПБ) фрактала производственного знания Q^m . После оценки значения векторов \vec{R}_1 и \vec{R}_2 для $\delta_j \dot{\delta} X$ исследуемой ПО, представленного в виде допустимой таблицы $T_{mpt}^{\xi^*}$, оцениваются глобальные свойства и локальные свойства $\delta_j \dot{\delta} X$ и строятся вектора \vec{R}_1 и \vec{R}_2 . На основе значений этих векторов выбираются признаки для ГПБ и ЛПБ.

Алгоритм A включает следующие этапы:

1-й шаг. Вычисление для всех признаков значения функции

$$\gamma_{ij} = \begin{cases} 1, & \text{если } \rho_i(j, c) \geq \bar{\rho}_{\tau_{\mathbb{C}}}, \\ 0, & \text{в противном случае;} \end{cases}$$

$$\gamma_j = \sum_{i=1}^n \gamma_{j,i}; \quad j, i = \overline{1, n}.$$

Здесь $\bar{\rho}_{t\odot}$ - среднее арифметическое среди всех значений, компонент вектора \vec{R}_1 , т.е.

$$\bar{\rho}_{t'} = \frac{\sum_{t=1}^{t'} \rho_t(j, c)}{t'}; \quad t' = C_m^2.$$

2-й шаг. Выбор $\delta_j(x_i)$ признаков для построения ГПБ. Рассмотрим случай, когда глобальный базис строится на основании только входных параметров. Из целесообразных соображений нам следует ограничить сверху лишь число тех признаков, из которых строится ГПБ (идеально, их всего 2 или 3). Допустим, что вид ГПБ найден. Следующий шаг заключается в нахождении значений, принимаемых объектом обучения на глобальном базисе, и несколько типов. Последнее реорганизует классы объектов и устанавливает на упорядоченных объектах новый квази-порядок, являющийся неким способом разбиения классов на группы объектов, схожих по доминирующему входным параметрам. Здесь надо отметить, если исследуемой ПО можно представить в виде черного ящика, тогда наиболее простой вид ГПБ представляет собой отношение для двух признаков и неравенства для трех признаков:

$$B = \left\{ \frac{\delta(x_{i1})}{\delta(x_{i2})} \right\}; \quad (2)$$

$$B = \{ \delta(x_{i1}) \leq \delta(x_{i2}) \leq \delta(x_{i3}); \delta(x_{i1}) \leq \delta(x_{i3}) \leq \delta(x_{i2}); \dots \}. \quad (3)$$

Далее на каждом из объектов классов K_1, K_2, \dots, K_c либо (в случае двух признаков) вычисляется $B(S_1) = \delta_1, \dots, B(S_m) = \delta_m$, либо (если три признака) выполняется одно из неравенств

$$B(S_1) = \delta_1, \dots, B(S_m) = \delta_m$$

Нужно заметить, что в случае двух-признакового базиса значения ГПБ (на объектах обучения) разбиваются на m типов (в случае с тремя параметрами разбиваются автоматически на число типов неравенств, представленных в ГПБ).

Таким образом, класс K_j разбивается на подклассы $K_{j,u}$, в которых сгруппированы объекты с j -м типом ГПБ:

$$K_j = \bigcup_u^{m\odot} K_{j,u}; \quad K_{j,l} \bigcap K_{j,c} = \emptyset; \quad j = 1, 2, \dots, l; \quad c = 1, 2, \dots, m.$$

3-й шаг. Выбор из группы важных признаков, требуемых для построения ЛПБ. Самая важная часть в ситуационном управлении имеет обобщение по структурам. К одному классу относятся ситуации, в которых для объекта управления принимаются одинаковые решения.

Результат

Для решения задач выявления закономерностей применяются процедуры, основанные на схемах индуктивной логики, С. Миля и названные методами индуктивных рассуждений. Рассмотрим, в частности, метод сходства [1]:

$$a, b, c \Rightarrow a$$

$$a, b, f \Rightarrow a$$

$$a, g, h \Rightarrow a$$

(4)

...

$$a, l, m \Rightarrow a$$

$$\underline{a \Rightarrow a}$$

т.е. при наличии посылки a следует α .

Задачи выявления закономерностей схожи по постановке с задачами построения фрактала продукционного знания.

Распишем задачу выбора признаков, необходимых для построения локальных базисов, методом сходства как задачу распознавания образов (PO). Для этой цели рассмотрим на примере группы объектов, имеющих U_1 -ый тип ГПБ. Пусть представлены два класса K_1 и K_2 , в каждом из которых по два объекта типа U_1 , т.е.

$$K_{1,u_1} = \begin{cases} S_1: \Phi_1^1 a_{1i_1}, a_{1i_2}, \dots, a_{1i_m}, \dots, a_{1i_{m+1}} \\ S_2: \Phi_2^1 a_{2i_1}, a_{2i_2}, \dots, a_{2i_m}, \dots, a_{2i_{m+1}} \end{cases}$$

$$K_{2,u_1} = \begin{cases} S_3: \Phi_3^2 a_{3i_1}, a_{3i_2}, \dots, a_{3i_m}, \dots, a_{3i_{m+1}} \\ S_4: \Phi_4^2 a_{4i_1}, a_{4i_2}, \dots, a_{4i_m}, \dots, a_{4i_{m+1}} \end{cases}$$

Проверим признак $i' \in \{i_1, i_2, \dots, i_{m+1}\}$.

Согласно определению (2), имеем следующий вид:

$$\{a_{ij}\} \sim B_j; i=1, 2, \dots, n_1; i \in \{i_1, i_2, \dots, i_{m+1}\} \setminus i';$$

$$a_{ji'} = a_1; j=1, 2;$$

$$a_{ki'} = a_2; k=3, 4;$$

$$K_1 \sim a_1;$$

$$K_2 \sim a_2;$$

Таким образом (2), с учетом (3) и (4) перепишется, как:

$$\begin{array}{c|cc|c} S_1 & a_1 & B_1 & \alpha_1 \\ S_2 & a_1 & B_2 & \alpha_1 \\ S_3 & a_2 & B_3 & \alpha_2 \\ S_4 & a_2 & B_4 & \alpha_2 \end{array}$$

Если для a_1, a_2 и a_3, a_4 выполняется условие б) или в) из (4), то $a=1, b=1, \alpha=1, a_i=1, B_j=1, a_i=1, i=\overline{1,2}, j=\overline{3,4}$ и $a, b \Rightarrow a, a_i, b_j \Rightarrow a_i$.

Рассмотрим на конкретном примере годность i' -го признака для построения ЛПБ по следующим этапам:

1. Отбор значений объектов класса K_1 группы u_1 , т.е. $K_{1,u_1} : S_1$ и S_2 по признаку i' , а именно $a_{1,i'} = a_1; a_{2,i'} = a_2$ и класса K_2 группы u_1 : $a_{3,i'} = a_3; a_{4,i'} = a_4$.

Пусть условно $\alpha_1 < \alpha_2$ и $\alpha_3 < \alpha_4$.

2. Проверка взаимного расположения этих значений. Здесь возможны несколько случаев:

а) $(\alpha_1 < \alpha_2) < (\alpha_3 < \alpha_4)$;

б) $(\alpha_3 < \alpha_4) < (\alpha_1 < \alpha_2)$.

3. Тогда в случае, когда имеем а) и б), признак i' i' годен для построения некоего ЛПБ в группе u_1 классов K_1 и K_2 , а в случае в) он не подходит.

4. Построение локального базиса. Самый простой способ сводится к тому, что сами значения признака объектов из группы u_1 классов K_1 и K_2 участвуют в построении ЛПБ, а разделяющей границей является α'' :

$$\alpha'' = \frac{\alpha_3 - \alpha_2}{2} + \alpha_3 \quad \text{для а);}$$

и

$$\alpha'' = \frac{\alpha_1 - \alpha_4}{2} + \alpha_4 \quad \text{для б).}$$

Резюмируя, мы хотели сказать следующее. Алгоритмический отбор признаков или отношений признаков для построения ГПБ и ЛПБ является самым ответственным моментом построения фрактала продукционного знания для ПО. Мы предложили одну из возможных процедур для построения ГПБ и ЛПБ, но это далеко не последняя процедура. Когда инженер знаний моделирует сложный объект, не всегда ему удается алгоритмически построить ГПБ и ЛПБ, особенно, когда априорная информация собрана из качественных признаков. В этом случае построения ГПБ и ЛПБ во многом зависят от профессионального уровня эксперта, с которым работает инженер знаний.

Заключение

В работе рассмотрены алгоритмические принципы построения продукционного знания и базы знаний, основанные на формализации семантики предметной области. Показано, что база знаний представляет собой не простое хранилище информации, а структурированную модель предметной области, сформированную в результате обработки и интерпретации данных. Предложенный алгоритмический подход позволяет последовательно осуществлять переход от сведений к данным, от данных к информации и далее — к знаниям. Выделение глобальных и локальных свойств признаков обеспечивает объективность формирования продукционных правил и отражает семантические закономерности исследуемого объекта. Использование корреляционного анализа и допустимых таблиц позволяет автоматизировать процесс получения знаний и повысить достоверность результатов. Полученные положения создают теоретическую и методологическую основу для проектирования интеллектуальных информационных систем различного назначения, включая системы диагностики, прогнозирования и

поддержки принятия решений. Рассмотренные алгоритмы могут служить базой для дальнейших исследований в области интеллектуальных технологий и инженерии знаний.

Литература

1. Поспелов Д.А. Представление знаний. Искусственный интеллект. Кн.г. Модели и методы: Справочник. / -М: Радио и связь, 1980. – 304с.
2. Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, 16(1), 3-9.
3. Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107-136.
4. Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Harvard Business Press.
5. Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford University Press.
6. Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163-180.
7. Russell, S. J., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Prentice Hall.
8. Tuomi, I. (1999). Data is more than knowledge: Implications of the reversed knowledge hierarchy for knowledge management and organizational memory. *Journal of Management Information Systems*, 16(3), 103-117.
9. Walker, L. O., & Avant, K. C. (2011). *Strategies for theory construction in nursing* (5th ed.). Prentice Hall.
10. Wang, Y. (2015). Formal cognitive models of data, information, knowledge, and intelligence. *WSEAS Transactions on Computers*, 14, 770-781.
11. Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, 58(4), 479-493.
12. Tulkunovna D. M., Rustamov E. Semantic Representation Of Production Knowledge //2022 International Conference on Information Science and Communications Technologies (ICISCT). – IEEE, 2022. – C. 1-4.
13. Rustamova E., Muhamediyeva D. T., Safarova L. U. Development and management of product knowledge base //Proc. of SPIE Vol. – 2024. – T. 13065. – C. 130650W-1.