

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ INFORMATION TECHNOLOGIES



УДК 004.912
<https://doi.org/10.37661/1816-0301-2021-18-4-7-16>

Оригинальная статья
Original Paper

Информационная поддержка принятия решений при описании проблемной ситуации

С. Ф. Липницкий

*Объединенный институт проблем информатики
Национальной академии наук Беларуси,
ул. Сурганова, 6, Минск, 220012, Беларусь
✉ E-mail: lipn@newman.bas-net.by*

Аннотация

Цели. Решается задача автоматизации информационной поддержки пользователя в системе принятия решений на этапе описания проблемной ситуации. Актуальность задачи связана с необходимостью сбора и обработки значительных объемов информации, поскольку при наличии большого количества факторов возможности человека часто оказываются недостаточными для поиска и систематизации нужных сведений. При решении задачи информационной поддержки пользователя на этапе описания проблемной ситуации преследуются три основные цели: построение математической модели соответствующих процессов, формирование и формализация совокупности основных понятий модели, разработка в рамках моделирования алгоритмов реализации взаимодействия пользователя с информационной системой.

Методы. Используются методы теории множеств, теории вероятностей и теории графов.

Результаты. Разработана математическая модель информационной поддержки пользователя на этапе описания проблемной ситуации. В процессе взаимодействия с пользователем система предъявляет ему для заполнения специальные шаблоны предложений и текстов. Вместе с шаблонами пользователь получает от системы тексты подсказок. Они синтезируются на основе разработанной автором ранее модели представления знаний в виде вербальных ассоциаций, т. е. семантических связей между словами и словосочетаниями, соответствующих ассоциативным отношениям между обозначаемыми ими сущностями в реальном мире.

Заключение. В качестве реализации предложенной в статье модели разработаны следующие алгоритмы: создания словаря коммуникативных фрагментов; создания фрагментно-слотовых шаблонов предложений, текстов и предметных областей; информационной поддержки пользователя. Словарь коммуникативных фрагментов создается в четыре шага в соответствии с их формальным определением. На каждом шаге последовательно проверяются четыре условия из данного определения. Фрагментно-слотовые шаблоны предложений формируются путем замены их базовых коммуникативных фрагментов слотами, а шаблоны текстов – как кортежи шаблонов их предложений. Фрагментно-слотовые шаблоны предметных областей создаются в виде реализации редукций бинарных отношений на множествах шаблонов предложений из соответствующих тематических корпусов текстов. Каждый тематический корпус текстов определяет некоторую предметную область.

Ключевые слова: коммуникативный фрагмент, корпус текстов, математическая модель, система принятия решений, фрагментно-слотовый шаблон

Для цитирования. Липницкий, С. Ф. Информационная поддержка принятия решений при описании проблемной ситуации / С. Ф. Липницкий // Информатика. – 2021. – Т. 18, № 4. – С. 7–16. <https://doi.org/10.37661/1816-0301-2021-18-4-7-16>

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию | Received 23.07.2021
Подписана в печать | Accepted 13.08.2021
Опубликована | Published 29.12.2021

Information support for decision making in problem situation description

Stanislav F. Lipnitsky

*The United Institute of Informatics Problems
of the National Academy of Sciences of Belarus,
st. Surganova, 6, Minsk, 220012, Belarus
✉E-mail: lipn@newman.bas-net.by*

Abstract

Objectives. The problem of automating the user information support in decision making system at the stage of describing the problem situation is solved. The relevance of the problem is associated with the need to collect and process significant amounts of information, since in the presence of a large number of factors, a person's capabilities are often insufficient to search and organize the necessary information.

When solving the problem of user information support at the stage of describing the problem situation, three main goals are pursued: building a mathematical model of the corresponding processes; formalization of the set of basic concepts of the model; development of the algorithms for implementation of user interaction with the information system.

Methods. Methods of set theory, probability theory and graph theory are used.

Results. A mathematical model of user information support at the stage of describing a problem situation has been developed. In the process of interacting with the user, the system suggests special templates of sentences and texts for filling. Along with templates, the user receives help texts from the system. They are generated on the basis of the previously developed model of knowledge representation in the form of verbal associations, that is, semantic links between words and phrases corresponding to associative relationships between the entities they designate in the real world.

Conclusion. As an implementation of the proposed model, the following algorithms have been developed: an algorithm for creating a dictionary of communicative fragments; algorithms for creating fragment-slot templates for sentences, texts and subject areas; an algorithm of user information support. The vocabulary of communicative fragments is created in four steps in accordance with their formal definition. At each step, four conditions from the given definition are tested sequentially. Fragment-slot templates of sentences are formed by replacing their basic communicative fragments with slots, and text templates – as tuples of templates of their sentences. Fragment-slot templates of subject areas are created in the form of implementation of binary relations reductions on the sets of sentence templates from the corresponding thematic text corpora. Each thematic body of texts defines a certain subject area.

Keywords: communicative fragment, text corpus, mathematical model, decision making system, fragment-slot template

For citation. Lipnitsky S. F. *Information support for decision making in problem situation description*. Informatika [Informatics], 2021, vol. 18, no. 4, pp. 7–16 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-4-7-16>

Conflict of interest. The author declare of no conflict of interest.

Введение. Проблемная ситуация в любой предметной области характеризуется наличием некоторой проблемы, а также условий и обстоятельств ее возникновения. Под проблемой обычно понимают совокупность сложных теоретических или практических задач, требующих постановки и решения. При этом правильное описание проблемной ситуации является важнейшим условием принятия адекватного ситуации решения [1–3].

Оценка эффективности решения и его принятие сопряжены, как правило, с необходимостью сбора и обработки значительных объемов информации. При наличии большого количества факторов, связанных с проблемной ситуацией, возможности человека часто оказываются недостаточными для поиска и систематизации нужных сведений. В подобных обстоятельствах используются специальные информационные системы – системы поддержки принятия решений.

В настоящей статье решается задача информационной поддержки пользователя (лица, принимающего решение) на этапе описания проблемной ситуации. В процессе взаимодействия с пользователем система предъявляет ему для заполнения специальные (фрагментно-слотовые) шаблоны предложений и тексты подсказок. Они синтезируются на основе разработанной автором модели представления знаний в виде вербальных ассоциаций, т. е. семантических связей между словами и словосочетаниями, соответствующих ассоциативным отношениям между обозначаемыми ими сущностями в реальном мире [4].

Итеративность процесса принятия решения. Задача принятия решения возникает всякий раз, когда имеется несколько альтернативных вариантов для достижения некоторой цели. При этом наиболее эффективное решение приходится выбирать по совокупности показателей (критериев), т. е. в условиях многокритериальности. В различных предметных областях процессы принятия решений имеют много общего. Они включают следующие основные этапы:

- описание проблемной ситуации и постановка задачи принятия решения;
- поиск вариантов (альтернатив) решения поставленной задачи;
- выбор критериев оценки альтернатив для описания вариантов решения;
- выявление ограничений на критерии;
- принятие решения с учетом результатов оценки альтернатив.

Некоторые из перечисленных этапов могут повторяться, т. е. процесс принятия решения имеет итеративный характер (рис. 1).

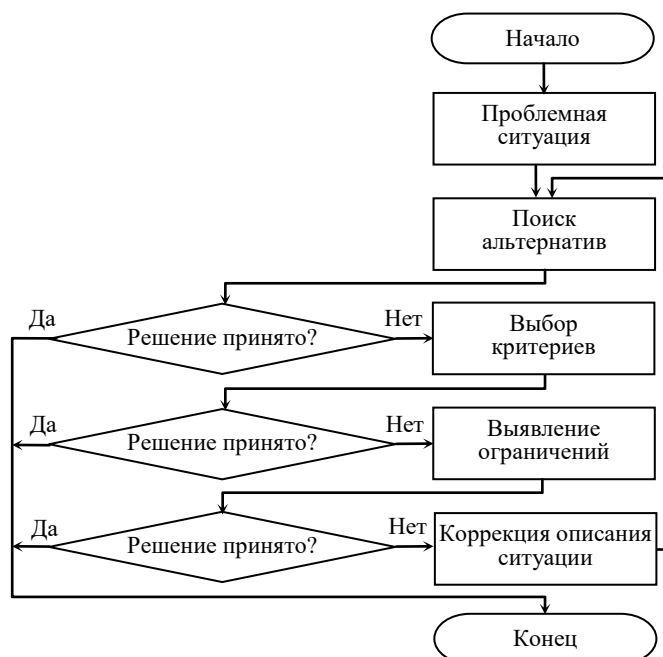


Рис. 1. Схема процесса принятия решения
 Fig. 1. Diagram of the decision making process

Создание словаря коммуникативных фрагментов. Традиционно синтез предложений языка (текста) реализуется как последовательная генерация лексем, затем синтаксических фраз и, наконец, предложений по известным синтаксическим правилам. Однако в соответствии с результатами исследования этой проблемы, изложенными в монографии [5], предложения строятся из готовых, хранящихся в памяти компонентов, названных коммуникативными фрагментами. Эти фрагменты не образуются по синтаксическим правилам, а извлекаются из памяти целиком.

При создании словаря коммуникативных фрагментов будем использовать формулы для вычисления информативности вербальной ассоциации между словами и словосочетаниями, предложенные автором в статьях [6, 7]. Под вербальными ассоциациями в компьютерной лингвистике понимают семантические связи между словами, предложениями и текстами в языке, соответствующие ассоциативным отношениям между обозначаемыми ими сущностями в реальном мире [8].

Определение понятия коммуникативного фрагмента. Обозначим через $\pi = a_1 a_2 \dots a_n$ произвольное предложение из тематического корпуса текстов Ct . (Тематический корпус – это совокупность текстов по конкретной тематике, объединение всех тематических корпусов образует полный корпус Cf .) Определим понятие коммуникативного фрагмента посредством следующих четырех условий:

1. Если $n = 1$, то слово a_1 цепочки π назовем коммуникативным фрагментом.
2. Если $n \geq 2$ и $I_{Cf}^{a_1 a_2} < I_{Cf}^{00}$ (I_{Cf}^{00} – пороговое значение информативности), то слово a_1 будем называть коммуникативным фрагментом. Информативность вербальной ассоциации между словами вычисляется по формуле

$$I_{Cf}^{ab} = \frac{n_{Cf}^{ab} + n_{Cf}^{Par_{ab}} + n_{Cf}^{Syn_{ab}}}{N_{Cf}}, \quad (1)$$

где n_{Cf}^{ab} – количество всех предложений в полном корпусе текстов Cf , в которых присутствуют слова a и b или их синонимы и словоизменения, а N_{Cf} – количество всех предложений в корпусе Cf . Параметры $n_{Cf}^{Par_{ab}}$ и $n_{Cf}^{Syn_{ab}}$ в формуле (1) указывают на число вхождений всех пар словоформ, являющихся словоизменениями слов a и (или) b (и, соответственно, их синонимами) и встречающихся в одном и том же предложении корпуса текстов Cf :

$$n_{Cf}^{Par_{ab}} = \sum_{\substack{c \in Par_a, d \in Par_b, \\ c \neq a \text{ и (или) } d \neq b}} n_{Cf}^{cd}, \quad n_{Cf}^{Syn_{ab}} = \sum_{\substack{d \in Syn_a, f \in Syn_b, \\ d \neq a \text{ и (или) } f \neq b}} n_{Cf}^{df}.$$

Словоизменения и синонимы для данных параметров содержат специальные лингвистические словари:

- словарь словоизменительных парадигм

$$Dic_{par} = \{(a, Par_a) \mid a \in W_{Cf}, a \in Par_a\},$$

состоящий из пар $\langle \text{словоформа}, \text{парадигма} \rangle$, где W_{Cf} – множество всех словоформ полного корпуса текстов Cf , а Par_a – совокупность всех словоизменений словоформы a ;

- словарь синонимичных словоформ

$$Dic_{syn} = \{(a, Syn_a) \mid a \in W_{Cf}, a \in Syn_a\},$$

включающий в себя пары $\langle \text{словоформа}, \text{синонимичные словоформы} \rangle$, в которых каждой словоформе a соответствует множество ее синонимов Syn_a .

3. Пусть $2 \leq m < n$. Подцепочку $a_1 a_2 \dots a_m$ цепочки π назовем коммуникативным фрагментом, если справедлива последовательность неравенств $I_{Cf}^{a_1 a_2} \geq I_{Cf}^{00}$, $I_{Cf}^{(a_1 a_2) a_3} \geq I_{Cf}^{00}$, ..., $I_{Cf}^{(a_1 a_2 \dots a_{m-1}) a_m} \geq I_{Cf}^{00}$, а $I_{Cf}^{(a_1 a_2 \dots a_m) a_{m+1}} < I_{Cf}^{00}$.

4. Если при $n \geq 2$ выполняется последовательность неравенств $I_{Cf}^{a_1 a_2} \geq I_{Cf}^{00}$, $I_{Cf}^{(a_1 a_2) a_3} \geq I_{Cf}^{00}$, ..., $I_{Cf}^{(a_1 a_2 \dots a_{n-1}) a_n} \geq I_{Cf}^{00}$, то цепочку $a_1 a_2 \dots a_n$ назовем коммуникативным фрагментом. Значения информативности вербальной ассоциации в пп. 3 и 4 вычисляются по формуле

$$I_{Cf}^{\pi b} = \frac{\sum_{a \in \pi} I_{Cf}^{ab}}{\sqrt{\sum_{a \in \pi} (I_{Cf}^{ab})^2}}, \quad (2)$$

являющейся частным случаем выражения

$$I_{Cf}^{\pi \rho} = \frac{\sum_{a \in \pi, b \in \rho} I_{Cf}^{ab}}{\sqrt{\sum_{a \in \pi, b \in \rho} (I_{Cf}^{ab})^2}}, \quad (3)$$

где $I_{Cf}^{\pi \rho}$ – информативность вербальной ассоциации между словосочетаниями (фразами) π и ρ , а $I_{Cf}^{\pi b}$ – между словосочетанием π и словом b .

Базовые и связующие коммуникативные фрагменты. Пусть Ct – некоторый тематический корпус текстов. Рассмотрим предметную область, определяемую корпусом Ct . Будем различать базовые и связующие коммуникативные фрагменты в зависимости от их информативности в данной предметной области. Обозначим через I_{Ct}^0 пороговое значение информативности коммуникативного фрагмента. Тогда коммуникативный фрагмент f будем называть базовым, если значение его информативности I_{Ct}^f удовлетворяет неравенству $I_{Ct}^f \geq I_{Ct}^0$. Если же $I_{Ct}^f < I_{Ct}^0$, то фрагмент f назовем связующим. Связующим, например, является коммуникативный фрагмент «предлагается новый подход к решению проблемы», а базовым – фрагмент «принятия решений в условиях неопределенности».

Обозначим через $Ft_{\text{баз.}}$ множество всех базовых коммуникативных фрагментов, а через $Ft_{\text{св.}}$ – множество всех связующих. Тогда множество всех коммуникативных фрагментов предметной области – это объединение множеств базовых и связующих фрагментов, т. е. $F_{Ct} = Ft_{\text{баз.}} \cup Ft_{\text{св.}}$.

Информативность коммуникативного фрагмента $f = ab \dots$ в тематическом корпусе текстов Ct вычисляется по формуле

$$I_{Ct}^f = \frac{I_{Ct}^a + I_{Ct}^b + \dots}{\sqrt{(I_{Ct}^a)^2 + (I_{Ct}^b)^2 + \dots}}, \quad (4)$$

где $I_{Ct}^a, I_{Ct}^b, \dots$ – значения информативности всех слов цепочки f [6, 7]. В свою очередь, информативность слова в корпусе Ct определяется выражением

$$I_{Ct}^a = \frac{n_{Ct}^a + n_{Ct}^{\text{Par}_a} + n_{Ct}^{\text{Syn}_a}}{n_{Cf}^a + N_{Cf}^{\text{Par}_a} + N_{Cf}^{\text{Syn}_a}}. \quad (5)$$

Значения частот n_{Cf}^a и n_{Ct}^a в формуле (5) ищутся в частотном словаре словоформ

$$Dic_a = \{ \langle a, n_{Cf}^a, n_{Ct_1}^a, n_{Ct_2}^a, \dots, n_{Ct_n}^a \rangle \mid a \in W_{Cf} \},$$

где W_{Cf} – множество всех словоформ полного корпуса текстов Cf , а $n_{Cf}^a, n_{Ct_1}^a, n_{Ct_2}^a, \dots, n_{Ct_n}^a$ – частоты встречаемости словоформы a во всех тематических корпусах текстов.

Алгоритм создания словаря коммуникативных фрагментов. На входе алгоритма – все предложения вида $\pi = a_1 a_2 \dots a_n$ из полного корпуса текстов C_f , на выходе – словарь коммуникативных фрагментов

$$Dic_f = \{ \langle f, I_{C_{t_1}}^f, I_{C_{t_2}}^f, \dots, I_{C_{t_m}}^f \rangle \mid f \in F \},$$

где $I_{C_{t_i}}^f$ ($i = \overline{1, m}$) – значение информативности коммуникативного фрагмента f в тематическом корпусе текстов C_{t_i} , F – множество всех коммуникативных фрагментов в полном корпусе текстов C_f . Алгоритм включает четыре шага:

1. Проверяется условие 1 из определения коммуникативного фрагмента. Если оно выполняется, то слово a_1 помещается в словарь Dic_f . По формуле (4) вычисляются и помещаются в Dic_f значения информативности слова a_1 в тематических корпусах текстов C_{t_i} ($i = \overline{1, m}$). Далее обрабатывается очередное предложение из корпуса C_f .

2. Проверяется условие 2 из определения коммуникативного фрагмента. При его выполнении слово a_1 цепочки π помещается в словарь Dic_f . Вычисляются и помещаются в Dic_f значения информативности слова a_1 в тематических корпусах текстов. Оставшиеся слова предложения π перенумеровываются, и процесс создания словаря коммуникативных фрагментов начинается с первого шага.

3. Проверяется условие 3 из определения коммуникативного фрагмента. Если оно выполнено, то подцепочка $a_1 a_2 \dots a_m$ предложения π помещается в словарь Dic_f . По формулам (3) и (4) вычисляются и помещаются в Dic_f значения информативности коммуникативного фрагмента в тематических корпусах текстов C_{t_i} ($i = \overline{1, m}$). Оставшиеся слова предложения π перенумеровываются, и алгоритм далее выполняется, начиная с первого шага.

4. Проверяется условие 4 из определения коммуникативного фрагмента. Если оно справедливо, то предложение π помещается в словарь Dic_f . По формулам (3) и (4) вычисляются и помещаются в Dic_f значения информативности коммуникативного фрагмента в тематических корпусах текстов C_{t_i} ($i = \overline{1, m}$). Далее обрабатывается очередное предложение из корпуса C_f , начиная с первого шага.

Построение фрагментно-слотовых шаблонов. Будем различать фрагментно-слотовые шаблоны предложений, текстов и предметных областей. Фрагментно-слотовый шаблон предложения – это цепочка, полученная заменой его базовых коммуникативных фрагментов слотами («пустыми» фрагментами). Фрагментно-слотовые шаблоны текстов образуются из кортежей шаблонов их предложений. Шаблоны предметных областей формируются как бинарные отношения на множествах шаблонов предложений из соответствующих тематических корпусов текстов.

Алгоритм создания фрагментно-слотовых шаблонов предложений. На входе алгоритма – предложение $\pi = a_1 a_2 \dots a_n$ из тематического корпуса текстов C_t , на выходе – фрагментно-слотовый шаблон предложения $h_\pi = s_1 s_2 \dots s_r$, где s_j ($j = \overline{1, r}$, $r \leq n$) – коммуникативный фрагмент или слот. Алгоритм состоит из трех шагов:

1. С использованием словаря Dic_f в предложении π выявляются коммуникативные фрагменты: $\pi = f_1 f_2 \dots f_r$.

2. В предложении $\pi = f_1 f_2 \dots f_r$ все базовые коммуникативные фрагменты заменяются слотами. В результате получаем фрагментно-слотовый шаблон предложения $h_\pi = s_1 s_2 \dots s_r$.

3. Аналогичным образом создаются фрагментно-слотовые шаблоны других предложений из текстов корпуса C_t .

Алгоритм создания фрагментно-слотовых шаблонов текстов. На входе алгоритма – текст T из тематического корпуса C_t в виде кортежа предложений $T = \langle \pi_1, \pi_2, \dots, \pi_m \rangle$, на выходе – фрагментно-слотовый шаблон $H_T = \langle h_{\pi_1}, h_{\pi_2}, \dots, h_{\pi_m} \rangle$ текста T . Алгоритм включает три шага:

1. Формируются фрагментно-слотовые шаблоны всех предложений $\pi_1, \pi_2, \dots, \pi_m$ текста T .

2. Сформированные шаблоны предложений $\pi_1, \pi_2, \dots, \pi_m$ помещаются в кортеж $H_T = \langle h_{\pi_1}, h_{\pi_2}, \dots, h_{\pi_m} \rangle$.

3. Аналогично создаются шаблоны других текстов из тематического корпуса St .

Алгоритм создания фрагментно-слотовых шаблонов предметных областей. Обозначим через $U_{St} = \{h_{\pi_i} \mid i = \overline{1, l}\}$ множество всех фрагментно-слотовых шаблонов предложений в предметной области, определяемой тематическим корпусом текстов St . Определим на множестве U_{St} отношение строгого порядка (антирефлексивное и транзитивное бинарное отношение) Ω_{St} . Обозначим через Ω_{St}^R редукцию $\Omega_{St}^R = \Omega_{St} \setminus (\Omega_{St})^2$ строгого порядка Ω_{St} . Содержательно редукция Ω_{St}^R соответствует отношению подчинения фрагментно-слотовых шаблонов предложений. Отношение Ω_{St}^R назовем *фрагментно-слотовым шаблоном предметной области, определяемой корпусом текстов St* . На практике элементами отношения Ω_{St}^R являются пары соседних фрагментно-слотовых шаблонов предложений, таких, что для любых шаблонов предложений h_{π_p} и h_{π_q} ($h_{\pi_p}, h_{\pi_q} \in \Omega_{St}^R$) тогда и только тогда, когда в данной предметной области существует фрагментно-слотовый шаблон текста, в котором шаблон h_{π_p} непосредственно предшествует шаблону h_{π_q} .

Приведем алгоритм формирования шаблона предметной области. На входе алгоритма – тематический корпус текстов St , на выходе – фрагментно-слотовый шаблон предметной области в виде отношения Ω_{St}^R . Алгоритм состоит из трех шагов:

1. Формируются фрагментно-слотовые шаблоны предложений из отобранных в корпусе St «хороших» текстов.
2. Формируются фрагментно-слотовые шаблоны «хороших» текстов.
3. Из сформированных шаблонов текстов выбираются все пары соседних шаблонов предложений и помещаются в множество Ω_{St}^R .

Информационная поддержка пользователя при описании проблемной ситуации. Сообщения пользователю синтезируются системой поддержки принятия решений путем заполнения слотов фрагментно-слотовых шаблонов дискурсивно-сочетаемыми коммуникативными фрагментами.

Отношение дискурсивной сочетаемости коммуникативных фрагментов. Под дискурсом в лингвистике понимают связную последовательность предложений, обладающую семантическим единством [9]. Для получения «хороших» предложений при их синтезе из коммуникативных фрагментов будем использовать отношение дискурсивной сочетаемости таких фрагментов. Понятие этого отношения введем следующим образом.

Определим на множестве F всех коммуникативных фрагментов в полном корпусе текстов Cf антирефлексивное бинарное отношение Δ_{Cf} , такое, что для любых фрагментов $f, g \in F$ соотношение $(f, g) \in \Delta_{Cf}$ выполняется тогда и только тогда, когда в некотором тексте $T \in Cf$ существует предложение π , в котором коммуникативный фрагмент f непосредственно предшествует фрагменту g . Отношение Δ_{Cf} будем называть отношением дискурсивной сочетаемости коммуникативных фрагментов в тематическом корпусе текстов Cf .

Дискурсивно-сочетаемые пары коммуникативных фрагментов $(f, g) \in \Delta_{Cf}$ будем хранить в специальном списке – словаре дискурсивно-сочетаемых коммуникативных фрагментов:

$$Dic_{fg} = \{(f, g) \mid f \in F_{Cf}, g \in F_{Cf}, (f, g) \in \Delta_{Cf}\}.$$

Вербально-ассоциативные сети предметных областей. Пусть по-прежнему $Ft_{баз.}$ – множество базовых, $Ft_{св.}$ – множество связующих, а $F_{St} = Ft_{баз.} \cup Ft_{св.}$ – множество всех коммуникативных фрагментов в тематическом корпусе текстов St . Определим на множестве F_{St} отноше-

ние толерантности Θ_{Ct} (рефлексивное и симметричное бинарное отношение), такое, что пара (f, g) любых фрагментов из множества Ft является элементом отношения Θ_{Ct} , т. е. $(f, g) \in \Theta_{Ct}$ тогда и только тогда, когда фрагменты f и g из этой пары содержатся хотя бы в одном предложении корпуса Ct и информативность вербальной ассоциации $I_{Ct}^{fg} \geq I_{Ct}^0$, где I_{Ct}^0 – пороговое значение информативности, а информативность I_{Ct}^{fg} вычисляется по формуле (3).

Обозначим через G_{Ct} граф отношения Θ_{Ct} . Пусть (f, g) – произвольное ребро этого графа. Если $(f, g) \in \Delta_{Ct}$, то для всех таких пар (f, g) вершины f и g соединим дугой, направленной от f к g . Обозначим полученный смешанный граф через Net_{Ct} .

Граф Net_{Ct} назовем вербально-ассоциативной сетью предметной области, определяемой тематическим корпусом текстов Ct (рис. 2).

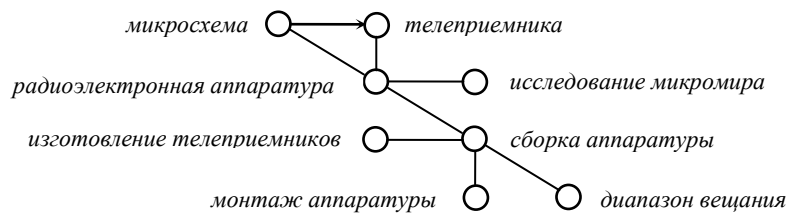


Рис. 2. Фрагмент вербально-ассоциативной сети

Fig. 2. Fragment of the verbal-associative network

Подграф $Net_{баз.}$ графа Net_{Ct} , порожденный множеством вершин $Ft_{баз.}$, будем называть базовой вербально-ассоциативной сетью предметной области, определяемой тематическим корпусом текстов Ct .

Алгоритм информационной поддержки. На входе алгоритма – перечень названий предметных областей, совокупность вербально-ассоциативных сетей для всех предметных областей и их фрагментно-слотовых шаблонов, на выходе – множество рекомендаций и подсказок пользователям системы поддержки принятия решений. Алгоритм включает следующие шаги:

1. Пользователю (в режиме диалога) предъявляется перечень названий предметных областей и предлагается выбрать требуемую. Пользователь указывает предметную область, которой соответствует тематический корпус текстов Ct .

2. В рамках выбранной предметной области пользователь имеет возможность выполнить одно из следующих действий:

– отметить одну или несколько вершин g_1, g_2, \dots базовой вербально-ассоциативной сети $Net_{баз.}$. Перейти к выполнению шага 3;

– сформировать одно или несколько предложений $\langle \rho_1, \rho_2, \dots \rangle$ текста описания проблемной ситуации. Перейти к выполнению шага 4.

3. Вычисляется информативность вербальной ассоциации между фрагментно-слотовыми шаблонами предложений из тематического корпуса текстов Ct и множеством $g_{баз.} = \{g_1, g_2, \dots\}$ отмеченных пользователем вершин базовой вербально-ассоциативной сети $Net_{баз.}$ по формуле

$$I_{Ct}^{h_{\pi_i} g_{баз.}} = \frac{\sum_{f \in h_{\pi_i}, g \in g_{баз.}} I_{Ct}^{fg}}{\sqrt{\sum_{f \in h_{\pi_i}, g \in g_{баз.}} (I_{Ct}^{fg})^2}}, \quad i = \overline{1, l}. \quad (6)$$

Информативность I_{Ct}^{fg} вычисляется по формуле (3). Далее фрагментно-слотовые шаблоны из множества $U_{Ct} = \{h_{\pi_i} | i = \overline{1, l}\}$ упорядочиваются по убыванию информативности $I_{Ct}^{h_{\pi_i} g_{баз.}}$ и предъявляются пользователю по его требованию в качестве форм для заполнения слотов

коммуникативными фрагментами. Пользователь выбирает один из шаблонов, например h_1 . Переход к выполнению шага 5.

4. Вычисляется информативность вербальной ассоциации между фрагментно-слотовыми шаблонами предложений из множества U_{C_i} и кортежем предложений $S = \langle \rho_1, \rho_2, \dots \rangle$ по формуле

$$I_{C_i}^{h_i, S} = \frac{\sum_{f \in h_{\pi_i}, \rho \in S} I_{C_i}^{f\rho}}{\sqrt{\sum_{f \in h_{\pi_i}, \rho \in S} (I_{C_i}^{f\rho})^2}}, \quad i = \overline{1, l}. \quad (7)$$

Информативность $I_{C_i}^{f\rho}$ вычисляется по формуле (5). Реализуется упорядочение фрагментно-слотовых шаблонов h_{π_i} по убыванию информативности $I_{C_i}^{h_i, S}$. Шаблоны предъявляются пользователю, который выбирает наиболее соответствующий его потребностям, например шаблон h_3 . Переход к выполнению шага 6.

5. Пользователь заполняет слоты фрагментно-слотового шаблона h_1 . Ему предлагаются варианты заполнения в виде совокупностей коммуникативных фрагментов, для которых информативность вербальной ассоциации с шаблоном h_1 не меньше некоторого порогового значения. Кроме того, эти фрагменты должны быть дискурсивно сочетаемыми с соседними фрагментами шаблона h_1 . Далее пользователю предъявляются шаблоны, удовлетворяющие отношению $\Omega_{C_i}^R$. Процесс их заполнения продолжается аналогичным образом.

6. Выполняется аналогично шагу 5.

Заключение. Предложена математическая модель информационной поддержки пользователя при принятии решений на этапе описания проблемной ситуации. Разработаны следующие реализующие данную модель алгоритмы:

- создания словаря коммуникативных фрагментов;
- создания фрагментно-слотовых шаблонов предложений, текстов и предметных областей;
- информационной поддержки пользователя.

Словарь коммуникативных фрагментов создается в четыре шага в соответствии с их формальным определением. На каждом шаге последовательно проверяются четыре условия из данного определения.

Фрагментно-слотовые шаблоны предложений формируются путем замены их базовых коммуникативных фрагментов слотами, а шаблоны текстов – как кортежи шаблонов их предложений.

Фрагментно-слотовые шаблоны предметных областей создаются в виде реализации редукций бинарных отношений на множествах шаблонов предложений из соответствующих тематических корпусов текстов. Каждый тематический корпус текстов определяет некоторую предметную область.

Информационная поддержка пользователя при описании проблемной ситуации осуществляется в режиме диалога и сводится к предъявлению ему для заполнения релевантных фрагментно-слотовых шаблонов и возможных вариантов их заполнения.

Список использованных источников

1. Кравченко, Т. К. Системы поддержки принятия решений / Т. К. Кравченко // Информационные технологии для современного университета ; под общ. ред. А. Н. Тихонова, А. Д. Иванникова. – М. : ГНИИ ИТТ «Информика», 2011. – С. 107–118.
2. Моисеенко, Е. В. Информационные технологии в экономике / Е. В. Моисеенко, Е. Г. Лаврушина ; ред. М. А. Касаткина. – М. : Софт, 2009. – С. 120–135.
3. Симанков, В. С. Методологическое обеспечение этапов поддержки принятия решений при синтезе сложных систем / В. С. Симанков, А. Н. Черкасов // Перспективы науки. – 2012. – № 12. – С. 85–89.
4. Липницкий, С. Ф. Модель представления знаний в информационных системах на основе вербальных ассоциаций / С. Ф. Липницкий // Информатика. – 2011. – № 4(32). – С. 21–28.

5. Гаспаров, Б. М. Язык, память, образ. Лингвистика языкового существования / Б. М. Гаспаров. – М. : Новое литературное обозрение, 1996. – 352 с.
6. Липницкий, С. Ф. Синтез запросов и поиск альтернатив в системе информационной поддержки принятия решений / С. Ф. Липницкий // Проблемы физики, математики и техники. – 2020. – № 2. – С. 91–95.
7. Липницкий, С. Ф. Интернет-поиск и лексико-семантическая обработка аналогов принятых решений в различных предметных областях / С. Ф. Липницкий // Информатика. – 2020. – Т. 17, № 4. – С. 73–82. <https://doi.org/10.37661/1816-0301-2020-17-4-73-82>
8. Мартинович, Г. А. Вербальные ассоциации и организация лексикона человека / Г. А. Мартинович // Филологические науки. – 1989. – № 3. – С. 39–45.
9. Темнова, Е. В. Современные подходы к изучению дискурса / Е. В. Темнова // Язык, сознание, коммуникация : сб. ст. – М. : МАКС Пресс, 2004. – Вып. 26. – С. 24–32.

References

1. Kravchenko T. K. *Decision support systems*. Informacionnye tehnologii dlja sovremennogo universiteta [*Information Technologies for a Modern University*]. In Tikhonov A. N., Ivannikov A. D. (eds.). Moscow, Gosudarstvennyj nauchno-issledovatel'skij institut informacionnyh tehnologij i telekommunikacij "Informika", 2011, pp. 107–118 (In Russ.).
2. Moiseenko E. V., Lavrushina E. G. Informacionnye tehnologii v jekonomike. *Information Technologies in the Economy*. In M. A. Kasatkina (ed.). Moscow, Soft, 2009, pp. 120–135 (In Russ.).
3. Simankov V. S., Cherkasov A. N. *Methodological support of decision support stages in the synthesis of complex systems*. Perspektivy nauki [*Prospects of Science*], 2012, no. 12, pp. 85–89 (In Russ.).
4. Lipnitsky S. F. *Model of knowledge representation in information systems based on verbal associations*. Informatika [*Informatics*], 2011, no. 4 (32), pp. 21–28 (In Russ.).
5. Gasparov B. M. Jazyk, pamjat', obraz. Lingvistika jazykovogo sushhestvovanija. *Language, Memory, Image. Linguistics of Linguistic Existence*. Moscow, Novoe literaturnoe obozrenie, 1996, 352 p. (In Russ.).
6. Lipnitskiy S. F. *Synthesis of queries and search for alternatives in the system of information support for decision-making*. Problemy fiziki, matematiki i tehniki [*Problems of Physics, Mathematics and Technology*], 2020, no. 2, pp. 91–95 (In Russ.).
7. Lipnitsky S. F. *Internet search and lexical-semantic processing of analogs when making decisions in various subject areas*. *Informatics*, 2020, vol. 17, no. 4, pp. 73–82 (In Russ.). <https://doi.org/10.37661/1816-0301-2020-17-4-73-82>
8. Martinovich G. A. *Verbal associations and the organization of the human lexicon*. Filologicheskie nauki [*Philological Sciences*], 1989, no. 3, pp. 39–45 (In Russ.).
9. Temnova E. V. *Modern approaches to the study of discourse*. Jazyk, soznanie, kommunikacija : sbornik statej [*Language, Consciousness, Communication: Collection of Articles*]. Moscow, MAKS Press, 2004, iss. 26, pp. 24–32 (In Russ.).

Информация об авторе

Липницкий Станислав Феликсович, доктор технических наук, главный научный сотрудник, Объединенный институт проблем информатики Национальной академии наук Беларуси.
E-mail: lipn@newman.bas-net.by

Information about the author

Stanislav F. Lipnitsky, Dr. Sci. (Eng.), Chief Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.
E-mail: lipn@newman.bas-net.by