

УДК: 004.853

Извлечение нечетких знаний при разработке экспертных прогнозных диагностических систем

В. А. Суздальцев, И. В. Суздальцев, Э. Г. Тахавова^а

Казанский национальный исследовательский технический университет им. А. Н. Туполева – КАИ,
Россия, 420111, г. Казань, ул. К. Маркса, д. 10

E-mail: ^а elzzy@yandex.ru

Получено 24.11.2021, после доработки — 02.11.2022.

Принято к публикации 07.11.2022.

Экспертные системы имитируют профессиональный опыт и мыслительный процесс специалиста при решении задач в различных предметных областях, в том числе в прогнозной диагностике в медицине и технике. При решении подобных задач применяются нечеткие модели принятия решений, что позволяет использовать профессиональные экспертные знания при формировании прогноза, исключая анализ данных непосредственных экспериментов. При построении нечетких моделей принятия решений используются типовые нечеткие ситуации, анализ которых позволяет сделать вывод специалистам о возникновении в будущем времени нестандартных ситуаций. При разработке базы знаний экспертной системы прибегают к опросу экспертов: инженеры по знаниям используют мнение экспертов для оценки соответствия между типовой текущей ситуацией и риском возникновения чрезвычайной ситуации в будущем. В большинстве работ рассматриваются методы извлечения знаний с точки зрения психологических, лингвистических аспектов. Множественные исследования посвящены проблемам контактного, процедурного или когнитивного слоев процесса извлечения знаний. Однако в процессе извлечения знаний следует отметить значительную трудоемкость процесса взаимодействия инженеров по знаниям с экспертами при определении типовых нечетких ситуаций и оценок рисков нестандартных ситуаций. Причиной трудоемкости является то, что число вопросов, на которые должен ответить эксперт, очень велико. В статье обосновывается метод, который позволяет инженеру по знаниям сократить количество вопросов, задаваемых эксперту, а следовательно, снизить трудоемкость разработки базы знаний. Метод предполагает наличие отношения предпочтения, определяемое на множестве нечетких ситуаций, что позволяет частично автоматизировать формирование оценок частоты наступления нечетких ситуаций и тем самым сократить трудоемкость созданий базы знаний. Для подтверждения проверки и целесообразности предложенного метода проведены модельные эксперименты, результаты которых приведены в статье. На основе предложенного метода разработаны и внедрены в эксплуатацию несколько экспертных систем для прогнозирования групп риска патологий беременных и новорожденных.

Ключевые слова: экспертная система, извлечение знаний, лингвистическая переменная, степень принадлежности, нечеткое правило

UDC: 004.853

Fuzzy knowledge extraction in the development of expert predictive diagnostic systems

V. A. Suzdaltsev, I. V. Suzdaltsev, E. G. Takhavova^a

Kazan National Research Technical University named after A. N. Tupolev – KAI,
10 K. Marx st., Kazan, 420111, Russia

E-mail: ^a elzzy@yandex.ru

*Received 24.11.2021, after completion – 02.11.2022.
Accepted for publication 07.11.2022.*

Expert systems imitate professional experience and thinking process of a specialist to solve problems in various subject areas. An example of the problem that it is expedient to solve with the help of the expert system is the problem of forming a diagnosis that arises in technology, medicine, and other fields. When solving the diagnostic problem, it is necessary to anticipate the occurrence of critical or emergency situations in the future. They are situations, which require timely intervention of specialists to prevent critical aftermath. Fuzzy sets theory provides one of the approaches to solve ill-structured problems, diagnosis-making problems belong to which. The theory of fuzzy sets provides means for the formation of linguistic variables, which are helpful to describe the modeled process. Linguistic variables are elements of fuzzy logical rules that simulate the reasoning of professionals in the subject area. To develop fuzzy rules it is necessary to resort to a survey of experts. Knowledge engineers use experts' opinion to evaluate correspondence between a typical current situation and the risk of emergency in the future. The result of knowledge extraction is a description of linguistic variables that includes a combination of signs. Experts are involved in the survey to create descriptions of linguistic variables and present a set of simulated situations. When building such systems, the main problem of the survey is laboriousness of the process of interaction of knowledge engineers with experts. The main reason is the multiplicity of questions the expert must answer. The paper represents reasoning of the method, which allows knowledge engineer to reduce the number of questions posed to the expert. The paper describes the experiments carried out to test the applicability of the proposed method. An expert system for predicting risk groups for neonatal pathologies and pregnancy pathologies using the proposed knowledge extraction method confirms the feasibility of the proposed approach.

Keywords: expert system, knowledge acquisition, linguistic variable, membership degree, fuzzy rule

Citation: *Computer Research and Modeling*, 2022, vol. 14, no. 6, pp. 1395–1408 (Russian).

1. Введение

В настоящее время широкое применение экспертные системы находят в области прогнозной диагностики функционирования различных сложных систем. Подобные задачи характерны для техники, медицины [Shortliffe, 1976] и других областей профессиональной деятельности. При решении задач прогнозной диагностики в таких системах устанавливается возможность наличия нештатных ситуаций (НС) в будущем времени на основе наблюдения ситуаций, возникающих в настоящем времени при функционировании сложной системы. Такое предвидение, как правило, осуществляется экспертами-профессионалами, имеющими высокую квалификацию и большой опыт. Экспертные системы (ЭС), имитирующие опыт эксперта, позволяют заменить специалиста и осуществлять прогнозирование НС с необходимой точностью [Козлова, Игнатъев, Самойлова, 2011; Рутковская, Пилински, Рутковски, 2013].

При построении подобных экспертных систем необходимо в первую очередь определить, извлечь, приобрести знания экспертов-профессионалов и сформировать базу знаний ЭС. Процесс приобретения знаний — важнейший этап разработки любой экспертной системы: «Приобретение знаний является самым большим препятствием при разработке экспертных систем» [Olson, Henry, 2016, с. 152]. В этой области существует достаточно много работ как отечественных, так и зарубежных авторов. Многие авторы исследуют проблему формализации знаний [Леденева, Сергиенко, 2008] и разработки моделей представления знаний [Гаврилова, Червинская, 1992]. Большое число работ также представляют исследования по структурированию знаний о предметной области [Гаврилова, Червинская, 1992; McDermott, 1980], разработке программных средств обеспечения процесса извлечения знаний [Kitto, Boose, 1989]. Кроме проблемы структурирования и моделирования знаний большое внимание уделяется анализу различных подходов к организации процесса взаимодействия инженера по знаниям и эксперта [Kitto, Boose, 1987; Wagner, Chung, Naidawi, 2003]. Много работ в этой области представляют исследования проблем контактного, процедурного, когнитивного слоев процесса извлечения знаний [Гаврилова, Червинская, 1992]. Большое внимание уделялось также методам интервьюирования экспертов при извлечении знаний [Wagner, Chung, Naidawi, 2003], отличающимся количеством экспертов или стратегиями опроса [Gammack, Young, 1985; Wagner, 2006]. Рассматриваются также различные аспекты прямых и косвенных методов опроса [Olson, Henry, 2016].

Процедуры приобретения знаний можно построить на основе обработки фактических данных, например, при использовании байесовского подхода [Шевелев, 2017] и других статистических методов [Городецкий, Курбанов, Тарасова, 2012; Rizaev, Takhavova, 2016] прогнозирования. Однако область применения этих методов существенно сужается из-за требований к статистической независимости признаков (параметров), характеризующих наступление НС, что часто не выполняется на практике в связи с наличием общих латентных факторных признаков ситуации, вызывающих изменение корреляционно связанных наблюдаемых признаков. Также следует отметить, что использование фактических данных с результатами исходов для формулирования правил прогнозирования часто затруднено, особенно в медицине, поскольку возникновение НС бывает просто недопустимым по необходимости сохранения здоровья пациентов. В таких случаях единственным способом сформировать правила прогнозирования является интервьюирование экспертов [Davis, 1977]. Целью интервьюирования экспертов является выявление и описание соответствия между типовыми ситуациями и частотой возникновения нештатных ситуаций в будущем при наличии этих ситуаций в настоящее время. Неопределенность, источником которой являются недостаток информации и невозможность точного описания поведения объектов диагностирования, является одной из основных проблем в задачах прогнозной диагностики. Моделирование рассуждений эксперта должно основываться на обработке качественной информации, выраженной профессиональным языком, в качестве которого обычно фигурирует естественный язык. Теория нечетких множеств стала теоретической основой для создания формализованного

описания таких предметных областей [Zadeh, 1983]. Нечеткие множества дают возможность разрабатывать формализованное описание объектов с использованием лингвистических переменных [Заде, 1976]. Лингвистические переменные позволяют описать состояние контролируемого объекта с использованием качественных значений признаков НС и реализовать идеи нечеткого логического вывода для решения задач принятия решений, когда и отсутствуют фактические данные об аномальных ситуациях [Аверкин и др., 1986; Пегат, 2013; Suzdaltsev, Suzdaltsev, Vogula, 2017].

Следует указать на еще одну проблему процесса приобретения знаний: процесс извлечения знаний требует много времени, когда речь идет о приобретении знаний при построении нечетких моделей прогнозирования для реальных сложных систем. Число ситуаций, по которым эксперт должен высказать свое мнение, очень велико, поэтому процесс извлечения знаний характеризуется высокой трудоемкостью. Возникает необходимость в разработке таких методов приобретения знаний, которые позволяют повысить эффективность процесса извлечения знаний за счет снижения трудоемкости интервьюирования эксперта.

2. Постановка задачи

Построим нечеткую модель процесса прогнозирования, опираясь на аппарат теории нечетких множеств. Образует лингвистическую переменную $\langle \beta_0, T_0, Y \rangle$. Элемент β_0 представляет имя лингвистической переменной «Частота наступления нештатной ситуации» с терм-множеством значений лингвистической переменной T_0 и базовым множеством Y , определяющим количественную меру частоты наступления нештатной ситуации: $Y = [0, 1]$.

Терм-множество T_0 значений лингвистической переменной является двухэлементным множеством:

$$T_0 = \{\alpha_{01}, \alpha_{02}\},$$

где α_{01} и α_{02} — имена нечетких переменных, определяющих соответственно значения «Низкая частота возникновения нештатной ситуации» и «Высокая частота возникновения нештатной ситуации».

Нечеткие переменные с именами α_{10} и α_{20} являются упорядоченными тройками:

$$\langle \alpha_{01}, Y, \tilde{M}_{01} \rangle, \quad \langle \alpha_{02}, Y, \tilde{M}_{02} \rangle,$$

где \tilde{M}_{01} и \tilde{M}_{02} — нечеткие множества с функциями принадлежности $\mu_{01}(y)$ и $\mu_{02}(y)$:

$$\tilde{M}_{01} = \left\{ \left\langle \mu_{M_{01}}(y), y \right\rangle \mid y \in Y \right\}, \quad \tilde{M}_{02} = \left\{ \left\langle \mu_{M_{02}}(y), y \right\rangle \mid y \in Y \right\}.$$

Возникновение любой нештатной ситуации может быть охарактеризовано набором из m признаков с именами β_i и соответствующими им лингвистическими переменными с терм-множествами значений T_i и базовыми множествами X_i , определяющими количественные значения i -го признака:

$$\langle \beta_i, T_i, X_i \rangle, \quad i = \overline{1, m}.$$

Элементы терм-множеств T_i содержат нечеткие переменные соответственно нечетких значений признаков нештатной ситуации:

$$T_i = \{\alpha_{ij} \mid j = \overline{1, n_i}\}, \quad i = \overline{1, m}.$$

Будем считать, что имена нечетких переменных выбраны таким образом, что терм-множества попарно не пересекаются:

$$T_{i_1} \cap T_{i_2} = \emptyset; \quad i_1, i_2 = \overline{1, m}; \quad i_1 \neq i_2.$$

Нечеткие переменные с именами α_{ij} являются нечетким тройками:

$$\langle \alpha_{ij}, X_i, \widetilde{M}_{ij} \rangle,$$

где \widetilde{M}_{ij} — нечеткие множества с функциями принадлежности $\mu_{ij}(x_i)$:

$$\widetilde{M}_{ij} = \left\{ \langle \mu_{M_{ij}}(x_i), x_i \rangle \mid x_i \in X_i, i = \overline{1, m} \right\}, \quad j = \overline{1, n}.$$

Пространство состояний объекта прогнозирования определяется декартовым произведением X :

$$X = X_1 \times X_2 \times \dots \times X_m.$$

Нечеткую ситуацию определим нечетким множеством [Мелихов, Берштейн, Коровин, 1990]

$$\widetilde{s} = \left\{ \langle \mu_s(\alpha), \alpha \rangle \mid \alpha \in T, T = \bigcup_{i=1}^m T_i \right\}.$$

В множестве нечетких ситуаций выделим подмножество типовых нечетких ситуаций:

$$S^* = \{ \widetilde{s} \mid (\forall t \in T (\mu_s(t) = 0 \vee \mu_s(t) = 1)) \& (\forall i = \overline{1, m} \exists! t \in T_i (\mu_s(t) = 1)) \}.$$

Количество типовых нечетких ситуаций определяется выражением

$$L = |S^*| = \prod_{i=1}^m n_i.$$

Для множества типовых нечетких ситуаций S^* построим множество сюръективных отображений $(f_i: S^* \rightarrow T, i = \overline{1, m})$ таких, что $\mu_s(f_i(s)) = 1, i = \overline{1, m}$.

Степень принадлежности состояния x ($x \in X$) к типовой нечеткой ситуации \widetilde{s}^* ($\widetilde{s}^* \in S^*$) определим выражением

$$w(x, \widetilde{s}^*) = \prod_{i=1}^m \left(\sum_{j=1}^{n_i} \mu_{M_{ij}}(x_i) \mu_{s^*}(\alpha_{ij}) \right).$$

Если с помощью эксперта установлена функция φ ($\varphi: S^* \rightarrow Y$), задающая оценку частоты наступления НС y ($y \in Y$) в типовой нечеткой ситуации \widetilde{s}^* ($\widetilde{s}^* \in S^*$), то оценить степень высокой частоты наступления нештатной ситуации в состоянии x ($x \in X$) можно следующим образом [Мелихов, Берштейн, Коровин, 1990]:

$$v(x) = \max_{\widetilde{s}^* \in S^*} (w(x, \widetilde{s}^*) \mu_{02}(\varphi(\widetilde{s}^*))).$$

При использовании данного подхода прогнозирования нештатной ситуации при построении ЭС необходимо определить:

- 1) лингвистические переменные с функциями принадлежности $\mu_{01}(y)$ и $\mu_{02}(y)$;
- 2) набор признаков и соответствующие им лингвистические переменные с функциями принадлежности $\mu_{ij}(x_i)$;
- 3) функцию φ , определяющую оценку частоты возникновения НС y ($y \in Y$) при наступлении типовой нечеткой ситуации ($\widetilde{s}^* \in S^*$).

Решение задач 1 и 2 достаточно проработано и представлено в различных источниках [Мелихов, Берштейн, Коровин, 1990; Пегат, 2013; Рутковская, Пилински, Рутковски, 2013].

Остановимся на решении последней задачи. Прежде всего заметим, что в реальности на практике величина L , представляющая количество типовых ситуаций, имеет достаточно большое значение — это является одной из причин высокой трудоемкости разработки базы знаний экспертной системы. Расчеты показывают, что даже при небольшом количестве лингвистических переменных (признаков) и их нечетких значений (нечетких значений признаков) количество вопросов, которые приходится задавать эксперту, достигает достаточно больших величин. Например, при количестве признаков, равном 7, и среднем количестве нечетких значений признаков, равном 3, количество вопросов, задаваемых эксперту, равно 2187. В этом случае при среднем времени ответа на один вопрос, равном трем минутам, общее время опроса превышает 100 часов. Поэтому возникает задача определения такой организации интервьюирования экспертов, которая обеспечивает снижение трудоемкости технологий подготовки базы знаний ЭС.

3. Предлагаемый метод снижения трудоемкости опроса

Задача уменьшения количества вопросов, которые необходимо задать эксперту при формировании правил прогнозирования, может быть решена с использованием дополнительной информации об отображении φ , т. е. для совокупности правил прогноза должны выполняться некоторые условия, которые устанавливаются экспертом до интервьюирования.

Часто при прогнозной диагностике предполагается, что на элементах каждого множества T_i имеет место отношение нестрогого порядка R_i для признаков, представленных в модели лингвистическими переменными с именами β_i ($i = \overline{1, m}$). Отношение R_i означает, что пара $\langle \alpha_{ij_1}, \alpha_{ij_2} \rangle$ принадлежит R_i , если значение α_{ij_1} является более характерным для возникновения нештатной ситуации, чем значение α_{ij_2} ($\alpha_{ij_1}, \alpha_{ij_2} \in T_i, j_1, j_2 = \overline{1, n_i}, i = \overline{1, m}$):

$$\forall i = \overline{1, m} \forall s, t \in S^* (\langle \alpha_i^s, \alpha_i^t \rangle \in R_i \rightarrow (\varphi(s) \geq \varphi(t)); \alpha_i^s = f_i(s), \alpha_i^t = f_i(t))$$

и может быть построено полное антисимметричное, рефлексивное и транзитивное отношение нестрогого предпочтения R на множестве нечетких ситуаций S^* :

$$R = \{ \langle s, t \rangle \mid \langle s, t \rangle \in (S^* \times S^*) \rightarrow (\forall i = \overline{1, m} (\langle \alpha_i^t, \alpha_i^s \rangle \notin R_i) \& \exists k = \overline{1, m} (\langle \alpha_k^s, \alpha_k^t \rangle \in R_k); \alpha_i^s = f_i(s), \alpha_i^t = f_i(t)) \}.$$

Ситуация s является более характерной для возникновения нештатной ситуации, чем ситуация t , если значение каждого признака α_i^t в случае ситуации t не является более характерным для возникновения нештатной ситуации, чем значение того же признака α_i^s у ситуации s , и существует хотя бы один признак k ($k = \overline{1, m}$), значение которого α_k^s в ситуации s является более характерным, чем значение того же признака α_k^t в ситуации t , $s, t \in S^*$.

Тогда справедливо следующее утверждение:

$$\forall s, t \in S^* (\langle s, t \rangle \in R \rightarrow (\varphi(t) \leq \varphi(s))).$$

Это означает, что если ситуация s имеет предпочтение над ситуацией t , то частота возникновения нештатной ситуации y_s ($y_s = \varphi(t)$) в случае возникновения ситуации s не будет меньше, чем частота возникновения НС для ситуации y_t ($y_t = \varphi(t)$) в случае возникновения ситуации t .

Учитывая свойства отношений R и R_i ($i = \overline{1, m}$), можно показать, что найдутся типовые ситуации \bar{s} и \underline{s} ($\bar{s}, \underline{s} \in S^*$), для которых будут выполняться следующие условия:

$$\forall t \in S^* (\varphi(t) \leq \varphi(\bar{s})); \quad \forall i = \overline{1, m} \forall t \in S^* (\langle \alpha_i^{\bar{s}}, \alpha_i^t \rangle \in R_i), \quad \alpha_i^{\bar{s}} = f_i(\bar{s}), \quad \alpha_i^t = f_i(t);$$

$$\forall t \in S^* (\varphi(\underline{s}) \leq \varphi(t)); \quad \forall i = \overline{1, m} \forall t \in S^* (\langle \alpha_i^t, \alpha_i^{\underline{s}} \rangle \in R_i), \quad \alpha_i^{\underline{s}} = f_i(\underline{s}), \quad \alpha_i^t = f_i(t).$$

Пусть y_s^- и y_s^+ обозначают границы значения частоты возникновения НС, когда имеем в наличии типовую ситуацию s ($s \in S^*$):

$$y_s^- \leq y_s \leq y_s^+.$$

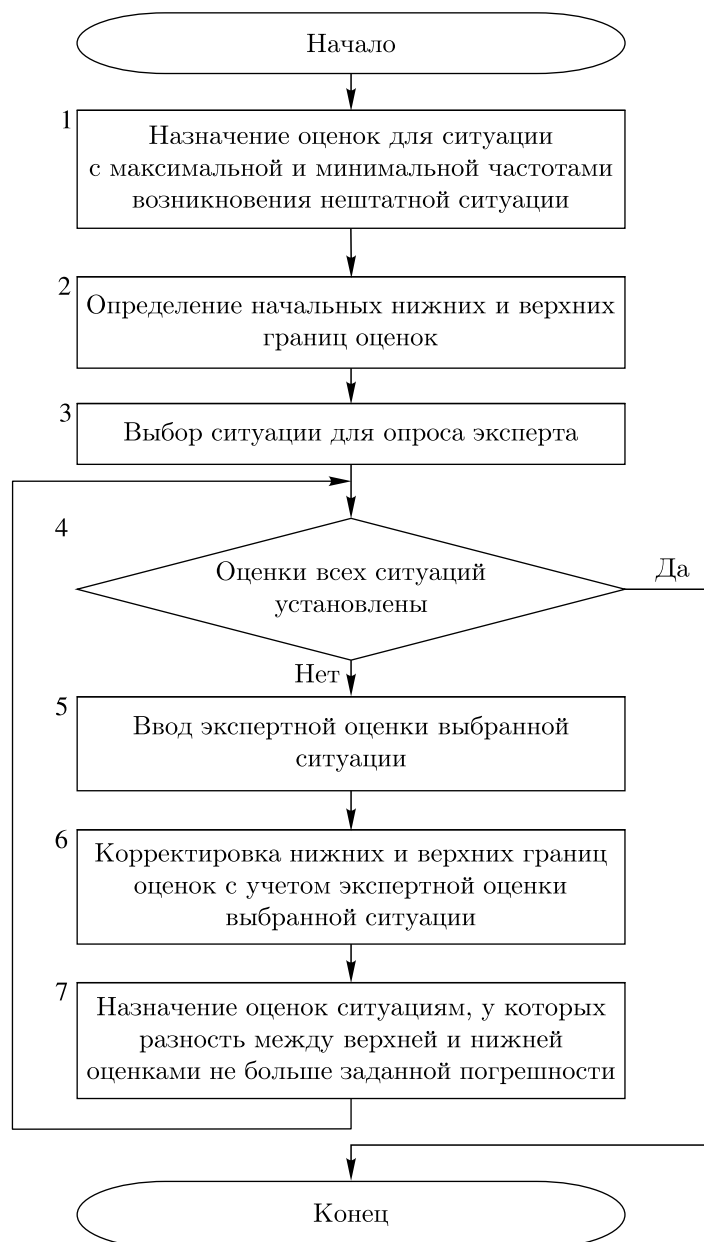


Рис. 1. Схема алгоритма процедуры интервьюирования эксперта

Схема алгоритма извлечения знаний при интервьюировании эксперта показана на рис. 1. В начале опроса эксперта для типовых ситуаций \bar{s} и \underline{s} значения $y_{\bar{s}}$ и $y_{\underline{s}}$ устанавливаются равными:

$$\begin{aligned} y_{\bar{s}} &= y_{\bar{s}}^+ = y_{\bar{s}}^- = 1, \\ y_{\underline{s}} &= y_{\underline{s}}^+ = y_{\underline{s}}^- = 0, \end{aligned}$$

а для каждой ситуации s ($s \in S^*$, $s \neq \bar{s}$, $s \neq \underline{s}$) возможно любое значение из вариации степени риска и $y_s^- = 0$ и $y_s^+ = 1$. В процессе опроса эксперта границы значений y_s^- и y_s^+ будут изменены.

Пусть эксперт оценил частоту возникновения нештатной ситуации y_s при наличии текущей типовой ситуации s ($s \in S^*$). В этом случае, одной стороны, нижняя граница для значения частоты возникновения нештатной ситуации для тех ситуаций из множества W_s ($W_s = \{w \mid w \in Coim_R s, y_w^+ \neq y_w^-\}$), которые являются предпочтительней по отношению к ситуации s , т.е. когда $\langle w, s \rangle \in R$ и $w \in W_s$, может быть увеличена на δ_{ws}^+ :

$$\delta_{ws}^+ = \begin{cases} y_s - y_w^- & \text{при } y_s > y_w^-, \\ 0 & \text{в противном случае.} \end{cases}$$

С другой стороны, для ситуации v из множества V_s ($V_s = \{v \mid v \in Im_R s, y_v^+ \neq y_v^-\}$), когда ситуация s ($\langle s, v \rangle \in R$, $v \in V_s$) является предпочтительнее, верхняя граница для оценки значения частоты возникновения нештатной ситуации может быть уменьшена на величину δ_{sv}^- :

$$\delta_{sv}^- = \begin{cases} y_v^+ - y_s, & \text{если } y_v^+ > y_s, \\ 0 & \text{в противном случае.} \end{cases}$$

Значения величин $q_s^+ = \sum_{w \in W_s} \delta_{sw}^+$, $q_s^- = \sum_{v \in V_s} \delta_{sv}^-$ предлагается использовать при отборе очередной типовой ситуации для оценки ее экспертом. Возможны различные варианты критериев отбора ситуаций для оценки частоты наступления НС в соответствии с формулами (1) и (2). При этом необходимо выбирать такую ситуацию s ($s \in S^*$), для которой отсутствует оценка, и значение одного из показателей была бы максимальна:

$$s = \arg \left(\max_{v \in S} (q_v^+, q_v^-) \right), \quad (1)$$

$$s = \arg \left(\max_{v \in S} (\min (q_v^+, q_v^-)) \right). \quad (2)$$

Однако в самом начале, перед опросом эксперта, значение y_s является неизвестной величиной. В этом случае для вычислений значений δ_{ws}^- и δ_{sv}^+ используется оценка y_s^* :

$$y_s^* = \frac{y_s^+ + y_s^-}{2}.$$

Рассмотрим пример. Пусть оценивается частота осложнений пациента после перенесенного заболевания, вызванного вирусной инфекцией. Описание лингвистических переменных, соответствующих признакам осложнения с нечеткими значениями, приведено в табл. 1.

Функции принадлежности нечетких множеств, определяющие эти нечеткие значения, показаны в табл. 2 и на рис. 2.

Пусть заданы отношения нестрогого порядка R_i так, что выполняются следующие условия:

$$\langle \alpha_{ij_1}, \alpha_{ij_2} \rangle \in R_i \leftrightarrow j_1 \geq j_2; \quad i = \overline{1, 2}; \quad j_1, j_2 = \overline{1, 3}.$$

Таблица 1

| | Имя лингвистической переменной ($i = 0$) | Имя первой лингвистической переменной признака ($i = 1$) | Имя второй лингвистической переменной признака ($i = 2$) |
|-----|--|--|--|
| | Частота возникновения осложнения | Кашель | Температура |
| J | Имя нечеткой переменной значения (α_{0j}) | Имя нечеткой переменной значения (α_{1i}) | Имя нечеткой переменной значения (α_{2i}) |
| 1 | Низкая частота осложнений | Отсутствие | Нормальная |
| 2 | Высокая частота осложнений | Редкий | Небольшое повышение |
| 3 | — | Частый | Повышенная |

Таблица 2

| J | Функции принадлежности нечетких множеств нечетких переменных первой лингвистической переменной «Кашель», $i = 1$ | Функции принадлежности нечетких множеств нечетких переменных второй лингвистической переменной «Температура», $i = 2$ |
|-----|---|--|
| | $x_1 \in [0, \infty)$ | $x_2 \in [0^\circ, \infty)$ |
| 1 | $\mu_{11}(x_1) = \begin{cases} 1 - 0,1 \cdot x_1 & \text{при } x_1 \leq 10, \\ 0 & \text{при } x_1 \geq 10 \end{cases}$ | $\mu_{21}(x_2) = \begin{cases} 1 & \text{при } x_2 \leq 36,6^\circ, \\ 25,4 - \frac{2}{3} \cdot x_2 & \text{при } 36,6^\circ \leq x_2 \leq 38,1^\circ, \\ 0 & \text{при } x_2 \geq 38,1^\circ \end{cases}$ |
| 2 | $\mu_{12}(x_1) = \begin{cases} 0,25 \cdot x_1 & \text{при } x_1 \leq 4, \\ 1 & \text{при } 4 \leq x_1 \leq 6, \\ 1 - 0,25 \cdot x_1 & \text{при } 6 \leq x_1 \leq 10, \\ 0 & \text{при } x_1 \geq 10 \end{cases}$ | $\mu_{22}(x_2) = \begin{cases} 0 & \text{при } x_2 \leq 36,6^\circ, \\ \frac{5}{3} \cdot x_2 - 61 & \text{при } 36,6^\circ \leq x_2 \leq 37,2^\circ, \\ 1 & \text{при } 37,2^\circ \leq x_2 \leq 37,5^\circ, \\ -\frac{5}{3} \cdot x_2 + 63,5 & \text{при } 37,5^\circ \leq x_2 \leq 38,1^\circ, \\ 0 & \text{при } x_2 \geq 38,1^\circ \end{cases}$ |
| 3 | $\mu_{13}(x_1) = \begin{cases} 0,1 \cdot x_1 & \text{при } x_1 \leq 10, \\ 1 & \text{при } x_1 \geq 10 \end{cases}$ | $\mu_{23}(x_2) = \begin{cases} 0 & \text{при } x_2 \leq 37,5^\circ, \\ \frac{2}{3} \cdot x_2 - 24,4 & \text{при } 36,6^\circ \leq x_2 \leq 38,1^\circ, \\ 1 & \text{при } x_2 \geq 38,1^\circ \end{cases}$ |

Тогда можно построить отношение предпочтения R на множестве типовых ситуаций (рис. 3). Запись $s_k: \langle \alpha_{1j}, \alpha_{2j} \rangle$ на рис. 3 означает, что в k -й типовой нечеткой ситуации ($k = \overline{1, 9}$) первый и второй признаки принимают j -е нечеткие значения α_{1j} и α_{2j} соответственно ($j = \overline{1, 3}$); стрелки указывают на наличие отношения предпочтения между типовыми нечеткими ситуациями, при этом стрелка указывает на типовую нечеткую ситуацию с меньшим предпочтением. Транзитивные элементы отношения предпочтения на рисунке не показаны.

Один из вариантов выполнения последовательности шагов при формировании оценок частот наступления НС представлен в табл. 3. В рассмотренном примере значение погрешности определения частоты наступления НС принимается равным 0,33. Перед опросом эксперта (шаг 0) устанавливаются оценки для ситуаций с самой низкой и высокой частотой наступления НС: соответственно 0 и 1. Для всех остальных ситуаций устанавливаются нижние и верхние границы оценок: 0 и 1 соответственно. В начале выполнения шага 1, в соответствии с первым критерием, в соответствии с выражением (1), определяется ситуация с номером 5, которая оценивается экспертом значением частоты, равной 0,2. Далее при завершении шага 1 пересчитываются значения нижних и верхних оценок частоты наступления НС. Для ситуации номер 2 и ситуации номер 4,

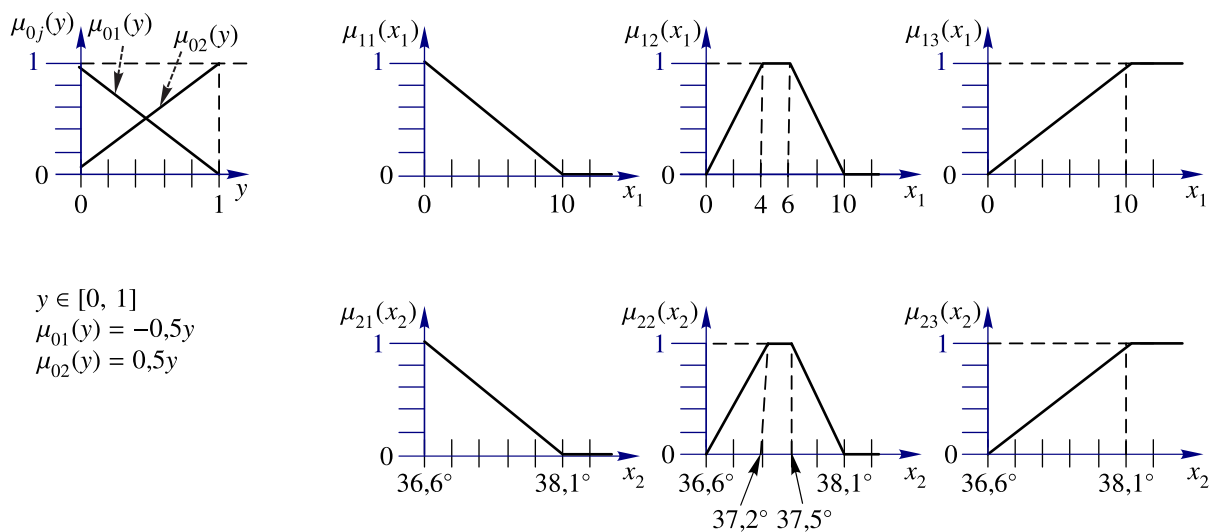


Рис. 2. Функции принадлежности нечетких множеств: μ_{01} и μ_{02} определяют нечеткие значения «Низкая частота осложнений» и «Высокая частота осложнений» соответственно лингвистической переменной «Частота наступления нештатной ситуации» (значение переменной y – количественная мера частоты наступления нештатной ситуации). μ_{11} , μ_{12} и μ_{13} определяют нечеткие значения «Отсутствует», «Редкий» и «Частый» соответственно для лингвистической переменной «Кашель» (x_1 – количественная мера частоты возникновения приступов кашля в единицу времени); μ_{21} , μ_{22} и μ_{23} определяют нечеткие значения «Нормальная», «Небольшое повышение» и «Повышенная» лингвистической переменной «Температура» соответственно (x_2 – количественная мера максимальной температуры пациента, возникшей во время заболевания)

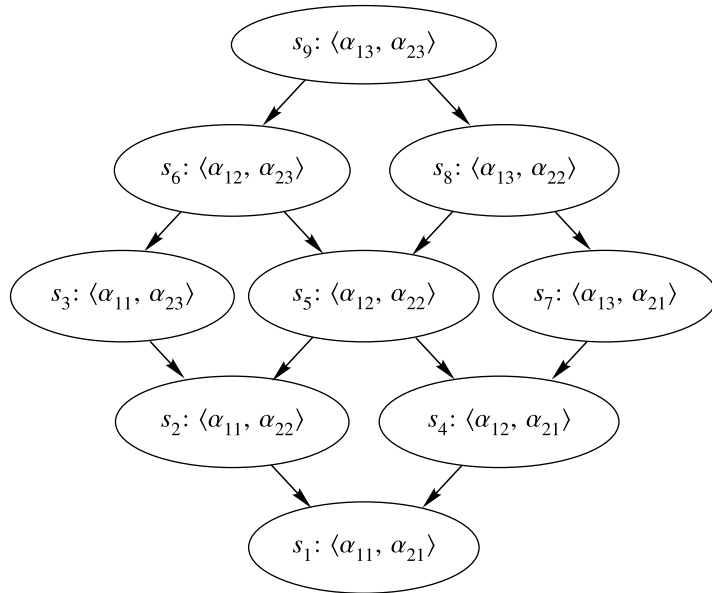


Рис. 3. Отношение предпочтения R между ситуациями. Запись $s_k: \langle \alpha_{1j_1}, \alpha_{2j_2} \rangle$ означает, что в k -й нечеткой ситуации первый и второй признаки соответственно принимают j_1 -е и j_2 -е нечеткие значения α_{1j_1} и α_{2j_2} , стрелки представляют наличие отношения предпочтения между нечеткими ситуациями (конец стрелки указывает на нечеткую ситуацию с меньшим предпочтением)

ввиду того что разность значения нижних и верхних оценок частот наступления НС оказывается меньше установленной погрешности 0,33, оценки частот наступления НС определяются авто-

Таблица 3. Последовательность формирования оценок

| | Номер k ситуации s_k | | | | | | | | |
|--|--------------------------|-------|------|-------|------|-------|------|--------|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Шаг 0 | | | | | | | | | |
| Нижние границы оценок (y_s^-) | — | 0 | 0 | 0 | 0 | 0 | 0 | 0 | — |
| Верхние границы оценок (y_s^+) | — | 1 | 1 | 1 | 1 | 1 | 1 | 1 | — |
| Оценки ситуации (y_s) | 0 | — | — | — | — | — | — | — | 1 |
| Шаг 1 | | | | | | | | | |
| Максимум: $\max_{s \in S} (q_s^+ + q_s^-)$ | — | 0 | 0,5 | 0 | 1 | 0 | 0,5 | 0 | — |
| Ввод оценки экспертом (*) | | | | | | | | | |
| Оценки ситуации (y_s) | 0 | — | — | — | 0,2* | — | — | — | 1 |
| Изменение оценок после ввода оценки экспертом (**) | | | | | | | | | |
| Нижние границы оценок (y_s^-) | — | 0 | 0 | 0 | — | 0,2 | 0 | 0,2 | — |
| Верхние границы оценок (y_s^+) | — | 0,2 | 1 | 0,2 | — | 1 | 1 | 1 | — |
| Оценки ситуации (y_s) | 0 | 0,1** | — | 0,1** | 0,2* | — | — | — | 1 |
| Нижние границы оценок (y_s^-) | — | — | 0,1 | — | — | 0,2 | 0,1 | 0,2 | — |
| Верхние границы оценок (y_s^+) | — | — | 1 | — | — | 1 | 1 | 1 | — |
| Шаг 2 | | | | | | | | | |
| Максимум | — | — | 0 | — | — | 0 | 0 | 0 | — |
| Ввод оценки экспертом (*) | | | | | | | | | |
| Оценки ситуации (y_s) | 0 | 0,1** | 0,8* | 0,1** | 0,2* | — | — | — | 1 |
| Изменение оценок после ввода оценки экспертом (**) | | | | | | | | | |
| Нижние границы оценок (y_s^-) | — | — | — | — | — | 0,8 | 0,1 | 0,2 | — |
| Верхние границы оценок (y_s^+) | — | — | — | — | — | 1 | 1 | 1 | — |
| Оценки ситуации (y_s) | 0 | 0,1** | 0,8* | 0,1** | 0,2* | 0,9** | — | — | 1 |
| Нижние границы оценок (y_s^-) | — | — | — | — | — | — | 0,1 | 0,2 | — |
| Верхние границы оценок (y_s^+) | — | — | — | — | — | — | 1 | 1 | — |
| Шаг 3 | | | | | | | | | |
| Максимум | — | — | — | — | — | — | 0 | 0 | — |
| Ввод оценки экспертом (*) | | | | | | | | | |
| Оценка экспертом (*) | 0 | 0,1** | 0,8* | 0,1** | 0,2* | 0,9** | 0,7* | — | 1 |
| Изменение оценок после ввода оценки экспертом (**) | | | | | | | | | |
| Нижние границы оценок (y_s^-) | — | — | — | — | — | — | — | 0,7 | — |
| Верхние границы оценок (y_s^+) | — | — | — | — | — | — | — | 1 | — |
| Оценки ситуации (y_s) | 0 | 0,1** | 0,8* | 0,1** | 0,2* | 0,9** | 0,7* | 0,85** | 1 |

матически. В данной ситуации они оказываются равными 0,1. После этого границы значения нижних и верхних оценок частоты наступления НС вновь пересчитываются. Далее, шаги 2 и 3 выполняются аналогично рассмотренному шагу 1. В табл. 3 символ «—» означает, что для ситуаций с определенными значениями оценок частоты значения нижних и верхних оценок частоты наступления НС не указываются. В рассмотренном примере вместо 7 вопросов, которые должны быть заданы эксперту, было задано 3 вопроса. Таким образом, предложенный подход позволяет сформировать тактику интервьюирования и организации процесса извлечения знаний, который обеспечивает сокращение времени интервьюирования эксперта.

Обобщенные результаты приведенных выше утверждений позволяют сформировать общие рекомендации для инженера по знаниям при организации процесса приобретения знаний коммуникативными методами при разработке баз знаний экспертных систем. Изложенные рекомендации также могут быть использованы для реализации проведения интервьюирования экспертов в автоматизированном режиме. В этом случае эксперт будет опрошен компонентом приобретения знания ЭС в автоматизированном режиме диалога, в процессе которого должны определять-

ся последовательность задаваемых вопросов и производиться оценки частот наступления НС в соответствии с вышеизложенными положениями.

4. Экспериментальные результаты

Для подтверждения применения предложенного метода были проведены эксперименты с имитацией ответов эксперта. Результаты проведенных экспериментов приведены в таблице 4.

В экспериментах принималось, что каждая типовая нечеткая ситуация характеризуется количеством признаков в диапазоне от 5 до 7, а каждый признак принимает три нечетких значения. В каждом эксперименте задавалась точность оценки частоты наступления НС: 0,33 либо 0,5. При имитации ответа эксперта устанавливалось значение оценки частоты наступления НС случайно в соответствии с равномерным законом распределения на отрезке допустимого интервала значений: $[y_s^-, y_s^+]$, $s \in S$. Равномерный закон распределения был выбран в связи с тем, что ни одно из событий, связанных с выбором оценки частоты наступления нештатных событий, не является объективно более возможным, чем другое, ввиду отсутствия априори информации о каком-либо выборе. Последний столбец представляет собой процентное отношение количества заданных вопросов в эксперименте к максимально возможному количеству вопросов.

Таблица 4. Результаты проведенных экспериментов

| Количество признаков | Максимальное количество вопросов (количество типовых ситуаций) | Требуемая точность, $y_s^+ - y_s^-$, $s \in S$ | Количество вопросов, заданных эксперту | Процент количества задаваемых вопросов, % |
|----------------------|--|---|--|---|
| 5 | 243 | 0,33 | 195 | 80 |
| 5 | 243 | 0,5 | 88 | 36 |
| 6 | 729 | 0,33 | 461 | 63 |
| 6 | 729 | 0,5 | 115 | 16 |
| 7 | 2187 | 0,33 | 1288 | 58 |
| 7 | 2187 | 0,5 | 232 | 11 |

Результаты экспериментов демонстрируют, что предложенная методика приводит к сокращению количества вопросов, задаваемых эксперту, и, как следствие, обеспечивает снижение трудоемкости процесса формирования оценок частоты возникновения НС. Требование точности, очевидно, также влияет на количество вопросов, задаваемых эксперту. Очевидно, что усиление требований к точности приводит к увеличению количества вопросов.

5. Заключение

Разработка ЭС представляет собой одно из направлений внедрения и применения искусственного интеллекта. Базы знаний экспертных систем аккумулируют формализованное представление навыков человека, способного качественно решать задачи. Для обеспечения автоматизированного решения задач привлекаются опытные профессионалы, по результатам собеседования с которыми формируются правила вывода для автоматизированной интеллектуальной системы. Поскольку предметные области реализации интеллектуальных систем в большей степени связаны с плохо структурированными задачами, уместно использовать подход, основанный на теории нечеткого моделирования. В этом случае нечеткие правила имитируют суждение лица, принимающего решение [Suzdaltsev, Suzdaltsev, Bogula, 2017]. Для формирования нечетких правил необходимо вовлечение экспертов в процесс построения базы знаний. Интервьюирование экспертов — дорогостоящая процедура (как по стоимости, так и по временным критериям). Поэтому важно обосновать разумный подход к организации интервью в процессе извлечения

знаний. Представленный метод был использован в процессе извлечения экспертных знаний для получения знаний с целью формирования нечетких правил в области медицины, когда разрабатывались экспертные системы прогнозирования патологий новорожденных и беременности [Аскаророва и др., 2007]. Практическое использование, как и результаты проведенного моделирования с имитацией ответа эксперта, показали, что за счет введения отношения порядка для набора значений признаков можно уменьшить количество вопросов, задаваемых эксперту. Так, при разработке экспертных систем для прогнозирования различных групп риска беременных и новорожденных, при подготовке базы знаний было достигнуто снижение количества вопросов эксперту примерно от 20% до 40%. Область медицины — это обширная область для реализации нечетких моделей рассуждений. Тем не менее описанный подход применим и при разработке ЭС прогнозной диагностики в других предметных областях, включая технические объекты, когда прогнозирование НС основано только на профессиональных, качественных оценках специалистов-экспертов при отсутствии фактических данных или результатов экспериментов.

Список литературы (References)

- Аверкин А. Н., Батыршин И. З., Блишун А. Ф., Силов В. Б., Тарасов В. Б.* Нечеткие множества в моделях управления и искусственного интеллекта / под ред. Д. А. Пospelova. — М.: Наука, 1986.
- Averkin A. N., Batyrshin I. Z., Blishun A. F., Silov V. B., Tarasov V. B.* Nечetkie mnozhestva v modelyakh upravleniya i iskusstvennogo intellekta [Fuzzy sets in models of control and artificial intelligence] / pod red. D. A. Pospelova. — Moscow: Nauka, 1986 (in Russian).
- Аскаророва З. И., Суздальцев В. А., Сабиров И. Х., Хасанов А. А.* Прогнозирование группы риска беременных по развитию аномалий сократительной деятельности матки // Сб. Всероссийской научной конференции, 30–31 мая. — Казань: Изд-во КГТУ, 2007.
- Askarova Z. I., Suzdal'tsev V. A., Sabirov I. Kh. Khasanov A. A.* Prognozirovanie gruppy riska beremennykh po razvitiyu anomalii sokratitel'noi deyatel'nosti matki [Forecasting the risk group of pregnant women for the development of anomalies of the contractile activity of the uterus] // Sb. Vserossiiskoi nauchnoi konferentsii, 30–31 maya. — Kazan': Izd-vo KGTU, 2007 (in Russian).
- Гаврилова Т. А., Червинская К. Р.* Извлечение и структурирование знаний для экспертных систем. — М.: Радио и связь, 1992. — 200 с.
- Gavrilova T. A., Chervinskaya K. R.* Izvlechenie i strukturirovanie znaniy dlya ekspertnykh system [Extraction and structuring of knowledge for expert systems]. — Moscow: Radio i svyaz', 1992 (in Russian).
- Городецкий А. Е., Курбанов В. Г., Тарасова И. Л.* Экспертная система анализа и прогнозирования аварийных ситуаций на электростанциях // Информационно-управляющие системы. — 2012. — № 4. — С. 59–63.
- Gorodetskii A. E., Kurbanov V. G., Tarasova I. L.* Ekspertnaya sistema analiza i prognozirovaniya avariinykh situatsii na elektrostantsiyakh [Expert system for analysis and prediction of emergency situations at power plants] // Information and control systems. — 2012. — No. 4. — P. 59–63 (in Russian).
- Заде Л. А.* Понятие лингвистической переменной и его применение к принятию приближенных решений. — М.: Мир, 1976. — 165 с.
- Zadeh L. A.* The concept of a linguistic variable and its application to approximate reasoning // Information Sciences. — Vol. 8, Issue 3. — P. 199–249. DOI: 10.1016/0020-0255(75)90036-5 (Russ. ed.: Zadeh L. A. Ponyatiye lingvisticheskoy peremennoy i yego primeneniye k prinyatiyu priblizhennykh resheniy. — Moscow: Mir, 1976.)
- Козлова Т. Д., Игнат'ев А. А., Самойлова Е. М.* Внедрение экспертной системы поддержки принятия решений для определения неисправностей технологической системы // Вестник Саратовского государственного технического университета. — 2011. — Т. 2, № 2. — С. 219–224.
- Kozlova T. D., Ignat'ev A. A., Samoilova E. M.* Vnedrenie ekspertnoi sistemy podderzhki prinyatiya reshenii dlya opredeleniya neispravnostei tekhnologicheskoi sistemy [Implementation of an expert decision support system for determining faults in a technological system] // Vestnik Saratovskogo gosudarstvennogo tekhnicheskogo universiteta. — 2011. — Vol. 2, No. 2. — P. 219–224 (in Russian).
- Леденева Т. М., Сергиенко М. А.* Оптимальное построение базы нечетких правил // Системы управления и информационные технологии. — 2008. — Т. 34, № 4. — С. 34–38.
- Ledeneva T. M., Sergiyenko M. A.* Optimal'noye postroyeniye bazy nechetkikh pravil [Optimal design of a fuzzy rule base] // Automation and Remote Control. — 2008. — Vol. 34, No. 4. — P. 34–38. DOI: 10.1134/S0005117912110173 (in Russian).

- Мелихов А. Н., Берштейн Л. С., Коровин С. Я.* Ситуационные советующие системы с нечеткой логикой. — М.: Наука, Гл. ред. физ.-мат. лит., 1990. — 272 с.
Melikhov A. N., Bershteyn L. S., Korovin S. Ya. Situatsionnyye sovetuyushchiye sistemy s nechetkoy logikoy [Situational advising systems with fuzzy logic]. — Moscow: Nauka, Gl. red. fiz.-mat. lit., 1990 (in Russian).
- Пегат А.* Нечеткое моделирование и управление / А. Пегат; пер. с англ. — 2-е изд. — М.: Бином, Лаборатория знаний, 2013. — 798 с.
Piegat A. Fuzzy modeling and control. — Heidelberg: Physical-Verlag, 2001. (Russ. ed.: *Piegat A.* Nечetkoye modelirovaniye i upravleniye. — Moscow: Binom, Laboratoriya znaniy, 2013.)
- Рутковская Д., Пилински М., Рутковски И.* Нейронные сети, генетические алгоритмы и нечеткие системы. — М.: Горячая линия – Телеком, 2013. — 384 с.
Rutkovskaya D., Pilinski M., Rutkovski I. Neyronnyye seti, geneticheskiye algoritmy i nechetkiye sistemy [Neural networks, genetic algorithms and fuzzy systems]. — Moscow: Goryachaya liniya – Telekom, 2013 (in Russian).
- Шевелев А. А.* Байесовский подход к оценке влияния внешних шоков на макроэкономические показатели России // Мир экономики и менеджмента. — 2017. — Т. 17, № 1. — С. 26–40.
Shevelev A. A. Baiesovskii podkhod k otsenke vliyaniya vneshnikh shokov na makroekonomicheskie pokazateli Rossii [Bayesian approach to evaluate the impact of external shocks on Russian macroeconomics indicators] // Mir ekonomiki i menedzhmenta. — 2017. — Vol. 17, No. 1. — P. 26–40 (in Russian).
- Boose J.* Expertise Transfer for Expert System Design. — Elsevier, 1986. — 312 p.
- Davis R.* Interactive Transfer of expertise: Acquisition of inference rules // 5th International Joint Conference on Artificial Intelligence. — Boston: MIT, 1977. — P. 321–328.
- Gammack J. K., Young R. M.* Psychological techniques for electing expert knowledge // Research and Development in Expert Systems. — Cambridge: University Press, 1985. — P. 105–112.
- Kitto C. M., Boose J. H.* Heuristics for expertise transfer: an implementation of a dialogue manager for knowledge acquisition // Inc. J. of Man. Machine Studies. — 1987. — Vol. 26, No. 2. — P. 183–202.
- Kitto C. M., Boose J. H.* Selecting knowledge acquisition tools and strategies based on application characteristics // International Journal of Man-Machine Studies. — 1989. — Vol. 31, No. 2. — P. 149–160.
- McDermott J.* R1: An expert in the computer system domain // The first National Conference of Artificial Intelligence. — Stanford: Stanford University Press, 1980. — P. 269–271.
- Olson J. R., Henry H.* Extracting expertise from experts: Methods for knowledge acquisition // Expert Systems. — 1987. — Vol. 4, No. 3. — P. 152–168.
- Rizaev I. S., Takhavova E. G.* Statistical methods of making decision on repair of technical means // 2016 2nd International Conference on Industrial Engineering, Applications and Manufacturing, ICIEAM 2016. — Proceedings 7911647.
- Shortliffe E. H.* Computer-based medical consultation: MYCIN. — New York: Academic Elsevier, 1976. — 264 p.
- Suzdaltsev V. A., Suzdaltsev I. V., Bogula N. Yu.* Fuzzy rules formation for the construction of the predictive diagnostics expert system // Proc. 20th IEEE International Conference on Soft Computing and Measurements, SCM 2017. — 2017. — P. 481–482.
- Wagner C.* Breaking the knowledge acquisition bottleneck through conversational knowledge management // Information Resources Management Journal. — 2006. — Vol. 19, No. 1. — P. 70–83.
- Wagner W. P., Chung Q. B., Naidawi M. K.* The impact of problem domains and knowledge acquisition techniques: a content analysis of P/OM expert system case studies // Expert Syst. Appl. — 2003. — Vol. 24, No. 1. — P. 79–86.
- Zadeh L. A.* The role of fuzzy logic in the management of uncertainty in expert systems // Fuzzy Sets and Systems. — 1983. — Vol. 11. — P. 199–227.