

# АВТОМАТИЗИРОВАННЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ В ОБРАЗОВАНИИ И НАУКЕ

## КОГНИТИВНЫЕ КАРТЫ ДИАГНОСТИКИ ЗНАНИЙ

В.А. Углев

Филиал ФГАОУ ВПО «Сибирский федеральный университет», г. Железногорск

Рассматривается подход к решению задачи эффективного управления индивидуальной траекторией обучения в автоматизированных обучающих системах. Для этих целей в обучающую систему вводятся когнитивные карты диагностики знаний.

**Ключевые слова:** автоматизированные обучающие системы, когнитивные карты диагностики знаний, индивидуализация траектории обучения.

## THE COGNITIVE MAPS OF KNOWLEDGE DIAGNOSIS

V.A. Uglev

The branch of Siberian Federal University, Zheleznogorsk

The article considers a method of solving the problem of effectiveness manage the individual trajectory of training in the automated educational systems. For this goal the cognitive maps of knowledge diagnosis were input into training system.

**Keywords:** automated educational systems, cognitive maps of knowledge diagnosis, training trajectory individualization.

### 1. Введение

Организация процесса дистанционного или самостоятельного обучения с помощью средств вычислительной техники до сих пор является сложной задачей [1, 2], не имеющей универсальных решений. При обработке информации о человеко-машинном взаимодействии данная проблема стоит особенно остро, так как имеет место ограниченный набор факторов, регистрируемых компьютером о пользователе (обучающемся). Это обстоятельство приводит к существенному повышению роли методов извлечения знаний [3] из протоколов работы автоматизированных систем, реализующих индивидуальную траекторию обучения. Рассмотрим один из подходов к анализу наборов статистических данных и извлечению из них знаний о пользователе на примере работы автоматизированных обучающих систем (АОС) нового поколения [4].

Проблема построения гипотез о процессе обучения при взаимодействии типа «человек–компьютер» имеет ряд сложностей, связанных с низкой формализацией сути учебного материала

(последовательность текстов, графиков и таблиц), с низкой эффективностью процесса автоматизированного контроля (применение итогового тестирования) и с ограниченными источниками данных об уровне обученности пользователя<sup>1</sup> [5–7]. Проблема требует системного взгляда, учитывающего наиболее важные факторы поведения обучаемого [8]. В таком случае источниками знаний о динамике процесса обучения являются частотно-временные показатели работы человека с различными подсистемами АОС, а также данные о результатах работы с контролирующим материалом (решение задач в электронной тетради, прохождение тестов). Результирующая информация по каждому пользователю представляется в виде временного ряда  $\{X\}$ , каждый элемент которого описывается кортежем типа:

$$x_i = \langle A', B, C, D' \mid A, D, E \rangle, \quad (1)$$

где  $i$  – порядковый номер элемента в привязке к временной шкале;  $A'$  – подмножество целей / компетенций, которые пользователь выбрал для развития при работе с АОС;  $B$  – статистика работы с учебным материалом (включая историю траек-

<sup>1</sup> Фактически АОС «видит» не человека, а последовательность перемещений мыши по экрану и сочетания нажатий клавиш с клавиатуры. Применение Web-камер, регистрирующих перчаток и иной экзотики, – приятные допущения, не отражающие реалий типовой ситуации дистанционного и самостоятельного обучения..

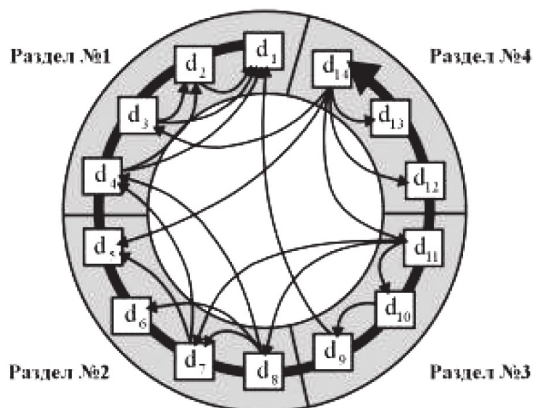


Рис. 1. Пример представления структуры учебного курса в виде семантической сети

тории перемещений между элементами дидактического материала);  $C$  – статистика прохождения контрольно-измерительного материала;  $D'$  – индивидуализированный состав учебного курса (включая контролирующий материал);  $A$  – набор целей / компетенций, которые должны быть развиты в процессе обучения;  $D$  – фактический состав учебного курса;  $E$  – метаинформация об электронном учебном курсе.

Анализируя структуру  $\{X\}$  на уровне матричного представления (таблицы базы данных), можно сделать ряд простейших выводов, касающихся динамики знаний в контрольных точках электронного курса. Но для эффективной работы АОС, формирующей адекватную индивидуальную траекторию обучения, этих данных недостаточно [9]. Требуется извлечь знания, определяющие дальнейшую стратегию поведения обучающей системы. Для этого требуется расширить описание модели учебного курса, дополнив её слоем метаинформации.

## 2. Модель курса и метазнания в АОС

Традиционная модель учебного курса включает в себя иерархию учебного и контролирующего материала (структуру курса  $D$ ) и метаданные о последовательности предъявления дидактиче-

ских единиц ( $E$ ). Если говорить об обучающих автоматизированных системах нового поколения, то в них оба эти компонента должны быть модифицированы. В фактический состав рабочего курса (его реализацию  $D'$ ) должны входить не все элементы  $D$ , а лишь та часть, которая позволяет сформировать адекватное подмножество учебного материала по (2), полностью включающего указанные пользователем цели обучения<sup>2</sup> и дополнения, без которых последовательное изучение материала будет невозможно (ядро курса  $D_{\text{kernel}}$ ):

$$D' = (D \subseteq D_{A'}) \cup D_{\text{kernel}}, \quad (2)$$

где  $D_{A'}$  – подмножество  $D$ , включающее те элементы, которые через связи с целями / компетенциями связаны с  $A'$ . Тогда  $E$  требует наличия не только описания стандартной траектории перемещения между элементами  $D$ , но и семантической сети [10], регламентирующей формирование траектории между элементами из  $D'$  и правила синтеза индивидуальной траектории обучения. Следовательно, часть знаний о семантике учебного материала необходимо формально размещать в составе модели учебного курса.

Наиболее компактной и наглядной формой представления структуры семантической сети удаётся достичь, если применять круговую схему представления структуры учебного материала (рис. 1). Здесь узлами сети являются дидактические единицы (включая характеристики важности, сложности, типа представления, развиваемые компетенции и пр.), а рёбрами – связи логических зависимостей типа «базовый – дочерний»<sup>3</sup>.

Из рисунка видно, что элементы  $E$  и  $D'$  определяют статическую часть метаинформации об учебном курсе в АОС, т.е. отвечают на вопросы: «что должно быть изучено?» и «в какой последовательности нужно изначально знакомиться с дидактическими единицами?». Из семантической сети можно выделить то, какие контрольно-измерительные материалы (задачи, упражнения, тестовые задания) будут включены в  $D'$ , т.е. получить эталонные показатели обучения.

<sup>2</sup> Состав подмножества  $A'$  определяется как минимальный объём элементов из  $A$ , отвечающий условию соответствия элементов подмножества пожеланиям пользователя при регистрации на курс в АОС: цели обучения указываются в виде набора нужных разделов курса, уточнения профиля обучения и/или множества развиваемых компетенций.

<sup>3</sup> Стрелка, идущая вдоль границы окружности, определяет нормативную последовательность изучения материала, а те связи, которые находятся внутри контура, описывают семантические зависимости элементов учебного курса.

Предложенная модель курса, которая позволяет получить ответы на вопросы, «каким должен быть индивидуализированный состав курса?» и «к чему должны стремиться показатели уровня обученности?», не даёт метазнаний, позволяющих индивидуализировать сам процесс обучения<sup>4</sup>. Необходимо учитывать динамику обучения, осуществляя ситуационное управление [11]. В автоматизированных обучающих системах за это традиционно отвечает механизм экспертных систем, которому на вход подаются оперативные данные о работе пользователя. Для осуществления управления работой АОС необходимо анализировать динамику учебного процесса, выявляя значимые показатели, а не ограничиваться констатацией уровня освоения отдельных дидактических единиц в баллах теста, как это делается в подавляющем большинстве автоматизированных обучающих систем.

Выделим показатели, которые *значимы* для управления процессом обучения:

- корреляция между фактически повторяемым материалом, рекомендуемым к изучению автоматизированной системой, и оценками во времени (ответ на вопрос «воспринимает ли пользователь рекомендации АОС?»);
- эффективность процесса самообучения (скорость, динамика успехов) при использовании АОС (ответ на вопрос «успешно ли продвигается обучение?»);
- эффективность трудозатрат (время, частота обращения) при работе с подсистемой электронного учебного курса и обучающего компьютерного тестирования (ответ на вопрос «достаточно ли пользователь работает с АОС?»);
- актуальность направления работы (выявление перечня разделов/материалов курса, которые необходимо повторить, временно отложить или считать освоенными) с учётом динамики успехов и целей обучения (ответ на вопрос «что пользователю наиболее актуально изучать в первую очередь при текущем уровне знаний?»).

Для реализации механизмов индивидуализации, оперирующих значимой информацией о процессе обучения, требуется обрабатывать статистику из  $B$  и  $C$  из (1). Для этого необходимо последовательно решить следующие задачи:

- извлечь из базы данных всю информацию, относящуюся к отдельным характеристикам процесса обучения;
- оценить динамику каждого показателя (осуществить предобработку);
- сгруппировать оценки показателей;
- направить совокупность показателей в подсистему принятия решений (в АОС это обычно экспертная система).

Таким образом, набор метазнаний о динамической части  $x_i$  требует комплексных мер по подготовке, прежде чем он будет подан на вход экспертной системы. Фактически требуется переработать весь массив хранящейся информации о динамике обучения конкретного пользователя и осуществить перевод информации в знания, т.е. перейти от количества к качеству. Очевидно, что за один шаг этого не добиться. Поэтому выделим задачи сбора, предобработки и группировки значимой информации о процессе обучения в отдельную процедуру, которая носит название *когнитивной карты диагностики знаний*.

### 3. Когнитивные карты диагностики знаний

Когнитивная карта диагностики знаний (ККДЗ), как следует из определения в [12], является сервисом АОС, автоматически подготавливающим (предобрабатывающим) информацию о процессе обучения (выявление пробелов в знаниях, оценка и отображение динамики обучения) с целью упростить комплексный экспертный анализ учебной ситуации и выработать адекватную реакцию системы на действия пользователя. Таким образом, ККДЗ формируется каждый раз, когда требуется реализовать обратную связь между АОС и учащимся: согласно [4, 13] это этапы оценивания знаний, адаптации состава обучающего теста, формирования траектории

<sup>4</sup>Здесь следует развести понятия «метаинформация» и «метазнания». Метаинформация – это такая информация, которая описывает служебные параметры структуры и семантических связей в типовом (не индивидуализированном) электронном курсе, а также однозначно продекларированные параметры модели пользователя (изначально указанные обучаемым цели и предпочтения). Метазнаниями является такая информация, которая выводится из имеющихся данных (включая метаинформацию) и ситуационно помогает выработать эффективную обратную (т.е. управляющую стратегией подачи материала в АОС) связь, используя методы искусственного интеллекта.

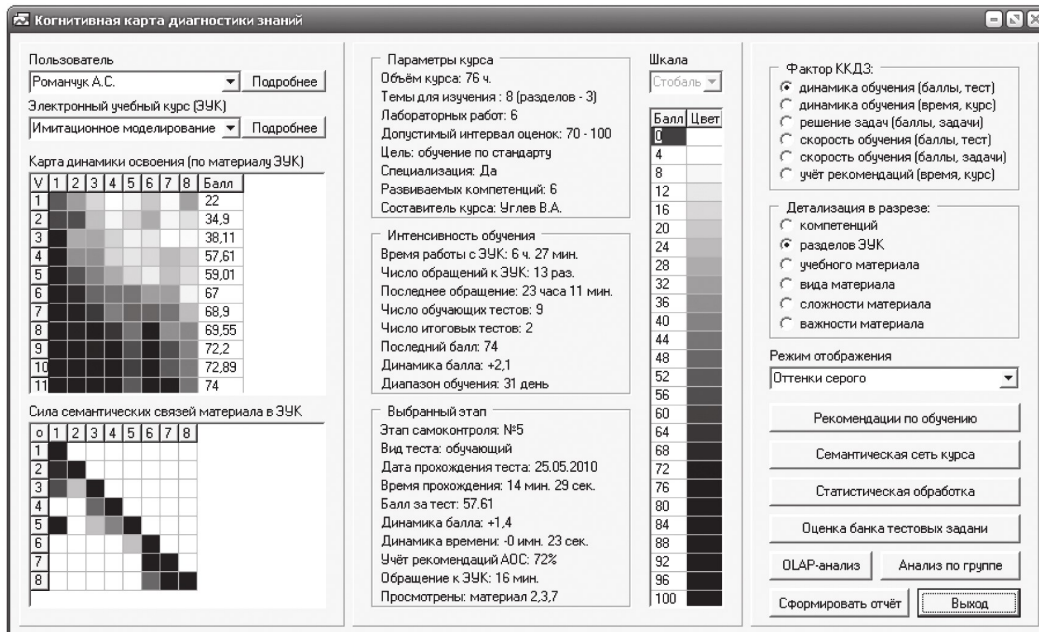


Рис. 2. Пример фрагмента ККДЗ для дисциплины «Имитационное моделирование»

обучения, синтеза подсказок или объяснения оценок в естественно-языковой форме. Фактически сервис формирования ККДЗ должен активизироваться перед вызовом отдельных экспертных систем, соответствующих перечисленным этапам, обрабатывая данные из  $x_i$ , хранящиеся в базе данных АОС.

Как было показано выше, метаинформация о процессе обучения включает статическую и динамическую части, поэтому и ККДЗ должна формироваться в памяти ЭВМ в виде комплекса показателей, которые будут отражать особенность учебного процесса. Учитывая ограниченность источников информации и различную степень достоверности отдельных показателей (например, при оценивании компьютерного теста оценивается вероятность угадываний верных ответов), в карте нужно по возможности обобщенно и компактно отразить картину обучения. При этом ККДЗ должна демонстрировать не только комплексные характеристики процесса обучения, но и позволять быстро выделять существенные тенденции, которые далее пойдут на вход продукционной экспертной системе. Это достигается за счёт сочетания методов когнитивной компьютерной графики, Data Mining, статистического анализа временных рядов и ментальных карт. Пример вывода на форму фрагмента простой

когнитивной карты диагностики знаний показан на рис. 2.

При синтезе ККДЗ, представленной на рис. 2, был использован целый комплекс алгоритмов, отражающих различные классы характеристик учебной ситуации. Во-первых, это набор обычных запросов, возвращающих статические параметры об электронном курсе (область «Параметры курса») и статистические показатели процесса обучения (область «Интенсивность обучения» и «Выбранный этап»). Во-вторых, это семантическая сеть электронного курса, отражаемая в левом нижнем углу формы. Определение силы влияния семантических связей на результат обучения выражается в виде комбинации тройки показателей:

- фактическая оценка дидактической единицы  $d_n$ ;
- влияние базовых зависимостей внутри структуры учебного материала  $\underline{Q}_n^*$ ;
- влияние дочерних зависимостей внутри структуры учебного материала  $\bar{Q}_n^*$ .

Значение влияния взаимозависимостей определяется по формулам (3)

$$\underline{Q}_n^* = \frac{\sum_{i \in M1} \underline{Q}_{in}}{\|M1\|} + \frac{\sum_{j \in M2} \underline{Q}_{jn}}{2\|M2\|} \quad \text{и} \quad \bar{Q}_n^* = \frac{\sum_{k \in M3} \bar{Q}_{kn}}{\|M3\|} + \frac{\sum_{l \in M4} \bar{Q}_{ln}}{2\|M4\|}, \quad (3)$$

где  $\underline{Q}_i, \underline{Q}_j, \overline{Q}_k, \overline{Q}_l$  – оценки связанных дидактических единиц из группы дочерних внешних, дочерних внутренних, родительских внутренних и родительских внешних соответственно; множества  $M1, M2, M3, M4$  содержат счётчики  $i, j, k, l$  указанных групп соответственно<sup>5</sup>.

В-третьих, это карта динамики проявления соответствующей размерности (фактора) ККДЗ. На рисунке ККДЗ представлена картой фактора «динамика обучения (баллы, тест)», где цвета обозначают процент освоения учебного материала в градации выбранной шкалы (по данным анализа результатов решения теста), столбцы – порядковые номера дидактических единиц, а строки – отдельные обращения к подсистеме тестирования. Эту область можно уподобить карте высот с той лишь разницей, что она строится относительно времени и позволяет наглядно увидеть динамику изменения учебных достижений.

В-четвёртых, в карту включается набор расчетных параметров, требующих целой последовательности действий. Примером такого параметра будет «Учёт рекомендаций АОС» ( $\delta$ ): для этапа контроля  $i$  сравнивается множество рекомендуемых к повторению элементов из  $D'$  после  $i-1$  шага с фактическим изучаемыми по формуле (4):

$$\delta = \left( \frac{\|D'' \cap D^*\|}{\|D^*\|} - \frac{|S|}{4 \cdot \|D''\|} \right) \cdot 100, \quad (4)$$

где  $S = \|(D^* \cup D'') \setminus (D^* \cap D'')\|$  – мощность симметрической разности;  $D^*$  – подмножество  $D'$ , рекомендованное к повторной проработке, а  $D''$  – фактически проработанное подмножество. Очевидно, что при смене размерностей отображения данных (списки «Фактор ККДЗ» и «Детализация в разрезе») будет изменена левая часть оконной формы, используя соответствующие алгоритмы расчёта показателей.

#### 4. Управление траекторией обучения по ККДЗ

Траектория обучения пользователя, как объект управления<sup>6</sup> является многопараметрической структурой. Это свидетельствует не только о необходимости комплексного обобщения множества промежуточных оценок работы обучаемого из ККДЗ, но и принятия решений на уровне модели знаний. В самой инженерии знаний существует целый ряд моделей, позволяющих представить в памяти компьютера структуру знаний. В данном подходе рассматриваются семантические сети и онтологии как гибкий инструмент описания логической взаимосвязи отдельных актов (концептов) и возможных путей принятия решений (аксиомы и продукционные правила) [3, 15, 16]. Следовательно, данные из ККДЗ должны быть далее обобщены и проанализированы с помощью экспертной системы, входящей в состав развитых АОС [17]. Здесь следует отметить ряд важных моментов:

Во-первых, большинство из показателей, представленных в ККДЗ, носят количественную природу и для логического анализа не приспособлены. Для преодоления этой трудности используется механизм нечёткой логики, осуществляющий фазификацию, т.е. перевод числовых данных в качественные значения лингвистической переменной [18].

Во-вторых, требуется осуществлять расчёт оценки владения не только отдельной дидактической единицы, но и оценки каждой компетенции из  $A'$ . Очевидно, что множества  $A'$  и  $D'$  будут составлять двудольный граф, определённый в  $E$ . Следовательно, тестовые задания, которыми обладает каждый элемент из  $D'$ , будут оценивать компетенции, относящиеся к конкретной дидактической единице. Возникает задача разработки метода оценки степени овладения компетенциями по результатам комплексной проверки (например, прохождения обучающего теста). В качестве

<sup>5</sup> В формулах (3) первое слагаемое отражает силу связей с соседними дидактическими единицами внутри анализируемого раздела, а второе слагаемое – с учебным материалом из других разделов (детально данный алгоритм описан в [14]).

<sup>6</sup> В литературе по данной тематике принято считать, что управление траекторией фактически эквивалентно управлению самим процессом обучения [21]. Принимая во внимание сложность решаемой задачи, это не всегда так. Поэтому первичные данные о траектории перемещений между элементами АОС – лишь исходные данные для преобработки (в данной работе с помощью ККДЗ) и дальнейшего принятия решений.

решения, позволяющего удовлетворительно сработать с множественными зависимостями объектов  $A' \leftrightarrow D'$ , можно воспользоваться коэффициентами уверенности Шортлиффа и Бучанана [19, 20]. Основная идея этого метода преодоления многозначности гипотез заключается в том, чтобы данные о состоянии показателя собирались итерационно (с каждым тестовым заданием) и сопровождалась априорной оценкой, подтверждающей одни состояния ( $MB$ ) гипотезы и опровергающей другие ( $MD$ ). Связь между мерой доверия и недоверия устанавливается с помощью коэффициента уверенности по формуле (5):

$$Cf(a_j, x_a \wedge x_{\bar{a}}) = MB(a_j, x_a) - MD(a_j, x_{\bar{a}}), \quad (5)$$

где  $a_j \in A'$ , а  $x_a$  и  $x_{\bar{a}}$  есть свидетельства в пользу и не в пользу гипотезы (усвоения компетенции) в результате прохождения обучающего теста. Для комбинации свидетельств в сложных гипотезах (ответ на задание демонстрирует различный исход для разных компетенций) используются следующие приближенные методы оценки по (6):

$$MB[H, x] = \frac{p(H/x) - p(H)}{1 - p(H)} \quad \text{и} \\ MD[H, x] = \frac{p(H) - p(H/x)}{p(H)}, \quad (6)$$

где  $x$  – количественная мера одного из факторов (тестовых заданий) для каждой компетенции из  $A$ . Так, в результате решения теста будет получен ряд коэффициентов уверенности, соответствующих степени проявления каждой компетенции из  $A'$  в результатах очередного этапа контроля.

Таким образом, в экспертную систему, вырабатывающую стратегию управления траекторией обучения, должны быть поданы агрегированные показатели, отражающие текущее состояние вспомогательных гипотез экспертной системы по таким вопросам, как:

- Имеет ли результат процесс самостоятельного обучения?
- Принимает ли во внимание учащийся рекомендации АОС?
- Насколько рационально действует пользователь при повторном обращении к учебному материалу?
- Какой фрагмент учебного материала наиболее актуален для получения наилучшей динамики показателей обученности?
- Насколько случайными являются показатели, зарегистрированные подсистемой контроля?

- Какую последовательность учебного материала требуется повторно проработать для повышения показателей для конкретной дидактической единицы / компетенции?

Реализация всех этих возможностей базируется на ККДЗ – инструменте извлечения, обобщения и компактного представления информации, важной при принятии решений в АОС. За счёт карты знаний существенно упрощается подача информации в экспертную систему, оперирующую не только ситуационной информацией о процессе обучения, но и знаниями экспертов. Так АОС становится прообразом индивидуального ситуационного центра [4, 22], оперативно реагирующего на изменения в массивах статистических данных о работе пользователя.

## 5. Заключение

Предложенный в статье формальный подход к выявлению и организации значимой информации о процессе обучения из статистических данных с помощью когнитивных карт диагностики знаний проходит апробацию в сочетании с методом обучающего тестирования в рамках единой автоматизированной обучающей системы. Интеграция сервиса когнитивных карт диагностики знаний в состав обучающей системы позволяет не только повысить адекватность индивидуализированного управления траекторией обучения, но и выработать правила синтеза аргументированных ответов системы при ведении диалога в естественно-языковой форме. Это способствует переходу автоматизированных обучающих систем к новому, более совершенному этапу развития, оперирующему не только структурированными наборами учебной информации, но и знаниями о процессе освоения дидактических единиц для решения задачи индивидуализированного управления траекторией обучения.

## ЛИТЕРАТУРА

1. Аванесов В.С. Методические и теоретические основы тестового педагогического контроля: дис. ... д-ра пед. наук. – М., 1994. – 326 с.
2. Беспалько В. Образование и обучение с участием компьютера (педагогика третьего тысячелетия). – М.: МОДЭК, 2002. – 352 с.
3. Гаврилова Т.А., Муромцев Д.И. Интеллектуальные технологии в менеджменте. – СПб.: Высшая школа, 2008. – 488 с.
4. Углев В.А. Автоматизированные обучающие системы и компьютерное тестирование: системный кризис и перспективы развития // Модернизация российского образования: тренды и перспективы. – Ч. 2. – Краснодар: АНО «Центр

социально-политических исследований «Премьер», 2011. – С. 3–18.

5. Морев И.А. Образовательные информационные технологии. – Ч. 2: Педагогические измерения: учеб. пособие. – Владивосток: Изд-во Дальневост. ун-та, 2004. – 174 с.

6. Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов. – М.: Прометей, 2000. – 168 с.

7. Углев В.А. О специфике индивидуализации обучения в автоматизированных обучающих системах // Философия образования. – 2010. – №2. – С. 68–74.

8. Тарасенко Ф.П. Прикладной системный анализ: учеб. пособие. – М.: КноРус, 2010. – 224 с.

9. Шихнабиева Т.Ш. Методические основы представления и контроля знаний в области информатики с использованием адаптивных семантических моделей: дисс. ... д-ра пед. наук. – М., 2009. – 302 с.

10. Uglev V.A., Ustinov V.A., Samrina F.I. Approach to the Semantic Links organization about the education material structure in Automated Education Systems // Neuroinformatic, her application and data analysis. – Krasnoyarsk: IVMSORAN, 2010. – P. 171–175.

11. Винер Н. Кибернетика или управление и связь в животном и машине. – М.: Советское радио, 1968. – 326 с.

12. Углев В.А. Применение когнитивных карт диагностики знаний для оптимизации процесса обучения в автоматизированных средах // Информатизация образования – 2010: педагогические аспекты создания информационно-образовательной среды: матер. Междунар. науч.-практ. конф. – Минск: БГУ, 2010. – С. 506–510.

13. Углев В.А. Модель индивидуализации траектории обучения в автоматизированных обучающих системах с при-

менением обучающих тестов и экспертных систем // Нейроинформатика, ее приложения и анализ данных: XVI Всерос. семинар. – Красноярск, 2008. – С. 153–157.

14. Углев В.А. Использование механизма нечёткой логики и семантических сетей для подготовки информации по оценке степени освоения учебного материала // Нейроинформатика, ее приложения и анализ данных: XVII Всерос. семинар. – Красноярск, 2009. – С. 161–163.

15. Гаврилова Т.А., Муромцев Д.И. Интеллектуальные технологии в менеджменте. – М.: Высшая школа менеджмента, 2008. – 488 с.

16. Девятков В.В. Системы искусственного интеллекта: учеб. пособие. – М.: МГТУ, 2001. – 352 с.

17. Цибульский Г.М. Автоматизированные обучающие системы / Г.М. Цибульский, А.М. Кутьин, Е.И. Герасимова, В.А. Ерошин // Вестник Красноярского государственного технического университета. – Вып. 33: Математические методы и моделирование. – Красноярск, 2004. – С. 267–286.

18. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений: Пер. с англ. – Вып. 3. – М.: Мир, 1976. – 168 с.

19. Джексон П. Введение в экспертные системы: Пер. с англ. – М.: Вильямс, 2001. – 624 с.

20. Buchanan B.G., Shortliffe E.H. Rule-based Expert Systems. – USA, 1984. – 748 p.

21. Triantafillou E. Computerized Adaptive Test – Adapting to What? // Informatics Education Europe II Conference, SEERC. – 2007. – P. 379–385.

22. Углев В.А. Интеллектуальные автоматизированные обучающие системы, как прообраз индивидуальных ситуационных центров // Информатика: проблемы, методология, технологии: матер. XI Междунар. науч.-метод. конф.: в 3 ч. – Ч. 2. – Воронеж: ВГУ, 2011. – С. 392–396.