

ИСПОЛЬЗОВАНИЕ АССОЦИАТИВНЫХ СЕТЕЙ ДЛЯ ГРАНУЛЯЦИИ МАТЕРИАЛОВ КУРСА

Вера Любченко, Виктор Крисилев

Abstract: An approach of constructive granulation checking for learning materials is proposed. There is defined an algorithm of associative network building and criteria of balanced concepts definition.

Keywords: concept, associative relation, associative network, granulation.

ACM Classification Keywords: I.2.4 Knowledge Representation Formalisms and Methods

Введение

Предметная область любого учебного курса имеет свои базовые понятия (концепты), свою систему иерархии и инкапсуляции этих понятий и свою систему деления на подобласти. Причем понятийные системы различных предметных областей различны, индивидуальны и непохожи друг на друга. В связи с этим стоит задача представления знаний предметной области в форме универсального поля знаний, представленного в универсальных терминах.

При разработке содержания (определении предметной области) учебного курса обязательно выясняют его внутреннюю логику, для чего

- определяют результаты, умения и уровень компетентности, ожидаемые от обучаемых по окончании курса;
- определяют концепты, формирующие содержание;
- определяют связи между концептами.

Анализ содержания учебного материала должен определить силу смысловой связи между концептами и позволить выполнить обоснованную грануляцию курса на тематические компоненты. Фактически, в ходе анализа проверяются структурированность и логическая связность учебного материала, являющиеся главными факторами, определяющими его качество.

Определение силы ассоциативных связей

Концепт – это конструкция, описывающая все объекты в данной категории или классе сущностей, их взаимодействия, свойства и связи между ними. Каждый концепт характеризуется своим контекстом, задаваемым тремя компонентами:

- множеством учебных целей, для достижения которых необходимо изучение концепта,
- предусловием – множеством категорий сущностей, знание которых необходимо для изучения данного концепта,
- постусловием – категорией сущностей, которая изучается в данном концепте.

Для определения концептной модели учебного материала следует задать множество концептов — C , множество учебных целей — G , множество предусловий — In , множество постусловий — Out . Тогда

$$C = \{c_i = \langle G_i, In_i, Out_i \rangle : G_i \in G, In_i \in In, Out_i \in Out\}, i = 1, \dots, n.$$

В [Любченко, 2006] показано, что формализация понятия ассоциативных связей может обеспечивать эффективное решение широкого круга задач интеллектуальной обработки данных. Определим ассоциативную связь как вид связи между двумя концептами, устанавливаемой исходя из заданного сочетания элементов их контекстов. На заданном множестве концептов определим два базовых отношения ассоциации.

Будем говорить, что между двумя концептами существует ассоциация по цели, если контексты концептов имеют общие элементы во множествах учебных целей:

$$R_G(c_i, c_j), \text{ если } G_i \cap G_j \neq \emptyset,$$

где $c_i, c_j \in C$ – концепты, $G_i, G_j \in G$ – множества целей этих концептов.

Будем говорить, что между двумя концептами существует ассоциация по логике, если постусловие одного из них входит в предусловие второго:

$$R_L(c_i, c_j), \text{ если } Out_i \subseteq In_j,$$

где $c_i, c_j \in C$ – концепты, $Out_i \in Out$ – постусловие концепта c_i , $In_j \in In$ – предусловие концепта c_j .

Отметим, что ассоциация по цели обладает свойством симметричности, а ассоциация по логике – нет.

Для расширения возможностей использования ассоциаций при анализе учебного материала целесообразно не просто зафиксировать факт существования ассоциации, а обеспечить возможность количественной оценки степени ассоциативной связи. Формализуем концепцию ассоциации с помощью меры ассоциативной связи – вещественной функции $ass : F^2 \rightarrow [0,1]$. Будем рассматривать два типа мер ассоциативной связи:

- 1) $ass_G : F^2 \rightarrow [0,1]$ – вычисление меры ассоциативной связи по цели, которую предлагается определять как значение отношения

$$ass_G(c_i, c_j) = \frac{|G_i \cap G_j|}{\min(|G_i|, |G_j|)};$$

- 2) $ass_L : F^2 \rightarrow [0,1]$ – вычисление меры ассоциативной связи по логике, которую предлагается определять как значение отношения

$$ass_L(c_i, c_j) = \frac{|Out_i \cap In_j|}{|In_j|}.$$

Возможность идентифицировать парные ассоциации между двумя концептами позволяет ввести в рассмотрение более сложные конструкции, которые могут быть полезны для анализа содержания учебного материала. Дадим определение.

Ассоциативная сеть – набор концептов учебного материала с определенными на них ассоциативными связями, математической моделью которой является ориентированный граф.

Для построения ассоциативной сети учебного материала предлагается использовать следующий алгоритм.

1. Построить дерево «курс – цели – концепты», корневая вершина которого соответствует анализируемому учебному курсу, вершины первого уровня – целям учебного курса, вершины второго уровня (листья) – концептам.
2. Для каждого концепта c_i определить его контекст $\langle G_i, In_i, Out_i \rangle$.

3. Объединить двумя противоположно направленными дугами листовые вершины, соответствующие концептам, между которыми существует ассоциация по цели.
4. Объединить дугами листовые вершины, соответствующие концептам, между которыми существует ассоциация по логике.
5. У всех листовых вершин удалить ребра, соединяющие их с вершинами первого уровня. Удалить компонент графа, который объединяет вершины первого уровня и корневую вершину.
6. Приписать каждой дуге графа весовой коэффициент равный мере соответствующей ассоциативной связи.
7. Если результирующий граф является гиперграфом, то для пар вершин, соединенных несколькими однонаправленными дугами, выполнить замену этих дуг дугой, весовой коэффициент которой вычисляется по формуле

$$ass(c_i, c_j) = ass_G(c_i, c_j) + ass_L(c_i, c_j).$$

Полученная ассоциативная сеть может быть использована для проверки сбалансированности определения концептов и обоснованности выделения тематических компонент.

Методика анализа грануляции курса

Грануляция курса – это объединение концептов курса в тематические компоненты. Желательно, чтобы тематические компоненты были сформированы равномерно (содержали одинаковое количество концептов) и не имели пересечений.

В [Krissilov, 2005] были сформулированы три критерия формирования предметной области интеллектуальной системы. На их основе с использованием количественной меры ассоциативной связи между концептами сформулируем три условия конструктивного определения концептов:

1. Условие излишней обособленности концепта c_i

$$\forall j, k : ass(c_i, c_j) < ass_{min} \ \& \ ass(c_k, c_i) < ass_{min}.$$

Выполнение этого условия означает, что концепт c_i практически не связан с другими концептами учебного материала. Чаще всего это вызвано излишней общностью концепта (в этом случае целесообразно разбить концепт на составные части), но могут быть и другие причины.

2. Условие излишней подробности концептов

$$ass(c_i, c_j) > ass_{max}.$$

Выполнение этого условия означает, что есть смысл объединить концепты c_i и c_j .

3. Условие сбалансированного определения концептов

$$ass_{min} \leq ass(c_i, c_j) \leq ass_{max} \ \forall i, j.$$

Выполнение этого условия означает, что разбиение на концепты выполнено с одной степенью детализации.

Значения ass_{min} и ass_{max} определяются методом экспертного оценивания.

Если определение концептов выполнено сбалансировано, то можно переходить к проверке обоснованности выделения тематических компонент. Представляется целесообразным выделять эти компоненты на основе принципа компактности [Загоруйко, 1999]. Он приводит к построению разбиения, обеспечивающего простоту добавления, изменения и удаления компонент, а также возможность повторного использования компонент. Введем в рассмотрение две метрики.

Связность тематической компоненты – это сила взаимосвязей между концептами данной компоненты, которую можно рассчитать как сумму весовых коэффициентов дуг между вершинами, соответствующими этим концептам.

Сцепление тематической компоненты – это сила взаимосвязей данной компоненты с остальными тематическими компонентами, которую можно рассчитать как сумму весовых коэффициентов дуг, соединяющих подграф данной тематической компоненты с остальными вершинами ассоциативной сети.

Эффективное разбиение на компоненты достигается максимизацией связности и минимизацией сцепления. Если для всех тематических компонент значение показателя связности больше значения показателя сцепления, то можно утверждать, что выделение тематических компонент выполнено обосновано.

Введенные метрики можно использовать для сравнения нескольких вариантов разбиения одного учебного материала на тематические компоненты. В этом случае для каждого варианта разбиения рассчитываются три показателя:

1. общая связность – сумма показателей связности всех тематических компонент разбиения,
2. общее сцепление – сумма показателей сцепления всех тематических компонент разбиения;
3. коэффициент разбиения – отношение значения общего сцепления к значению общей связности.

Лучшим из предложенных вариантов разбиения является тот, которому соответствует меньший коэффициент разбиения.

Заключение

В работе описана методика анализа грануляции курса с использованием ассоциативной сети концептов учебного материала. Методика позволяет выполнить проверку сбалансированности определения концептов и обоснованности выделения тематических компонент.

Следует обратить внимание на то, что данная методика не позволяет решить вопрос построения оптимального временного упорядочения тематических компонент. Эту проблему можно решить, введя в рассмотрение ассоциации по времени, что позволит учесть для каждого концепта степень зависимости от ранее изученных концептов.

Очевидным развитием предложенной методики является анализ ассоциативных цепочек. *Ассоциативная цепочка* – упорядоченная последовательность концептов, связанных ассоциативными связями. Можно ввести в рассмотрение косвенные ассоциации – связи между концептами через ассоциативную цепочку – и определить меру для их измерения. Тогда ассоциативная функция любой цепочки, выделенной в ассоциативной сети учебного материала, должна быть невозрастающей функцией.

Список литературы

- [Загоруйко, 1999] Загоруйко. Н.Г. Прикладные методы анализа данных и знаний. – Новосибирск: Издательство Института математики, 1999.
- [Любченко, 2006] Любченко В.В., Крисиллов В.А. Метод поиска фрейма по шаблону на основе ассоциаций // Труды Одесского политехнического университета. – 2006. – Спецвыпуск. – С. 60 – 63.
- [Krissilov, 2005] Krissilov V., Shabadash D. Towards the Problems of an Evaluation of Data Uncertainty in Decision Support Systems // Information Theories & Applications. – Vol.3. – P. 376 – 380.

Сведения об авторах

Вера Любченко – Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса 65044, Украина; e-mail: yira.lyubchenko@gmail.com

Виктор Крислов – Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса 65044, Украина; e-mail: victork@405.com.ua