

УДК 004.89

О.И. Морозова, А.Ю. Соколов, В.М. Хуссейн

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков

МЕТОД НЕЧЕТКОГО СТРУКТУРНОГО АНАЛИЗА ОНТОЛОГИЙ

Рассматривается представление знаний в виде онтологии для улучшения процесса обучения. В предлагаемом подходе к решению задачи, каждый обучаемый добавляет отредактированную онтологию относительно своего видения структуры рассматриваемой информации. На основании онтологий, полученных сообществом обучаемых, формируется так называемая идеальная онтология представления информации. Для сравнения построенных онтологий с идеальной онтологией используются графы онтологий. Для построения графов онтологий, в данной работе предлагается использовать метод нечеткого структурного анализа онтологий.

Ключевые слова: онтология, текстовое представление знаний, нечеткие множества, метод нечеткого структурного анализа, матрица полудостижимости, онтологический граф, множества уровней.

Введение

На сегодняшний день в системах обучения активно используются такие подходы, в которых основным являются данные, представленные в электронном виде, а также всевозможные системы тестирования и самооценки обучаемого.

В тоже время, знания, получаемые обучаемыми, формируются индивидуально и являются их персональной информацией, тем самым возникает сложность размещения такой информации в компьютерной среде.

В данной работе предлагается представление знаний в виде онтологии [1 – 3]. Построение онтологии позволяет достаточно просто определить наиболее значимые для представления знаний концепты и связи между ними. Представление знаний в виде онтологии является непростым процессом с точки зрения не только большого количества различных источников информации, но и в связи с различным пониманием обучаемыми значимости понятий. В связи с этим, предлагается формирование онтологического представления информации не только преподавателями и экспертами в рассматриваемой предметной области, но и самими обучаемыми.

Таким образом, возникает идея использовать модель самоорганизации для процесса формирования онтологии [4]. Другими словами, каждый обучаемый добавляет отредактированную собственную онтологию относительно своего видения структуры рассматриваемой информации.

На основании онтологий, полученных сообществом обучаемых, формируется так называемая идеальная онтология представления информации. Идеальная онтология является динамической базой знаний, которая постоянно совершенствуется и развивается. При формировании онтологии понятиям присваиваются веса. Согласно значениям весов, онтология реализуется в виде графического представления иерархического расположения понятий, называемого графом онтологии.

Однако на практике возникает необходимость минимизировать расстояние между онтологией, составленной определенным обучаемым и идеальной онтологией. В данной работе этот процесс осуществляется посредством сравнения графов онтологий. Для построения таких графов предлагается использовать метод нечеткого структурного анализа онтологий [5].

1. Предварительные замечания для метода нечеткого структурного анализа онтологий

На основе теории нечетких множеств [6], предлагается метод нечеткого структурного анализа для представления онтологии в виде графа онтологии.

Важным требованием для структурного анализа сложных систем является представление необходимых данных в таком виде, что графическое представление может быть разработано достаточно легко. Таким образом, основной целью этого метода является описание и иллюстрация формальной процедуры для построения графического представления иерархического расположения элементов с учетом необходимой информации относительно отношения каждого элемента к каждому другому элементу. В нашем случае таким элементом является понятие.

Для получения графического представления онтологии, необходимо сформировать множества различных уровней. Для этого применяют процедуру получения нечеткой матрицы полудостижимости, из которой затем выделяют необходимые множества. Данная процедура разрешает автоматическое развитие графической структуры, которая отображает иерархию. Приведем несколько свойств по отношению к нечетким подмножествам и нечетким отношениям [7].

Определение 1. Пусть матрица B и вектор Y , размерностью $n \times n$ и n соответственно, являются нечеткой матрицей и нечетким вектором, элементы которых b_{ij} , y_j принадлежат закрытому интервалу $[0, 1]$. Тогда нечеткой композицией C для матрицы B и вектора Y называется следующие выражение

$$c_i = \bigvee_{j=1}^n (b_{ij} \wedge y_j), \quad i = 1, 2, \dots, n \quad (1)$$

где c_i является элементом S .

Определение 2. Нечеткая логическая сумма D и нечеткое логическое произведение E матриц A и B определяются следующим образом:

$$d_{ij} = a_{ij} \vee b_{ij} \quad i, j = 1, 2, \dots, n \Leftrightarrow D = A \oplus B; \quad (2)$$

$$e_{ij} = a_{ij} \wedge b_{ij} \quad i, j = 1, 2, \dots, n \Leftrightarrow E = A \otimes B. \quad (3)$$

Под операцией дизъюнкции в нечетких отношениях понимается операция \max для двух элементов.

Определение 3. Нечеткое бинарное (двоичное) отношение и его дополнение в прямом произведении пространств S и S ($S \times S$), характеризуется соответственно отношениями f_R и $f_{\bar{R}}$ следующим образом

$$\left. \begin{aligned} f_R : S \times S &\rightarrow [0, 1]; \\ f_{\bar{R}} : S \times S &\rightarrow [0, 1]; \end{aligned} \right\} \quad (4)$$

где отношения f_R и $f_{\bar{R}}$ задается выражением

$$f_{\bar{R}}(s_i, s_j) = (1 - f_R(s_i, s_j)) / (1 + \lambda f_R(s_i, s_j)), \quad (5)$$

где параметр λ является действительным числом из интервала $-1 < \lambda < \infty$.

Пусть p действительное число, принадлежащее полуоткрытому интервалу $(0, 1]$. В соответствии с таким числом p , далее приведены некоторые определения относительно нечетких бинарных отношений.

Определение 4. (Нечеткий рефлексивный закон). Когда $f_R(s_i, s_i) \geq p$ для $\forall (s_i, s_i) \in S \times S$, тогда такое отношение называется нечетко рефлексивное.

Определение 5. (Нечеткий иррефлексивный (нерефлексивный) закон). Когда $f_R(s_i, s_i) < p$ для $\forall (s_i, s_i) \in S \times S$, тогда такое отношение называется нечетко иррефлексивное.

Определение 6. (Нечеткий симметричный закон). Когда $f_R(s_i, s_j) \geq p$ и $f_R(s_j, s_i) \geq p$ для $\forall s_i, s_j \in S (i \neq j)$, тогда такое отношение называется нечетко симметричное.

Определение 7. (Нечеткий асимметричный закон). Когда одно из $f_R(s_i, s_j)$ или $f_R(s_j, s_i)$ для $\forall s_i, s_j \in S (i \neq j)$ меньше чем p , тогда такое отношение называется нечетко асимметричное.

Определение 8. (Нечеткий транзитивный закон). Пусть s_i, s_j и $s_k \in S (i \neq j \neq k)$. Когда $f_R(s_i, s_k) > \max_{s_i} \left[\min(f_R(s_i, s_j), f_R(s_j, s_k)) \right]$ для любого $(s_i, s_j), (s_j, s_k)$ и (s_i, s_k) , тогда такое отношение называется нечетко транзитивное.

Определение 9. (Нечеткий полутранзитивный закон). Пусть $m_{ik} = \bigvee_{j=1}^n (f_R(s_i, s_j) \wedge f_R(s_j, s_k)) \geq p$

для $\forall (s_i, s_j), (s_j, s_k) \in S \times S (i \neq j \neq k)$. Когда $f_R(s_i, s_k) \geq m_{ik}$ для любого (s_i, s_k) , тогда такое отношение называется нечетко полутранзитивное.

2. Метод нечеткого структурного анализа онтологий

Пусть онтология представляет собой множество понятий $S = \{s_1, s_2, \dots, s_n\}$. Кроме этого, построена нечеткая матрица подчиненности A , которая представляет нечеткие отношения подчиненности между элементами S на основе определенного контекстуального отношения

$$A = [a_{ij}], \quad i, j = 1, 2, \dots, n, \quad (6)$$

где A – квадратная матрица $n \times n$; a_{ij} – элемент матрицы A , который задается нечетким бинарным отношением f_R , следующим образом:

$$a_{ij} = f_R(s_i, s_j), \quad 0 \leq a_{ij} \leq 1, \quad i, j = 1, 2, \dots, n. \quad (7)$$

Это показывает степень, с которой s_i подчиняется s_j . Для того, чтобы показать, что степень подчинения больше, чем при определенной степени, мы вводим параметр p в качестве порога. Как упоминалось ранее, значение p должно быть взято из полуоткрытого интервала $(0, 1]$.

Определение 10. Если выполняется нечеткий полутранзитивный закон в нечеткой матрице подчиненности A , тогда матрица A называется нечеткой матрицей полудостижимости.

Определение 11. Множество верхнего уровня $L_t(s)$, множество промежуточного уровня $L_i(s)$, множество нижнего уровня $L_b(s)$, и множество изоляционного уровня $L_{is}(s)$, соответственно, определяются следующим образом:

$$L_t(s) = \left\{ s_i \mid \bigvee_{j=1}^n a_{ij} < p \leq \bigvee_{j=1}^n a_{ji} \right\}; \quad (8)$$

$$L_i(s) = \left\{ s_j \mid \bigvee_{k=1}^n a_{kj} \geq p, \bigvee_{k=1}^n a_{jk} \geq p \right\}; \quad (9)$$

$$L_b(s) = \left\{ s_i \mid \bigvee_{j=1}^n a_{ji} < p \leq \bigvee_{j=1}^n a_{ij} \right\}; \quad (10)$$

$$L_{is}(s) = \left\{ s_j \mid \bigvee_{k=1}^n a_{kj} < p, \bigvee_{k=1}^n a_{jk} < p \right\}. \quad (11)$$

Каждый элемент множества верхнего уровня не подчиняется никому другому элементу. Каждый элемент множества промежуточного уровня подчиняется каждому другому элементу и имеет некоторое подчинение элементов к себе. Каждый элемент множества нижнего уровня не имеет подчинения элементов к се-

бе. Каждый элемент множества изоляционного уровня не подчиняется никому другому элементу, и не имеет никакого подчинения элементов к себе.

Пример таких множеств показан на рис. 1. Множество верхнего уровня должно быть отделено в каждой иерархии заранее.

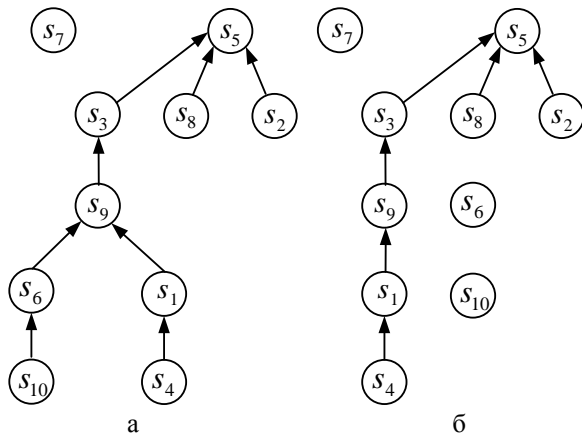


Рис. 1. Граф онтологии

Итак, для того, чтобы выделить множества различных уровней необходимо получить нечеткую матрицу полудостижимости понятий.

Рассмотрим процедуру получения нечеткой матрицы полудостижимости, которая состоит из следующих шагов.

Шаг 1: Степенные матрицы для матрицы подчиненности A вычисляются следующим образом:

$$A^1 = A, A^2 = A^1 \circ A, \dots, A^i = A^{i-1} \circ A, \dots \quad (12)$$

Пусть A^* является логической суммой степенных матриц A^1, A^2, \dots, A^n и определяется по формуле:

$$A^* = \bigoplus_i A^i = A^1 \oplus A^2 \oplus \dots \oplus A^n. \quad (13)$$

Шаг 2: Если элемент a_{ij}^* матрицы A^* меньше порога p , тогда a_{ij}^* равно нулю.

Шаг 3: Матрица, полученная от логической суммы A и A^* , становится нечеткой матрицей полудостижимости A' .

$$A' = A \oplus A^*, \quad a_{ij}' = a_{ij} \vee a_{ij}^*. \quad (14)$$

На основе матрицы полудостижимости понятий становится возможным описать контекстное отношение между понятиями и создать структурную графическую объяснительную модель.

3. Пример использования метода для решения задачи анализа онтологий

Рассмотрим пример формирования идеальной онтологии и построения графа онтологии с помощью метода структурного анализа онтологий.

Для выделения множеств уровней необходимо задать порог. В данном случае избран порог $p = 0,4$. Обучаемыми было выделено 10 основных понятий:

- 1) основание;
- 2) отношение;
- 3) рассуждение;
- 4) аргумент;
- 5) логика;
- 6) тезис;
- 7) диалектика;
- 8) высказывание;
- 9) доказательство;
- 10) утверждение.

Веса понятиям присвоены в соответствии с результатами обработки экспертных оценок, которые получены в процессе опытной эксплуатации системы двумя группами студентов. Веса, соответствующие понятиям, приведены в табл. 1.

Таблица 1

Матрица подчиненности понятий ($p = 0,4$)

№	1	2	3	4	5	6	7	8	9	10
1	0,0	0,2	0,0	0,2	0,7	0,0	0,1	0,2	0,0	0,2
2	0,7	0,0	0,7	0,2	0,7	0,7	0,1	0,2	0,7	0,2
3	0,0	0,2	0,0	0,2	0,7	0,0	0,1	0,2	0,0	0,2
4	0,7	0,2	0,7	0,0	0,7	0,7	0,1	0,2	0,7	0,2
5	0,3	0,2	0,3	0,2	0,0	0,3	0,1	0,2	0,3	0,2
6	0,0	0,2	0,0	0,2	0,7	0,0	0,1	0,2	0,0	0,2
7	0,1	0,1	0,1	0,1	0,1	0,1	0,0	0,1	0,1	0,1
8	0,7	0,2	0,7	0,2	0,7	0,7	0,1	0,0	0,7	0,2
9	0,0	0,2	0,0	0,2	0,7	0,0	0,1	0,2	0,0	0,2
10	0,7	0,2	0,7	0,2	0,7	0,7	0,1	0,2	0,7	0,0

Применив процедуру получения нечеткой матрицы полудостижимости, получаем следующие множества уровней:

$$L_t(s) = \{s_5\}; \quad L_b(s) = \{s_2, s_4, s_8, s_{10}\};$$

$$L_i(s) = \{s_1, s_3, s_6, s_9\}; \quad L_{is}(s) = \{s_7\}.$$

На основе полученных множеств формируется граф онтологии. На рис. 1, а приведено графическое изображение полученных множеств уровней.

Однако обучаемый мог присвоить другие веса понятиям, но при этом его онтология будет соответствовать идеальной. Для этого необходимо увеличить задаваемый порог. Для матрицы подчиненности понятий приведенной в табл. 2, порог p задается, например, равным 0,5. При этом граф онтологии соответствует графу, приведенному на рис. 1, а.

Таблица 2

Матрица подчиненности понятий ($p = 0,5$)

№	1	2	3	4	5	6	7	8	9	10
1	0,0	0,3	0,0	0,3	0,8	0,0	0,1	0,3	0,0	0,3
2	0,8	0,0	0,8	0,3	0,8	0,8	0,1	0,3	0,8	0,3
3	0,0	0,3	0,0	0,3	0,8	0,0	0,1	0,3	0,0	0,3
4	0,8	0,3	0,8	0,0	0,8	0,8	0,1	0,3	0,8	0,3
5	0,4	0,3	0,4	0,3	0,0	0,4	0,1	0,3	0,4	0,3
6	0,0	0,3	0,0	0,3	0,8	0,0	0,1	0,3	0,0	0,3
7	0,1	0,1	0,1	0,1	0,1	0,1	0,0	0,1	0,1	0,1
8	0,8	0,3	0,8	0,3	0,8	0,8	0,1	0,0	0,8	0,3
9	0,0	0,3	0,0	0,3	0,8	0,0	0,1	0,3	0,0	0,3
10	0,8	0,3	0,8	0,3	0,8	0,8	0,1	0,3	0,8	0,0

Как упоминалось ранее, на практике возникает необходимость минимизировать расстояние между онтологией, составленной определенным обучаемым и идеальной онтологией. Предлагается осуществить этот процесс посредством сравнения графов онтологии. В качестве примера рассмотрим ситуацию, когда обучаемый составил матрицу подчиненности понятий, приведенную в табл. 3. Понятия 6 и 10 обучаемый не выделил как основные. Порог берем такой же, как и для идеальной онтологии ($p = 0,4$).

Таблица 3

Вторая матрица подчиненности понятий ($p = 0,4$)

№	1	2	3	4	5	6	7	8	9	10
1	0,0	0,2	0,0	0,2	0,7	0,0	0,1	0,2	0,0	0,0
2	0,7	0,0	0,7	0,2	0,7	0,0	0,1	0,2	0,7	0,0
3	0,0	0,2	0,0	0,2	0,7	0,0	0,1	0,2	0,0	0,0
4	0,7	0,2	0,7	0,0	0,7	0,0	0,1	0,2	0,7	0,0
5	0,3	0,2	0,3	0,2	0,0	0,0	0,1	0,2	0,3	0,0
6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
7	0,1	0,1	0,1	0,1	0,1	0,0	0,0	0,1	0,1	0,0
8	0,7	0,2	0,7	0,2	0,7	0,0	0,1	0,0	0,7	0,0
9	0,0	0,2	0,0	0,2	0,7	0,0	0,1	0,2	0,0	0,0
10	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Для данной матрицы подчиненности понятий получаем следующие множества уровней:

$$L_t(s) = \{s_5\}; L_b(s) = \{s_2, s_4, s_8\};$$

$$L_1(s) = \{s_1, s_3, s_9\}; L_{is}(s) = \{s_6, s_7, s_{10}\}.$$

Как видно, понятия 6 и 10 переместились во множество изоляционного уровня $L_{is}(s)$. На рис. 1.6 приведено графическое построение полученных множеств уровней. На основе полученных множеств формируется другой граф онтологии. Затем осуществляем сравнение 4 множеств уровней для каждого из графов. В результате получаем два массива: массив, показывающий, каких элементов не хватает в графе для того, чтобы приблизить его к графу идеальной онтологии и массив, показывающий какие элементы лишние и их нет в графе идеальной онтологии. Преобразование онтологической модели в представление онтологии в виде графа, а затем во множественную модель позволяет решить задачу сравнения двух онтологий путем операций сравнения множеств.

МЕТОД НЕЧІТКОГО СТРУКТУРНОГО АНАЛІЗУ ОНТОЛОГІЙ

О.І. Морозова, О.Ю. Соколов, В.М. Хуссейн

У роботі розглядається подання знань у вигляді онтології для поліпшення процесу навчання. У запропонованому підході до вирішення завдання, кожен учень додає відредаговану онтологію щодо свого бачення структури розглянутої інформації. На підставі онтологій, отриманих спільною групою, що навчаються, формуються так звані ідеальна онтологія подання інформації. Для порівняння побудованих онтологій з ідеальною онтологією використовуються графи онтологій. Для побудови графів онтологій, в даній роботі пропонується використовувати метод нечіткого структурного аналізу онтологій.

Ключові слова: онтологія, текстове представлення знань, нечіткі множини, метод нечіткого структурного аналізу онтологій, матриця напівдосяжності, онтологічний граф, множини рівнів.

METHOD OF FUZZY STRUCTURAL ANALYSIS OF ONTOLOGY

O.I. Morozova, O.Yu. Sokolov, W.M. Hussein

The representation of knowledge in the form of ontology to improve the learning process is considered in this article. In the proposed approach of solving the problem, each student adds the edited ontology about his vision of the structure under consideration of information. On the basis of ontology obtained by learners' community, so-called ideal ontology of information representation is formed. To compare formed ontologies with ideal ontology the graphs of ontology are used. The method of fuzzy structural analysis of ontology is proposed to be used to construct graphs of ontologies.

Keywords: ontology, the text representation of knowledge, fuzzy sets, method of fuzzy structural analysis of ontology, semi-reachability matrix, graph of ontology, level sets.

Заключение

В работе предложено решение задачи формирования общей базы знаний обучения сообществом обучаемых. Для представления информации выбрано ее описание в виде онтологий. Для сравнения построенных различными обучаемыми онтологий используются графы онтологий. Для построения графов онтологий в работе используется метод нечеткого структурного анализа онтологий, основанный на процедуре получения нечеткой матрицы полудостижимости, из которой затем выделяют необходимые множества различных уровней. При рассмотрении метода был приведен пример для лучшего понимания основ данной теории. В дальнейшем планируется внедрение онтологической системы обучения с использованием принципов самоорганизации как одной из составных частей портала направленного обучения.

Список литературы

1. Никоненко А.А. Обзор баз знаний онтологического типа / А.А. Никоненко // Искусственный интеллект: журнал ИПШ и НАН Украины. – Наука і освіта. – Донецьк, 2009. – № 4. – С. 208-219.
2. Онтологии и тезаурусы: модели, инструменты, приложения: учебный курс Интернет-университета инф. технологий [Электронный ресурс] / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В.Д. Соловьев. – Режим доступа к ресурсу: <http://www.intuit.ru/department/expert/ontoth/2/>.
3. Загоруйко Ю.А. Информационная модель портала научных знаний / Ю.А. Загоруйко, О.И. Боровикова // Информационные технологии. – М., 2009. – № 12. – С. 2-7.
4. Хакен Г. Информация и самоорганизация: Макроскопический подход к сложным системам / Г. Хакен. – М.: Мир, 1991. – 240 с.
5. Tazaki Eiichiro. Structural modeling in a class of systems using fuzzy sets theory / Eiichiro Tazaki, Michio Amagasa // Fuzzy Sets and Systems. – 1979. – № 2. – P. 87-103.
6. Zadeh L.A. Fuzzy sets / L.A. Zadeh // Information and Control. – 1965. – № 8. – P. 338-353.
7. Кофман А. Введение в теорию нечетких множеств / А. Кофман. – М.: Радио и связь, 1982. – 432 с.

Поступила в редколлегию 21.04.2010

Рецензент: д-р техн. наук, проф. Н.Д. Кошевой, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.