

**Психологический подход
к обработке текстовой информации**
Psychological approach towards text information processing

А. С. Баканов

Институт психологии РАН,

Москва, Россия

Arseny Bakanov

Institute of Psychology of the Russian Academy of Sciences,

Moscow, Russia

В настоящей статье описывается способ обработки текстовой информации в процессе человеко-компьютерного взаимодействия с информационными системами. Обработка осуществляется с целью структуризации и графического представления текстовой информации. В работе приведено описание проведенных экспериментов, приводятся полученные результаты экспериментов. В описываемом подходе используются как психологические методики, так и контент-анализ текста с использованием специально разработанного словаря. Описаны оригинальные способы представления структуры текста и визуализации информации.

Ключевые слова: Обработка текстовой информации, информационные системы, принятие решений.

The author examines the method of text information processing through human-computer interaction with information systems. Text information is processed to be structurized and presented graphically. The conducted experiments and their findings are described. This approach is to apply psychological methods along with text content analysis using dedicated glossary. Unique methods of text structure presentation and information visualization are reviewed.

Keywords: Text information processing, information system, decision making.

Введение

Статья посвящена исследованию проблем, связанных с взаимодействием человека с информационными системами, в частности вопросам обработки человеком текстовой информации, в процессе человеко-компьютерного взаимодействия. В настоящей статье описывается психологический подход к обработке текстовой информации. Также в данной статье предлагается оригинальный способ представления структуры текста и визуализации информации.

В статье, кратко представлены результаты исследований, проводимых с 2012 года по настоящее время в Институте психологии РАН. Для исследования способов обработки человеком текстовой информации, были проведены исследования двух типов: в исследованиях первого типа первоначально предъявлялся стимульный материал (в виде вопросов, ключевых слов, графических образов), а затем испытуемому предъявлялась текстовая информация, после прочтения которой, испытуемый принимал решение о сортировке (т. е. определял департамент или подразделение организации, куда документ должен быть направлен для исполнения), отвечал на вопросы по тексту и проходил психологическое тестирование; в исследованиях второго типа, эксперимент начинался с предъявления текстовой информации, после прочтения которой, испытуемый принимал решение о сортировке, отвечал на вопросы по тексту и проходил психологическое тестирование.

Испытуемым предъявлялись специализированные тексты определенной тематики, как прошедшие предварительную обработку – т.е. с выделенными опорными словами, так и тексты без предварительной обработки.

В ходе исследований, использовалась установка контроля движения взгляда испытуемого (www.smivision.com). С ее помощью были получены экспериментальные данные: траектория взгляда, диаметр зрачка испытуемого, скорость перемещения взгляда, как в процессе чтения текста, так и в процессе принятия решения.

Взаимодействие человека с информационными системами

Взаимодействие человека с техническими системами вообще и с информационными системами в частности, изучается различными научными направлениями: информационные технологии, инженерная психология и т.д. Это обусловлено актуальностью, а также широким и повсеместным использованием информационных систем [Шрайберг, 2013], [Баканова, 2007], [Ташев и др., 2013]. Вопросам, связанным с проектированием информационных систем, а также проблемам взаимодействия человека с информационными системами посвящено значительное количество работ [Баканов, 2009]. Среди исследований, касающихся проектирования информационных систем необходимо выделить работу Вишневого В.М. [Вишневецкий, 2006], а также необходимо отметить работу коллектива авторов: Т. В. Атанасовой, Т.Н. Савченко, Г.М. Головиной и др., посвященную вопросам информационного взаимодействия человека с информационными системами, в которой описаны психологические механизмы взаимодействия человека с интеллектуальной информационной средой обитания [Атанасова Т.В. и др., 2010].

В процессе взаимодействия с информационной средой, к человеку по различным информационным каналам (с использованием различных телекоммуникационных устройств) могут поступать значительные объемы информации для переработки и последующего принятия решения. Количество каналов информации, по которым информация поступает к человеку, продолжает стремительно увеличиваться, как и количество поступающей информации. В процессе взаимодействия, человеку приходится обрабатывать массивы текстовой и графической информации, учитывать значительное количество различных факторов, а также решать задачи многокритериального выбора. Для человеческой системы переработки информации многокритериальные задачи представляют собой особо сложный класс задач [Петровский А.Б., 2004]. Наличие многих критериев приводит к нагрузке на человеческую систему переработки информации, заставляя человека использовать различные, зачастую оригинальные эвристики для того, чтобы решить поставленную задачу [Ларичев О. И., 1987].

Описание проведенных экспериментов

В соответствии с целью работы было проведено экспериментальное исследование взаимодействия человека со специализированной информационной системой. В ходе исследования моделировался процесс чтения текстовой информации. Испытуемому необходимо было прочитать документ – специализированный текст, предъявленный на мониторе компьютера, и затем ответить на предложенные вопросы – принять решение о сортировке (decision making), т. е. определить департамент или подразделение организации, к которому по тематике относится прочитанный текст.

Необходимо отметить, что в состав информационной системы был включен программный модуль, реализующий функции системы поддержки принятия решений. Данный модуль осуществлял предварительную обработку текста документа, выделял опорные слова цветом и визуализировал структуру документа. В рамках проводимых экспериментов представляло научный интерес исследовать процесс обработки текстовой информации человеком в процессе чтения специализированного документа, как «экспертами», т. е. людьми, имеющими опыт работы по обработке и сортировке документов, так и «новичками», т. е. людьми, не имеющими такого опыта работы. В процессе исследования, решения принимались как с помощью системы поддержки принятия решений, так и без помощи системы поддержки принятия решений. В качестве испытуемых в исследованиях участвовали:

- 1) эксперты – лица имеющие опыт работы в качестве ЛПР;
- 2) новички – лица, не имеющие опыта работы в качестве ЛПР – студенты высших учебных заведений г. Москвы. Исследования проводились с использованием оборудования отслеживающего траекторию движения глаз пользователя в процессе считывания информации, а также в процессе принятия решений (www.smivision.com). Параллельно с проведением экспериментов, проводились исследования когнитивного стиля импульсивный/рефлексивный [Холодная, 2002], а также стиля саморегуляции поведения (ССПМ) по методике В. И. Моросановой [Моросанова, Индина, 2011]. Испытуемым предъявлялся некоторый документ содержащий текст, после прочтения, которого необходимо было принять решение о выборе из

некоторого числа альтернатив (принять решение о сортировке документа). Испытуемым последовательно предъявлялись тексты как прошедшие предварительную обработку системой поддержки принятия решений, так и без обработки. Обработка заключалась в определении структуры текста и проведении контент-анализа с использованием специально разработанного словаря (тезауруса). В результате обработки визуализировалась структура текста, а опорные слова выделялись цветом [Баканов А.С., 2009]. Последовательность предъявления текстов (с обработкой и без обработки) на мониторе компьютера, менялась, чтобы нивелировать привыкание испытуемого, после прочтения текста испытуемый принимал решение. Экспериментатор наблюдал и фиксировал (в том числе и в формате *.mpeg) траекторию движения глаз испытуемого, как в процессе чтения текста, так и в процессе принятия решения. На этапе принятия решения испытуемым предъявлялись (также на мониторе компьютера) вопросы двух типов. К вопросам первого типа относились вопросы, предъявляемые в следующем порядке:

1) Выберите альтернативы (из некоторого списка) к которым относится данный документ/текст.

2) Выберите альтернативы (из числа выбранных ранее) к которым данный документ/текст относится в наибольшей степени.

3) Выберите только одну альтернативу (из выбранных ранее).

Таким образом, в процессе ответов на вопросы первого типа в приведенном выше порядке выявлялась структура ментальных репрезентаций пользователя, сформированная по прочтении конкретного документа/текста см. рисунок 1.

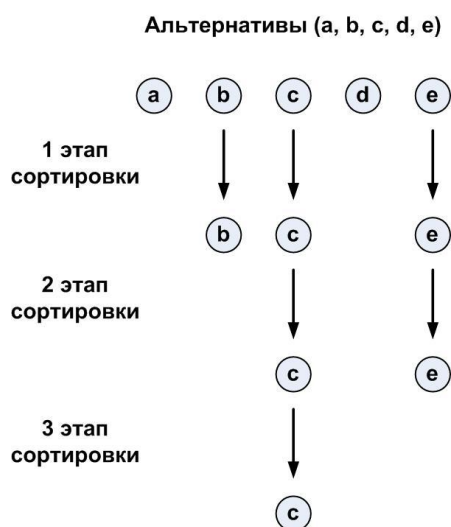


Рис. 1

Необходимо отметить, что существует достаточное количество определений термина ментальная репрезентация, мы же будем использовать следующее определение: ментальная репрезентация – субъективный образ объективной реальности, отражение внутреннего и внешнего мира в сознании человека. Или применительно к данному исследованию субъективный образ прочтенного документа/текста в сознании испытуемого. При ответе на вопрос (также предъявляемый на экране монитора) второго типа испытуемый должен был количественно оценить степень, в которой данный документ/текст относится к выбранной альтернативе. Траектория движения взгляда, скорость движения взгляда, а также диаметр зрачка испытуемого в процессе принятия решения фиксировалась экспериментатором.

Кроме траектории взгляда, экспериментатор в процессе проводимого исследования фиксировал время в течении которого испытуемый читал предъявленный текст, а также время в течении которого испытуемый принимал решение после прочтения текста как с выделенными цветом (опорными) словами, так и без выделенных цветом опорных слов. В настоящее время

исследования продолжаются. К настоящему времени в исследованиях приняло участие более 70 человек.

В ходе проведения экспериментов удалось выявить, что большинство испытуемых читали предлагаемый текст дважды. Причем во второй раз взор испытуемого перемещался от одного опорного слова (фрагмента текста) к другому опорному слову (фрагменту текста) в независимости были ли они выделены в процессе обработки текста или нет. В процессе движения от одного опорного слова к другому, взор испытуемого «перескакивал» через строки и абзацы, иногда возвращаясь к отдельным фрагментам текста. Можно предположить, что таким образом испытуемый пытался запомнить наиболее важные (для него) слова или фрагменты текста. Такие последовательности слов можно представить в виде:

$$A(X_1, X_2, X_3, \dots, X_n)$$

$$B(Y_1, Y_2, Y_3, \dots, Y_m)$$

И т.д.

Где $X_1, X_2, X_3, \dots, X_n, Y_1, Y_2, Y_3, \dots, Y_m$ опорные слова (невыделенные либо выделенные в процессе обработки) и фрагменты текста;

A, B – некоторые предикаты.

Последовательность опорных слов также можно представить в виде:

$$A_1(X_1, X_2), A_2(X_2, X_3), \dots, A_{n-1}(X_{n-1}, X_n)$$

$$B_1(X_1, X_2), B_2(X_2, X_3), \dots, B_{n-1}(X_{n-1}, X_n)$$

И т.д.

Представляя предикатное описание графически, получаем последовательность опорных слов в виде семантической сети рис.2.

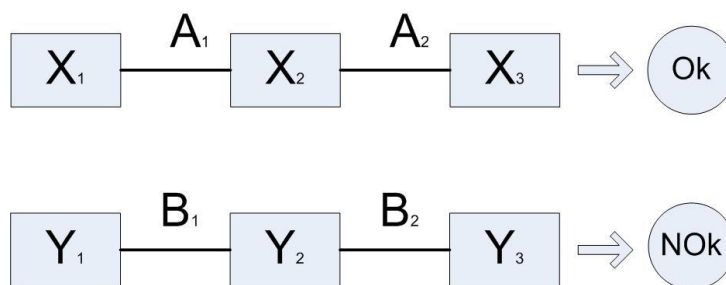


Рис. 2

Символами «Ok» и «NOK» на рис. 2 показано решение эксперта задачи о сортировке, т.е. о том, что данный документ/текст относится к выбранной альтернативе.

Краткие выводы

Обобщая все вышеизложенное можно сделать следующие выводы:

1) Представляя отношения дугами графа, мы визуализируем алгоритм рассуждения эксперта о предметной области.

2) Двигаясь по дугам графа, мы приходим к результату – решению эксперта в задаче о сортировке.

3) Расширяя множество $X_1, X_2, X_3, \dots, X_n, Y_1, Y_2, Y_3, \dots, Y_m$ опорных слов с помощью специализированного словаря мы можем получить совокупность продукционных правил «ЕСЛИ-ТО».

Библиография

1. [Атанасова и др., 2010] Атанасова Т., Савченко Т.Н., Головина Г.М., Баканов А.С. Интеллектуальная информационная среда обитания и субъективное восприятие качества жизни // Методы исследования психологических структур и их динамики. Труды ИП РАН. М., 2010.
2. [Баканов, 2009] Баканов А.С. Особенности психологического подхода к моделированию человеко-компьютерного взаимодействия // Вестник ГУУ. 2009. №6. С. 15–18.
3. [Баканова, 2007] Баканова Н. Б. Использование программно-технических комплексов для повышения эффективности контроля в системах документооборота // «Электросвязь». 2007. № 6. С. 51–53.
4. [Вишневский, 2006] . Вишневский В.М. Теоретические основы проектирования компьютерных сетей // – М. Техносфера 2003 с. 512 ISBN: 5-94836-011-3
5. [Ларичев, 1987] Ларичев О.И., Петровский А.Б. Системы поддержки принятия решений. Современное состояние и перспективы развития. // Итоги науки и техники. Серия Техническая кибернетика. М. ВИНТИ, 1987. т.21, с.131–164.
6. [Моросанова, Индина, 2011] Моросанова В.И., Индина Т.А. Регуляторные и личностные основы принятия решений. – СПб.; М.: Нестор-История, 2011, – 282 с.
7. [Петровский, 2004] Петровский А.Б. Многокритериальное принятие решений по противоречивыми данным: подход теории мультимножеств. // Информационные технологии и вычислительные системы, 2004, №2, 56–66.
8. [Холодная, 2002] Холодная М. А. Когнитивные стили: О природе индивидуального ума. Учебное пособие – М.: ПЕР СЭ, 2002. – 304 с.
9. [Ташев и др., 2013] Ташев Т., Баканова Н., Ташева Р. Исследование верхней границы пропускной способности коммутационного узла при входящем трафике типа «горячей точки». International Journal «Information Technologies & Knowledge», Vol.7, No. 2, 2013, pp. 182–189.
10. [Шрайберг, 2013] Шрайберг Я. Л. Доступ к библиотечно-информационным ресурсам сферы образования и науки: первые результаты Федерального проекта Министерства образования и науки РФ/ Я. Л. Шрайберг //Формирование и использование электронных ресурсов библиотек России: материалы ежегод. совещ. руководителей федер. и центр. регион. б-к России (Москва, 23–24 окт. 2012 г.). – М. : Пашков дом, 2013. – С. 88–93. ISBN 978-5-7510-0579-6.