

УДК 004:383.4:371.69

О.В. Соколова

ПРЕЦЕДЕНТНЫЙ ПОДХОД К ФОРМИРОВАНИЮ ЭКСПЕРТНЫХ ГРУПП

Запропоновано підхід до формування експертної групи за допомогою бази даних прецедентів. Розглянуто функціонування прецедентних систем підтримки прийняття рішень, які подано у вигляді СВР-цикла.

Введение. В настоящее время значительное распространение получают коллективные методы подготовки и принятия решений, которые применяются в самых разнообразных сферах. Одной из таких сфер применения является сфера создания качественного контента для электронных средств обучения [1,2].

Постановка задачи. Таким образом, возникает задача формирования экспертной группы для создания контента по конкретной дисциплине.

Основная часть. Известные подходы позволяют формировать экспертные группы (ЭГ) различных уровней [2]. Наиболее простой является однородная структура ЭГ, при которой все эксперты имеют одинаковые права и обязанности. Для решения задачи получения качественного контента электронного учебника с помощью специально сформированного сайта ЭГ целесообразно разделить на три относительно независимо функционирующие подгруппы, каждая из которых имеет однородную структуру. Данные подгруппы могут функционировать параллельно, обеспечивая анализ объекта экспертизы с разных, заранее определенных позиций, или же последовательно, когда результаты анализа, выполненного первой подгруппой экспертов, служат ориентирующей информацией для второй и третьей подгруппы. Алгоритм формирования ЭГ можно представить состоящим из следующих шагов:

1. Идентификация проекта (анализ структурной проблематики исследования), определение ключевых показателей проекта.
2. Определение ограничений на количество экспертов в ЭГ.
3. Обращение к БД хранилища прецедентов (ХП).
4. Поиск оптимального решения в БД прецедентов (ХП): описание текущей ситуации предметной области и выявление полученного описания понятий из имеющихся категорий; построение описания на языке представления ситуаций; поиск эталонной ситуации, в пределах заданного интервала близости; сужение поиска путем включения в поиск учета атрибутов объектов понятий; оптимизация процесса нахождения оптимальной выборки кандидатов.
5. Выбор кандидатов ЭГ.
6. Корректировка списка экспертов с применением модифицированного документационного метода.
7. Формирование ЛПП новой ЭГ и запись в БД прецедентов, если не найдены совпадения.

Процесс функционирования прецедентных СППР обычно представляется в виде так называемого СВР-цикла, состоящего из четырех основных фаз: выбор из ХП наиболее уместного прецедента или множества прецедентов, обычно на основе заданного отношения подобия; использование выбранных прецедентов для принятия решения; пересмотр и коррекция (в случае необходимости) принимавшихся ранее в выбранных прецедентах решений; сохранение в хранилище ХП принятого решения и сложившейся ситуации в качестве нового прецедента или соответствующее изменение выбранного прецедента, что может быть полезным в дальнейшем при решении аналогичных задач. Построение СППР прецедентного типа предполагает решение следующего круга задач: разработка способа представления знаний о ситуации и возможных решениях; разработка метода выбора прецедентов; разработка метода верификации и адаптации решений; разработка метода хранения и индексации прецедентов; разработка метода машинного обучения.

Обачно, СВР-цикл выполняется при непосредственном взаимодействии с ЛПП. Многие прецедентные системы только извлекают из ХП наиболее уместные прецеденты и оставляют процесс адаптации на усмотрение ЛПП, поскольку автоматизация процесса адаптации является сложной задачей и практически не поддается обобщению, а зачастую в адаптации нет необходимости. Если же адаптация необходима, то для ее проведения можно использовать БЗ предметной области, содержащую набор необходимых фактов и правил, а так же методы индуктивного и (или) дедуктивного машинного

обучения. Будем считать, что прецедент e состоит из кортежа $\langle s, r, h, z \rangle$, где ситуация $s \in S$ и связанное с ней решение $r \in R$ выполняется группой экспертов $h \in H$ под руководством ЛПП $z \in Z$.

Ситуация $s \in S$ может задаваться с помощью формул специализированного языка L . Каждой поставленной ситуации S может соответствовать несколько решений, таким образом можно допустить, что существуют прецеденты вида $\langle s, r \rangle$ и $\langle s, r' \rangle$, которые отличаются при условии что $r \neq r'$.

Данные, поступающие в модуль информационно-аналитической модели по анализу формирования ЭГ, представим множеством прецедентов M :

$$M = \{ \langle s_1, r_1, h_1, z_1 \rangle, \langle s_2, r_2, h_2, z_2 \rangle, \dots, \langle s_i, r_i, h_i, z_i \rangle, \dots, \langle s_n, r_n, h_n, z_n \rangle \}.$$

Каждый прецедент $e_i \in \{e\}$ может рассматриваться как условная импликация:

$$s_i \Rightarrow r_i, i = \overline{1, n}.$$

Таким образом, если задана некоторая ситуация:

$$s_i \approx s_j, j = \overline{1, n}, i \neq j,$$

и существует прецедент $e_i = \langle s_i, r_i, h_i, z_i \rangle$, то можно утверждать, что r_j является приближенным решением для ситуации s_i . Чем ближе ситуация s_i к ситуации s_j , тем правдоподобнее является решение r_j , которое так же является решением для s_i [17]. Для нахождения степени близости ситуации s_i к ситуации s_j и соответственно оценки близости решения r_j к искомому используется функция подобия ζ , на основе которой строится отношение подобия между прецедентами и выводится мера подобия SM . В работе принято, что система прецедентов представляет собой структуру

$$\langle DM, SM_{G(L)}, U_L \rangle,$$

где DM – база данных прецедентов, $SM_{G(L)}$ – мера подобия, заданная на множестве интерпретаций G языка L , описывающего входные ситуации; U_L – множество формул языка L .

Список условий P запишем в виде

$$P = \{ (p_1, w_1), (p_2, w_2), \dots, (p_k, w_k) \},$$

где w_1, \dots, w_k – весовые факторы условий, а p_1, \dots, p_k – предикаты, характеризующие условия.

При формировании и занесении нового прецедента в БД прецедентов уместно каждому признаку присвоить вес, учитывающий его относительную ценность.

Степень близости прецедента по всем признакам вычисляется, используя обобщенную формулу вида:

$$\sum_j w_j \cdot sim(x_{ij}, x_{kj}),$$

где w_j – вес j -го признака; sim – функция подобия; x_{ij}, x_{kj} – значения признака x_j для текущего случая и прецедента, соответственно.

Рассчитав по формуле значение для каждой модели, получим следующее множество:

$$SM = (S_1, S_2, \dots, S_n).$$

После вычисления степеней близости все прецеденты выставляются в единый ранжированный список.

Данный метод прост в реализации, но для его выполнения необходимо выделение значительного ресурса памяти, так как в процессе нахождения значения зависимой переменной для новой записи используется вся существующая БД.

Выбор меры близости считается ключевым моментом, от которого зависит поиск подходящих прецедентов. В каждой конкретной задаче выбор производится на основании своего набора правил с учетом главных целей исследования, физической и статистической природы используемой информации и т. п. Близость объектов определяется через отношения подобия между ними. Рассмотрим подход с применением Евклидова расстояния и меры сходства Хэмминга.

Пусть имеются образцы X_i и X_k в N -мерном пространстве признаков. Для ситуации, когда при оценивании претендентов известны точные количественные значения их характеристик $V_i(x)$ и их функции полезности $m_i[V_i(x)]$, математическая модель задачи формирования многофакторной оценки, альтернативы $x \in X$ может иметь вид:

$$\Phi(x) = \sum_{i=1}^n a_i m_i[V_i(x)], \sum_{i=1}^n a_i = 1,$$

при принципе оптимальности

$$x^\circ = \arg \max_{x \in X} \sum_{i=1}^n a_i m_i[V_i(x)] \text{ или } x^\circ = \arg \max_{x \in X} \sum_{i=1}^n a_i \overline{m}_i[V_i(x)],$$

где $\overline{m}_i[V_i(x)] = 1 - m_i[V_i(x)]$ – функция потери полезности.

В ситуации, когда точные количественные значения a_i не известны, но известна качественная информация по взаимной важности оценивания критериев:

$$V_1(x) \succ, \dots, V_n(x).$$

Для выбора лучшего кандидата из получаемого множества кандидатов X выделяется подмножество x_1° :

$$x^\circ = \arg \max_{x \in X} m_i[V_i(x)].$$

Так как x_1° состоит из ЭГ (множество экспертов), решаем задачу выбора кандидатов из множества x_1° по критериям важности. При отсутствии количественной информации о коэффициентах a_i , принимаются условия равенности важности критериев $a_i = 1/n$, $i = \overline{1, n}$, а модель оценивания обобщенной полезности альтернативы $x \in X$ представляется в виде:

$$\Phi(x) = \frac{1}{n} \left\{ \sum_{i=1}^n m_i[V_i(x)] \right\}.$$

Принцип оптимальности тогда принимает вид:

$$x^\circ = \arg \max_{x \in X} \frac{1}{n} \left\{ \sum_{i=1}^n m_i[V_i(x)] \right\}.$$

В ситуации, когда для части критериев $V_i(x)$, $i = \overline{1, n}$ весовые коэффициенты ВК известны, а для других оценка превосходства отсутствует, предусматривается рассматривать два множества критериев: множество критериев Y с известными весовыми коэффициентами ВК a_i и множество S , для которого весовые коэффициенты a_i неизвестны:

$$x^\circ = \arg \max_{x \in X} \left\{ \sum_{\substack{i=1 \\ V_i(x) \in Y}}^y a_i m_i[V_i(x)] + \frac{1}{s} \left[1 - \sum_{i=1}^y a_i \right] \sum_{\substack{i=1 \\ V_i(x) \in S}}^s m_i[V_i(x)] \right\}.$$

Подмножество наиболее близких прецедентов используется для получения упорядоченных оценок подобия эталона эксперта в кандидаты для ЭГ. На основе полученных данных принимается решение о включении эксперта в группу по решению данной проблемы.

Выводы. Описанный подход предлагается использовать для экспертизы по решению задачи формирования информационного обеспечения для компьютеризированного обучения по конкретной дисциплине.

После решения задачи формирования экспертных групп и оценки компетентности экспертов проводится экспертиза по методу Дельфы, что обеспечивает построение обоснованной начальной траектории обучения.

ЛИТЕРАТУРА:

1. Карпенко Д.С. Система автоматического повышения качества тестовых заданий и мониторинг процесса усвоения знаний / Д.С. Карпенко, О.М. Карпенко, Е.Н. Шлихунова. – К.: Наукова думка, 2008. – 165 с.
2. Гнатієнко Г.М. Експертні технології прийняття рішень. Монографія / Г.М. Гнатієнко, В.Є. Снітюк – К.: ТОВ «Маклаут», 2008. – 444 с.
3. Кухаренко В.М. Дистанційне навчання: Умови застосування. Дистанційний курс: Навчальний посібник. 3-є вид. / Кухаренко В.М., Рибалко О.В., Сиротенко Н.Г. – Харків: НТУ „ХПІ”, „Торсинг”, 2002. – 320 с.

4. Моисеева М.В. Интернет-обучение: технологии педагогического дизайна / М.В. Моисеева, В.Е. Полат – М.: Издательский дом «Камерон», 2004. – 216 с.

СОКОЛОВА Оксана Валентиновна – аспирант кафедры экономической кибернетики Херсонского национального технического университета.

Научные интересы:

– методы и модели компьютеризированного обучения.