



Представление, обработка и контроль знаний на основе элементарных вербальных единиц

Проскурнин Алексей Александрович

Научный руководитель: к.т.н., доцент Филиппович Ю.Н.

Две фазы функционирования АСКЗ

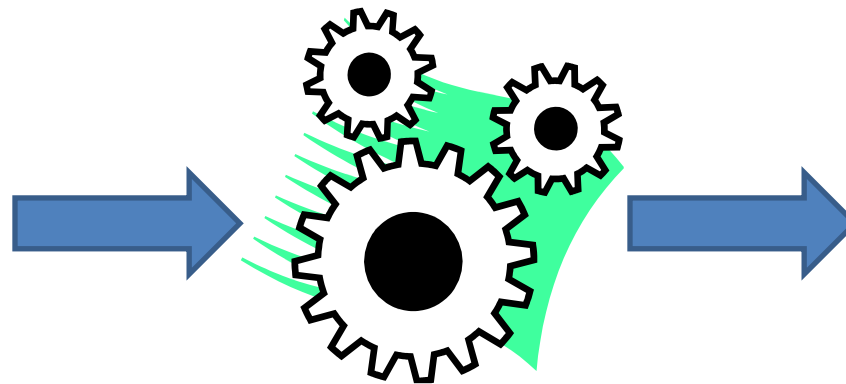
I. ФАЗА ВЗАИМОДЕЙСТВИЯ



II. ФАЗА ИНТЕРПРЕТАЦИИ



память компьютера



алгоритм
интерпретации

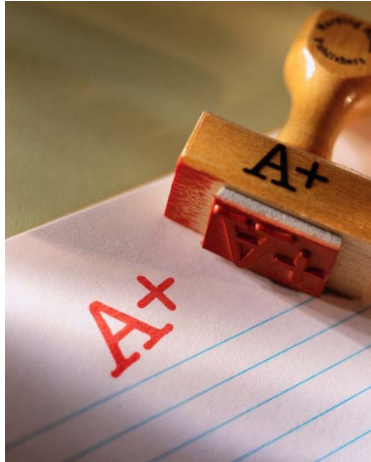


память компьютера

память компьютера

Две базовые парадигмы построения АСКЗ

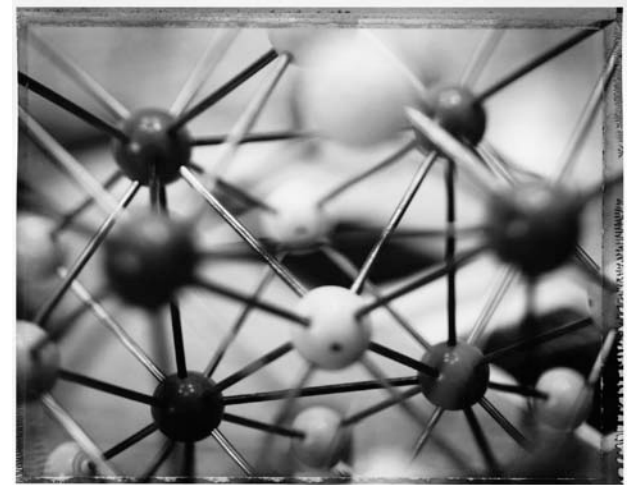
Скалярная
(отметка, ранг):
на **СКОЛЬКО** знает?



ПАРАДИГМА
ОЦЕНКИ ЗНАНИЙ

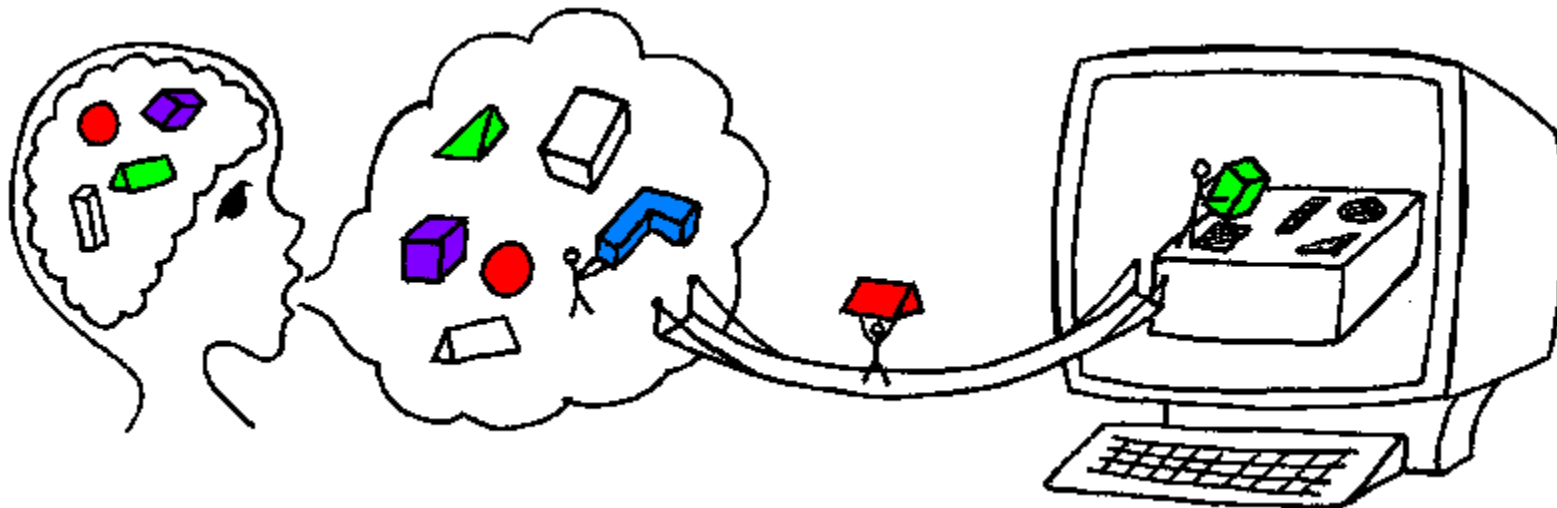


«Развернутая»,
структурная,
содержательная:
ЧТО знает/не знает?



ПАРАДИГМА
ВЫЯВЛЕНИЯ И АНАЛИЗА
ЗНАНИЙ

Понятие автоматизированного выявления знаний (АВЗ)



Автоматизированная система контроля знаний (**АСКЗ**) - автоматизированная система, основная функция которой состоит в формировании модели персональных знаний пользователя (МПЗ) в результате взаимодействия система-пользователь

Модель персональных знаний человека (**МПЗ**) - любая информация о знаниях человека в некоторой предметной области.

Автоматизированный контроль знаний (**АКЗ**) - процесс взаимодействия человека и компьютера, в результате которого в памяти компьютера формируется МПЗ человека.

Автоматизированное выявление знаний (АВЗ) - АКЗ, при котором МПЗ представляет собой данные, обладающие свойствами внутренней интерпретируемости, структурированности, и связности.

Актуальность

Актуальность обуславливается также необходимостью устранения следующих характерных недостатков в существующих АСКЗ, предназначенных для формирования содержательного описания знаний обучаемого:

1. Использование методов АКЗ, максимально учитывающих специфику конкретной предметной области, и неприменимых для других областей.
2. Моделирование в АСКЗ учебного курса (предмета), а не предметной области как таковой.
3. Поверхностность, низкий уровень детализации описания знаний обучаемого в АСКЗ.
4. Необходимость наличия достаточно больших репрезентативных выборок обучаемых.
5. С одной стороны, закрытая форма заданий обладает существенными недостатками; с другой стороны, для открытой формы заданий, подразумевающей анализ свободного естественно-языкового ответа обучаемого в виде предложения, крайне трудно гарантировать высокую адекватность такого анализа.
6. Применение в АСКЗ таких методов формирования МПЗ, при которых принципиально невозможно пояснить пользователю ход рассуждений, с помощью которых была получена модель его знаний.
7. Отсутствие в АСКЗ инструментов анализа знаний обучаемого, выявленных системой.
8. Ориентированность АСКЗ на один сеанс взаимодействия с обучаемым, после которого ему/ей выдаются результаты, а не на многие сеансы, подразумевающие «диалоги» с системой в течение продолжительного периода времени, когда с каждым сеансом взаимодействия пополняется (уточняется) МПЗ обучаемого.

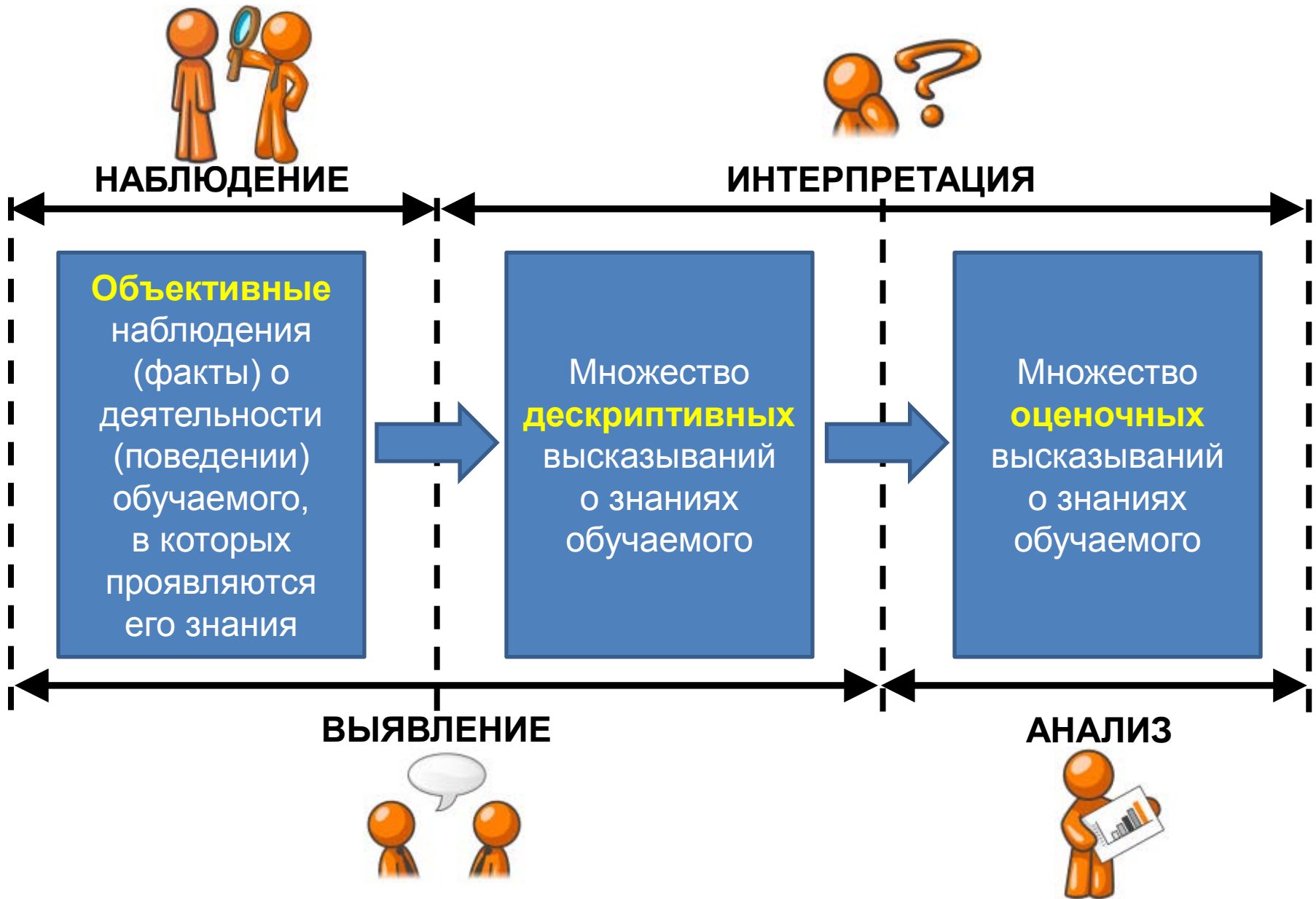
Цель и задачи исследования

Целью диссертационного исследования является разработка и практическая реализация в виде программной системы новой методики АКЗ, ориентированной на решение задачи формирования содержательного, структурного описания персональных знаний обучаемого, с учетом устранения описанных выше недостатков в существующих подходах к решению этой задачи.

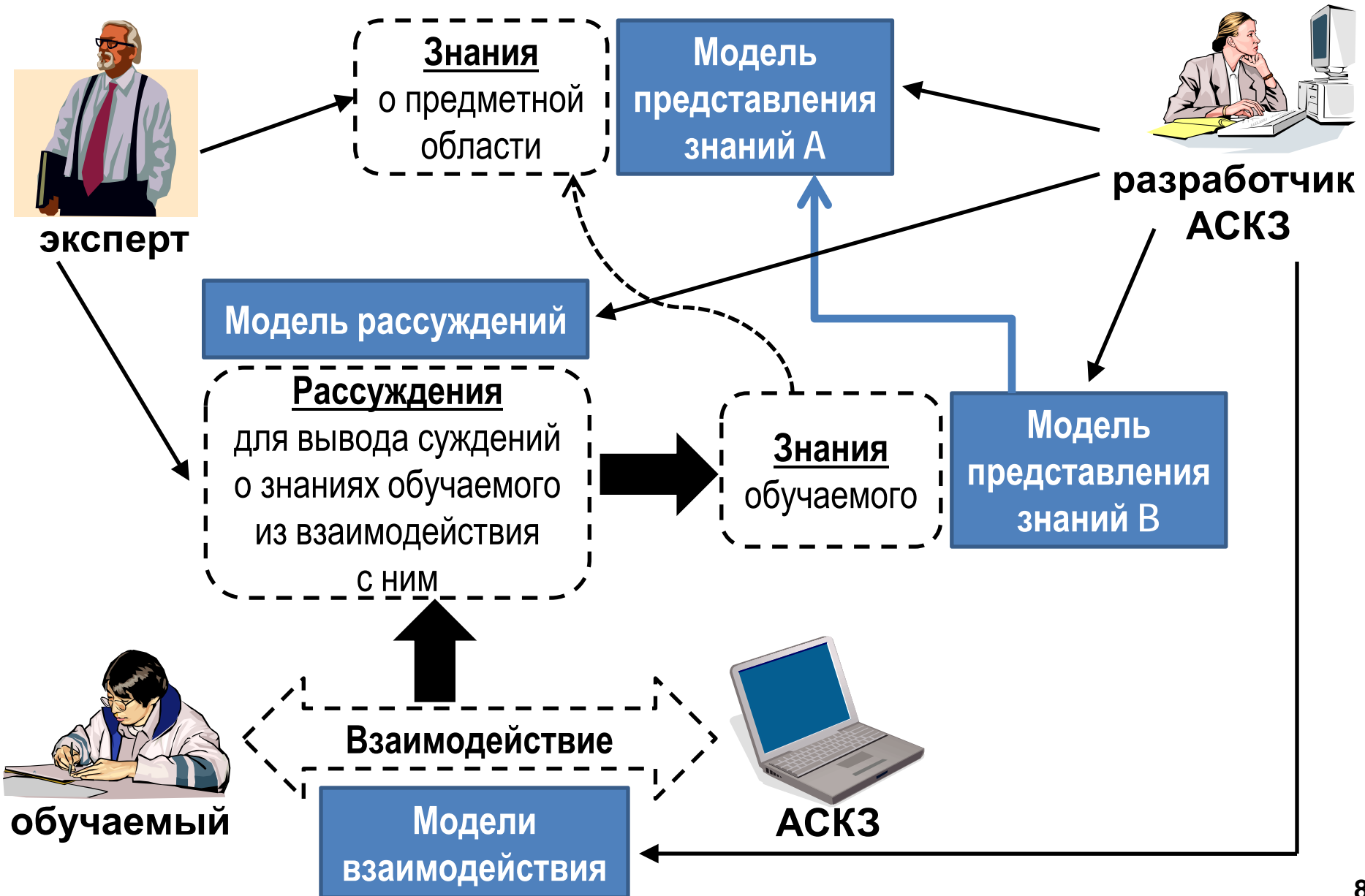
Для достижения этой цели в диссертационной работе решены следующие основные **задачи**:

1. Проведены анализ и классификация существующих методов и моделей АКЗ, а также моделей представления знаний и моделей рассуждений, используемых для решения задачи АКЗ.
2. Разработаны математические модели описания экспертных знаний и знаний обучаемого о предметной области для решения задачи АКЗ.
3. Разработаны алгоритмы АКЗ, основанные на использовании различных способов вывода на знаниях.
4. На основе предложенных моделей и алгоритмов разработана методика АКЗ.
5. Разработана и апробирована программная система, реализующая предлагаемую методику АКЗ.
6. Проведены экспериментальные исследования методики АКЗ.

Идея выявления и анализа знаний

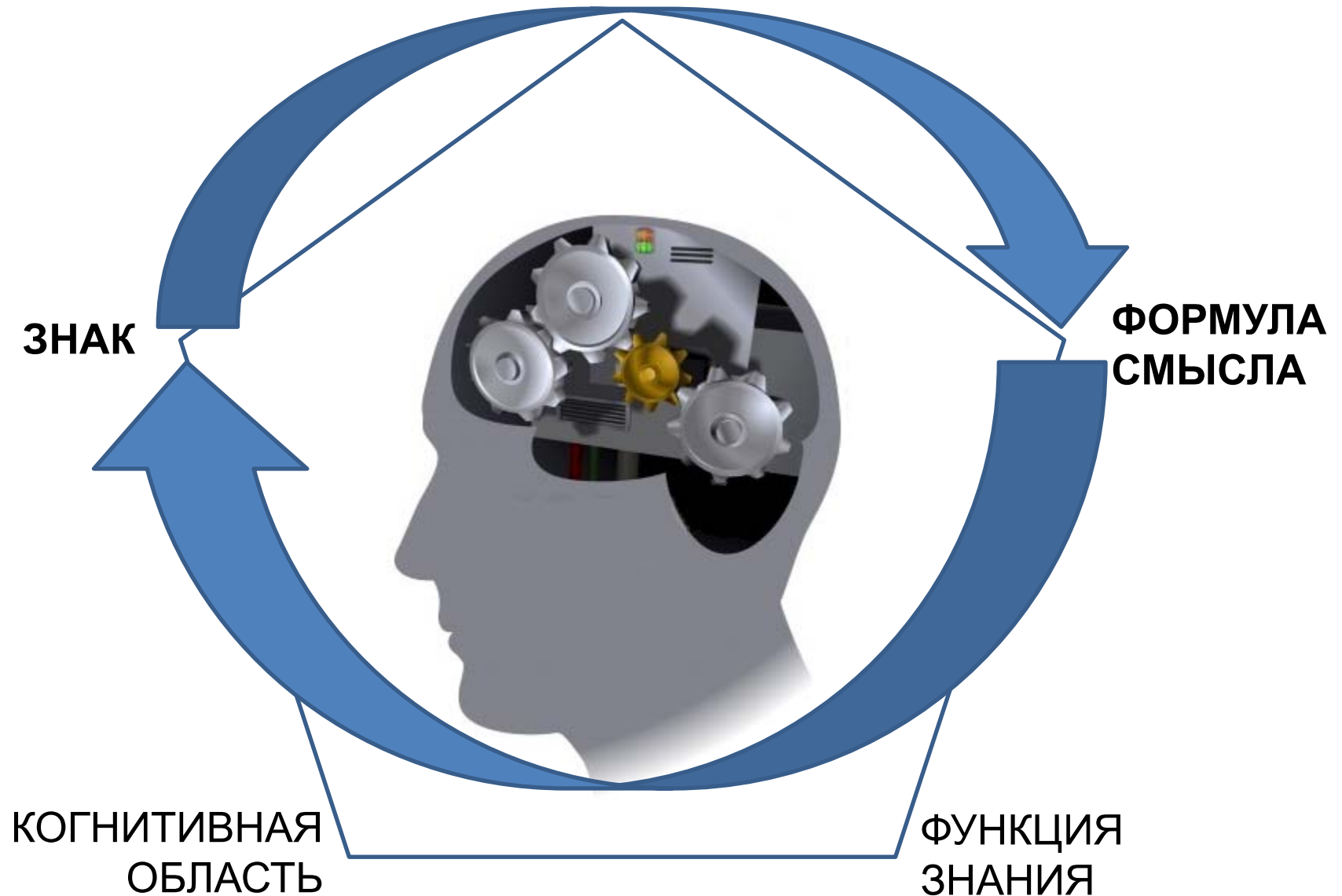


Концептуальная модель интеллектуальной АСКЗ



Понятие фигуры знания, предложенное Ю.Н. Карауловым

СПОСОБ ЗАДАНИЯ СМЫСЛА



Примеры фигур знания

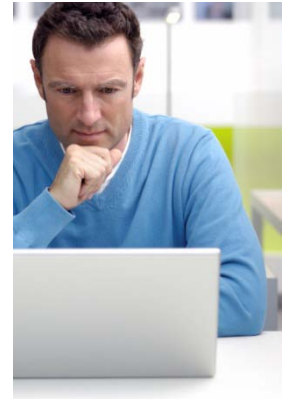
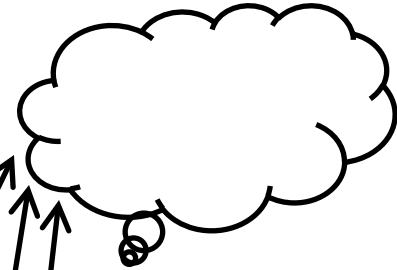
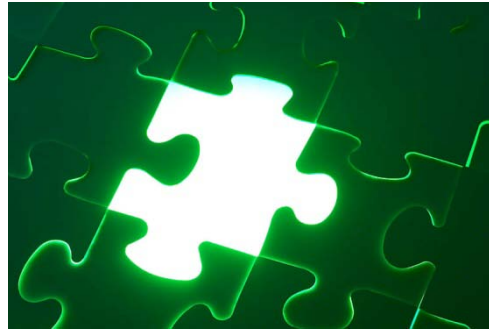
Знак	Способ задания смысла	Формула смысла	Область	Функция знания
Байт	Дескрипция	Восемь бит.	Единицы измерения	Рецепт
Мегабайт	Дефиниция	Единица измерения количества информации, равная 2^{20} байт.	Единицы измерения	Рецепт
Стек	Дескрипция	Он работает по принципу LIFO.	Структуры данных	Рецепт
Компилятор	Дескрипция	Вид транслятора, который обычно противопоставляется интерпретатору.	Программирование	Рецепт
Транзакция	Дескрипция	Она заканчивается при выполнении оператора COMMIT или оператора ROLLBACK.	Базы данных	Рецепт

Пример переходов «формула смысла - знак» и соответствующих им ассоциативных связей

Формула смысла	Знак	Ассоциативные связи
Восемь бит.	Байт	Бит↔Байт Число 8↔Байт
Единица измерения количества информации, равная 2^{20} байт.	Мегабайт	Единица измерения↔Мегабайт Информация↔Мегабайт Байт↔Мегабайт
Он работает по принципу LIFO.	Стек	LIFO↔Стек
Вид транслятора, который обычно противопоставляется интерпретатору.	Компилятор	Транслятор↔Компилятор Интерпретатор↔Компилятор
Она заканчивается при выполнении оператора COMMIT или оператора ROLLBACK.	Транзакция	COMMIT↔Транзакция ROLLBACK↔Транзакция

Применение фигур знания для решения задачи АКЗ

Популярный язык программирования, разработанный Никлаусом Виртом.



+ +

Паскаль
Pascal

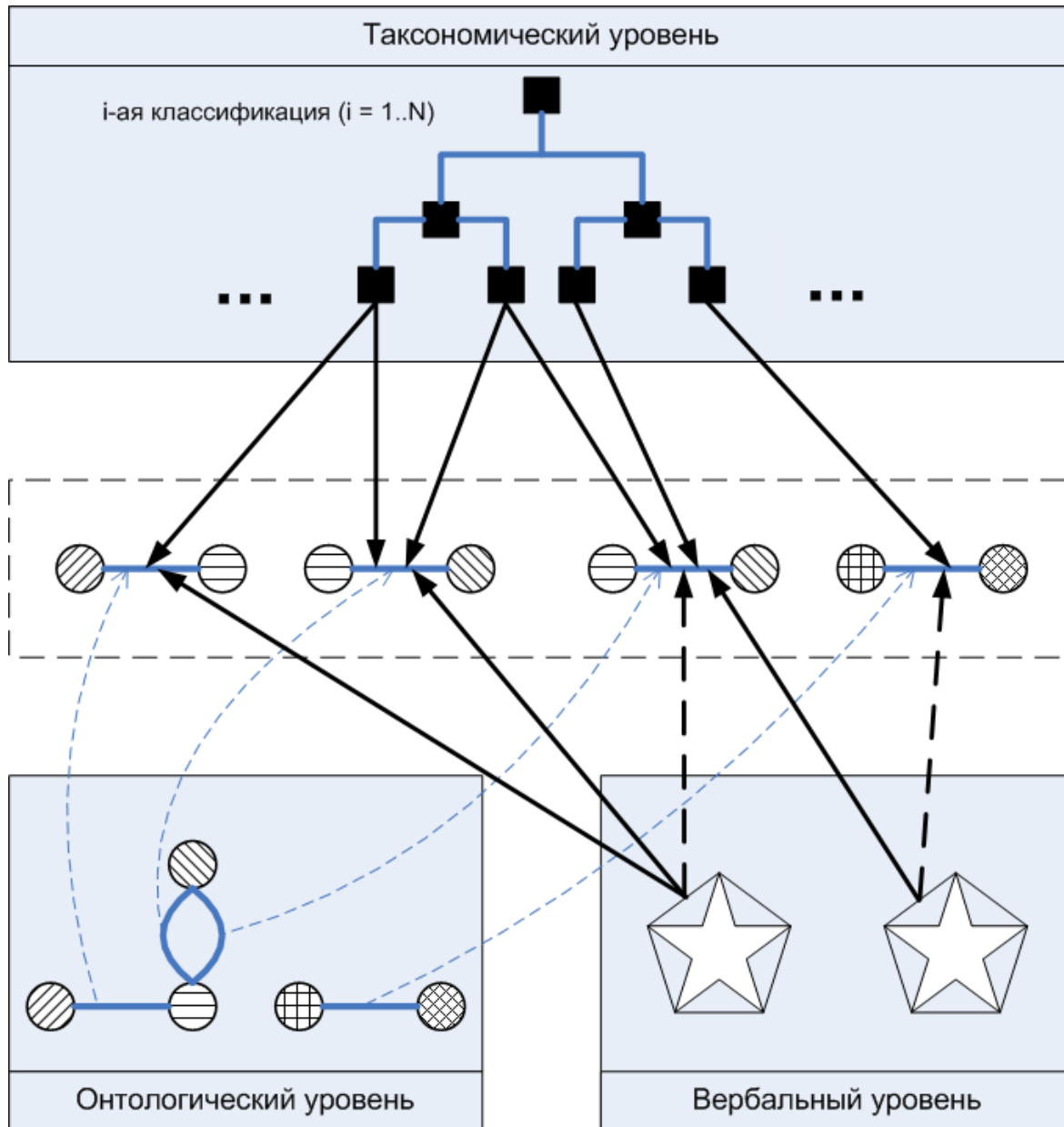
язык программирования-
Паскаль
Никлаус Вирт-Паскаль

Бейсик
[]

~~Никлаус Вирт-Паскаль~~

-

Упрощенное схематичное изображение модели предметной области



Обозначения:



Понятие (концепт)



Целевая единица знания

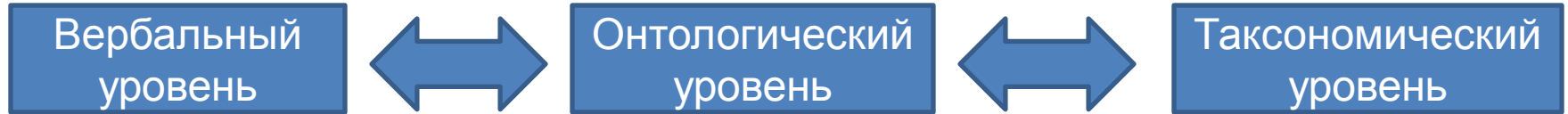


Классификационная единица



Фигура знания

Представление знаний о предметной области



$$M_{\text{ПР.ОБЛ.}} = \langle M_{\text{ОНТ.}}, M_{\text{ВЕРБ.}}, M_{\text{ТАКС.}}, ML_{\text{ОНТ.-ВЕРБ.}}, ML_{\text{ОНТ.-ТАКС.}} \rangle$$

$$M_{\text{ОНТ.}} = \langle T, E, A, L, f_{ET}, f_{LEEA}, H_L \rangle$$

$$M_{\text{ВЕРБ.}} = \langle Q, P, V, Z, W, f_{QPV}, R_{QZW} \rangle$$

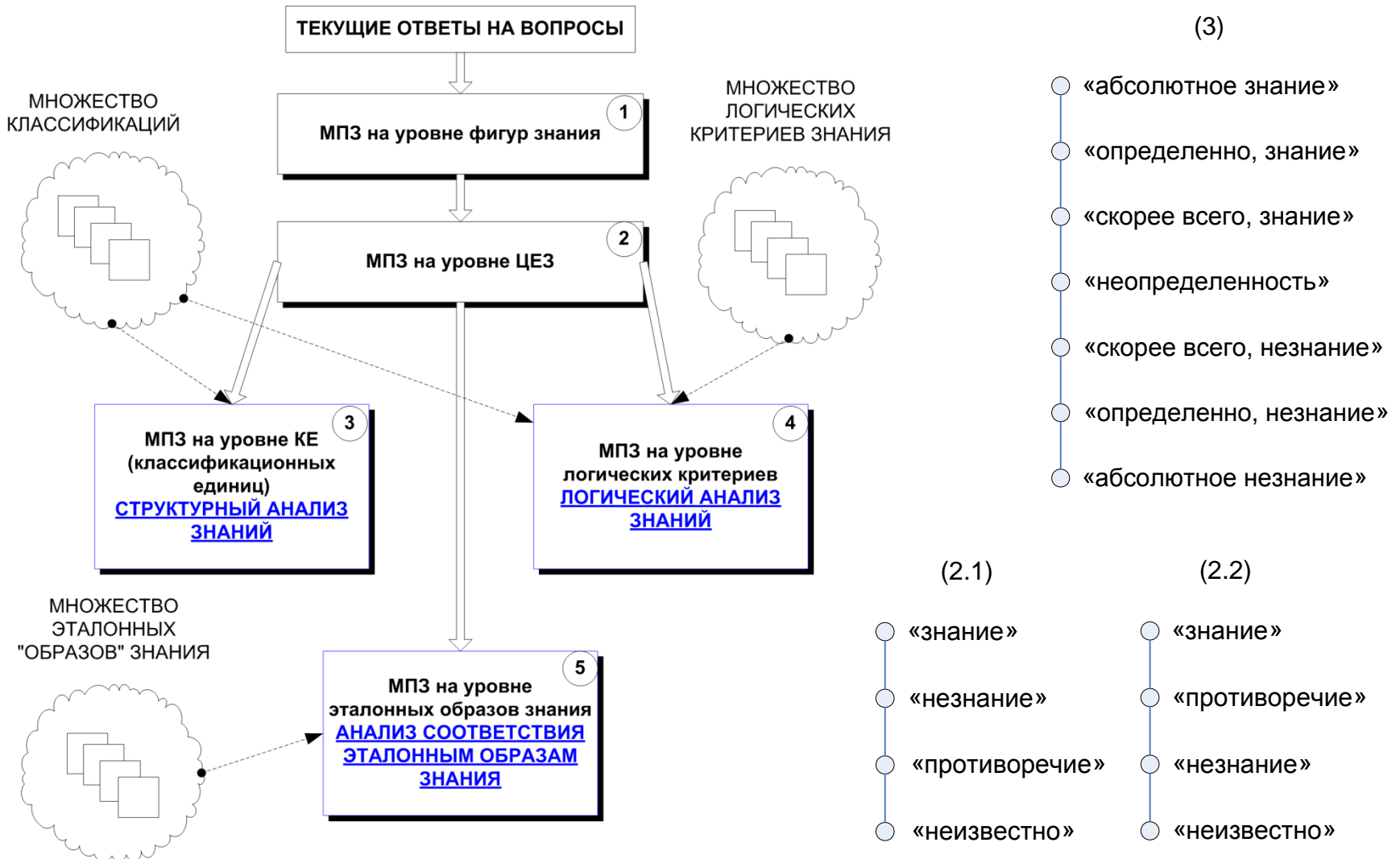
$$M_{\text{ТАКС.}} = \langle D, C, f_{CD}, R_{CC}, (\tilde{R}_d^{3H.})_{d \in D}, (\tilde{R}_d^{HE3H.})_{d \in D}, (H_{Cd})_{d \in D}, I, G \rangle$$

$$ML_{\text{ОНТ.-ВЕРБ.}} = \langle f^{3H.}, f^{HE3H.}, f_{QE} \rangle$$

$$ML_{\text{ОНТ.-ТАКС.}} = \langle (R_{CLd})_{d \in D}, (f_{EGi})_{i \in I} \rangle$$

Представление персональных знаний обучаемого

$$M_{ПЗ} = \langle M_{ПР.ОБЛ.}, S_Q, S_L, S_C, f_Q, f_L, f_C, f_{HL}, f_{HC}, f_I, f_{DC}, f_{DHC}, f_{DI} \rangle$$



Формальная постановка задачи

ДАНО:

Знание (экспертное) о предметной области:

$$M_{\text{ПР.ОБЛ.}}$$

Текущие ответы испытуемого на вопросы:

$$f_{\text{ОТВ.}} : Q \rightarrow B$$

$$f_{\text{ОТВ.}} : Q \rightarrow \Delta^*$$

ОПРЕДЕЛИТЬ:

МПЗ испытуемого (выявление и анализ знаний):

$$f_Q : Q \rightarrow S_Q \quad f_L : L \rightarrow S_L$$

$$f_C : C \rightarrow S_C \quad f_{HL} : H_L \rightarrow S_L$$

$$f_{HC} : \bigcup_{d \in D} (H_{Cd}) \rightarrow S_C \quad f_I : I \rightarrow [0,1]$$

$$f_{DHC} : \bigcup_{d \in D} (H_{Cd}) \rightarrow [0,1] \quad f_{DC} : C \rightarrow [0,1]$$

$$f_{DI} : I \rightarrow [0,1]$$

Типы моделируемых рассуждений

В работе реализовано моделирование **следующих типов рассуждений** человека-эксперта, осуществляющего контроль знаний:

1. Рассуждения, связанные с формированием суждений о присутствии в сознании человека определенных когнитивных единиц на основании информации, полученной от человека в процессе взаимодействия.
2. Рассуждения, связанные с формированием суждений о степени знания одних когнитивных единиц, исходя из степени знания других единиц, когда и те, и другие единицы находятся на одном уровне иерархии в структурной декомпозиции предметной области.
3. Рассуждения, связанные с формированием суждений о степени знания некоторого множества элементов, исходя из знания/незнания этих элементов (операция «свертки»).
4. Рассуждения, связанные с выбором следующего вопроса, который будет задан человеку в процессе адаптивного взаимодействия, исходя из текущей МПЗ человека и выбранной стратегии взаимодействия.

Рассуждения 1-го типа: логика аргументации А4

Шаг 1. Введем мультимножества: $m_{ЗН.} : L \rightarrow N$ $m_{НЕЗН.} : L \rightarrow N$

Шаг 2. $\forall l \in L : m_{ЗН.}(l) = 0$ $m_{НЕЗН.}(l) = 0$

Шаг 3. В цикле по всем $q \in Q$

Если ответ на вопрос верный, то добавить в $m_{ЗН.}$
множество ЦЕЗ $\{l \in f^{ЗН.}(q)\}$

Если ответ на вопрос неверный, то добавить в $m_{НЕЗН.}$
множество ЦЕЗ $\{l \in f^{НЕЗН.}(q)\}$

Шаг 4. В цикле по всем $l \in L$

$m_{ЗН.}(l)$
МНОЖЕСТВО
аргументов

$m_{ЗН.}(l) = 0 \wedge m_{НЕЗН.}(l) = 0$

НЕИЗВЕСТНОСТЬ

$m_{ЗН.}(l) > 0 \wedge m_{НЕЗН.}(l) = 0$

ЗНАНИЕ

$m_{НЕЗН.}(l)$

МНОЖЕСТВО
контраргументов

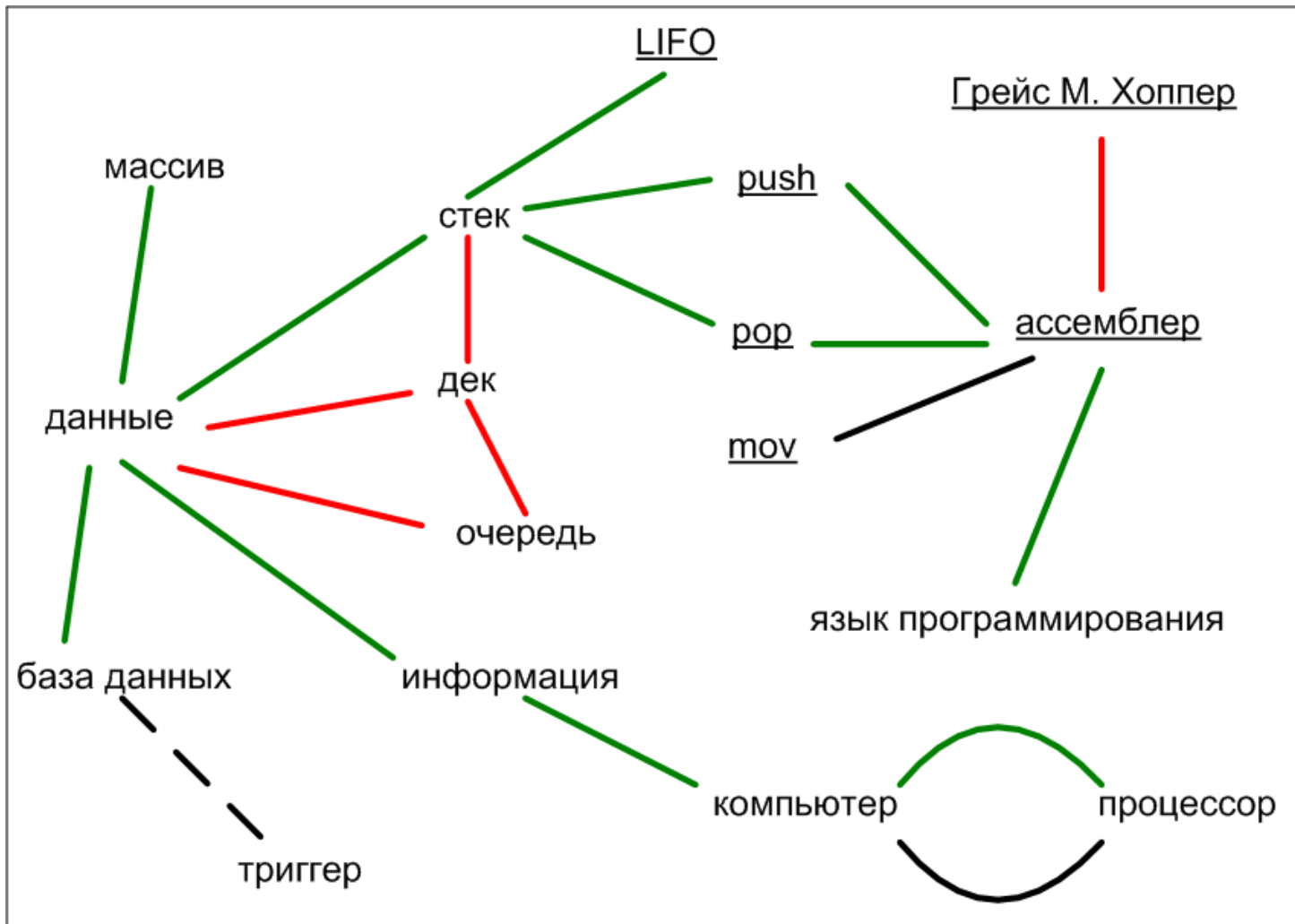
$m_{ЗН.}(l) = 0 \wedge m_{НЕЗН.}(l) > 0$

НЕЗНАНИЕ

$m_{ЗН.}(l) > 0 \wedge m_{НЕЗН.}(l) > 0$

ПРОТИВОРЕЧИЕ

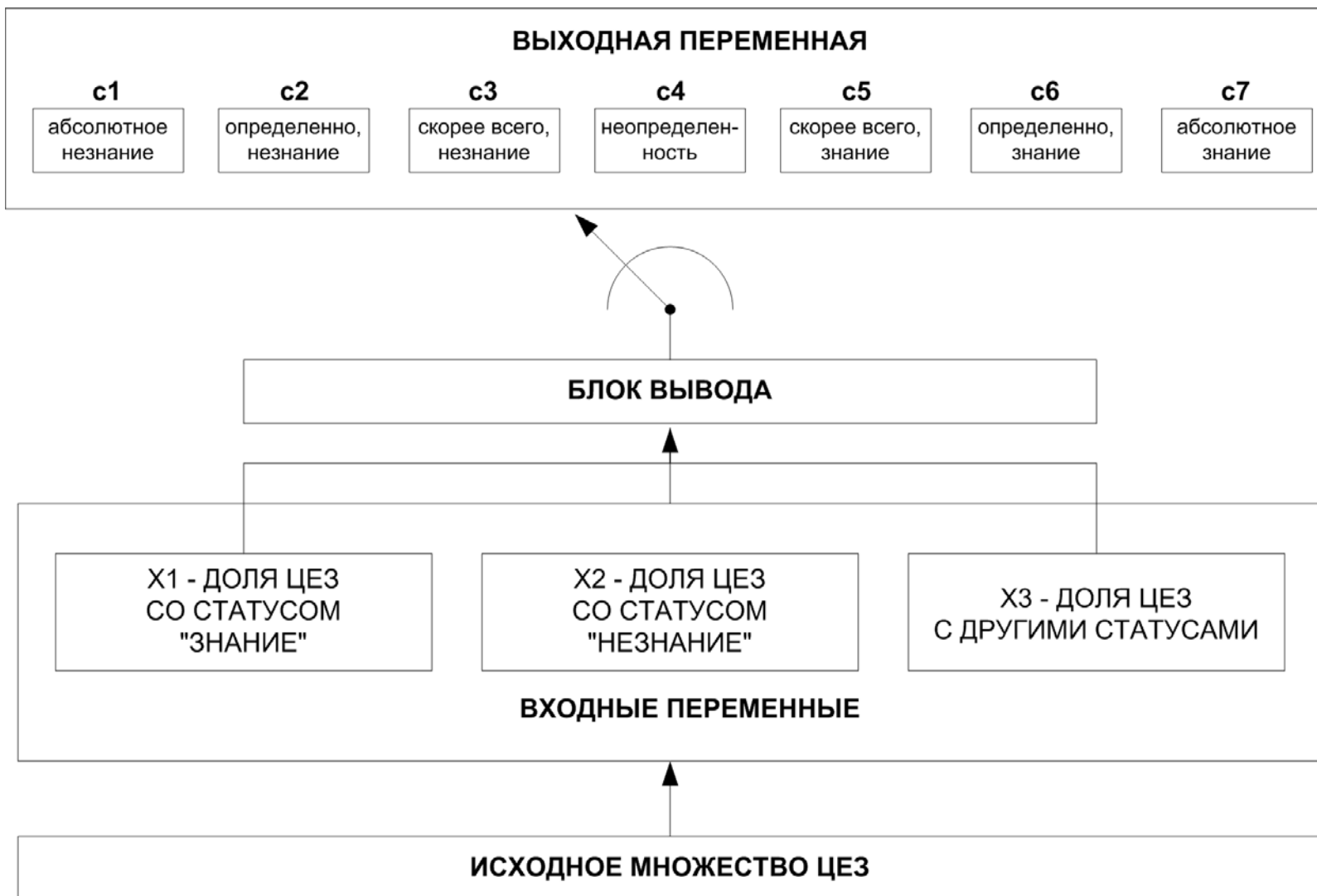
Пример фрагмента МПЗ на уровне ЦЕЗ



Каждому i -му ребру соответствует высказывание: «В персональное знание обучаемого входит i -ая ЦЕЗ». Используется 4 истинностных значения:

- фактическая истина
- фактическая ложь
- неизвестность
- - - фактическое противоречие

Рассуждения 3-го типа: операция «свертки»

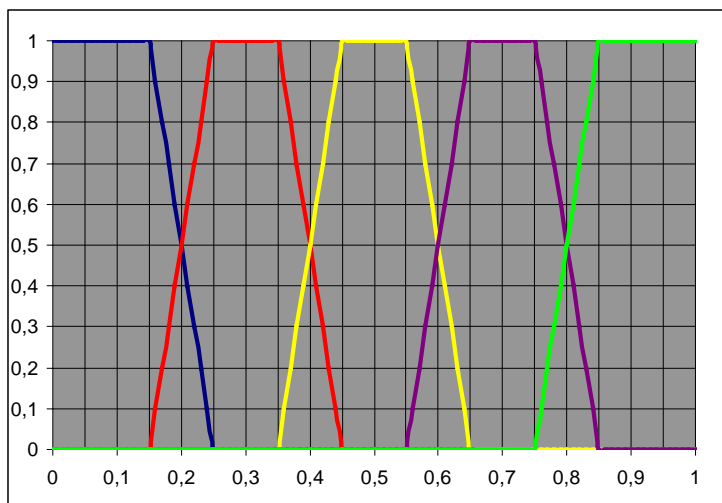


Варианты реализации операции «свертки» (1)

Вариант 1. Использование разности $DX = X_{3H.} - X_{HE3H.}$ $DX \in [-1, 1]$

Вариант 2. Система нечеткого логического вывода

Функции принадлежности:



Правила вывода:

$$\mu^{c1}(X_{3H.}, X_{HE3H.}) = \mu_{X_{3H.}}^{oH} \wedge \mu_{X_{HE3H.}}^{oB}$$

$$\mu^{c2}(X_{3H.}, X_{HE3H.}) = [\mu_{X_{3H.}}^{oH} \wedge \mu_{X_{HE3H.}}^B] \vee [\mu_{X_{3H.}}^H \wedge \mu_{X_{HE3H.}}^B]$$

$$\mu^{c3}(X_{3H.}, X_{HE3H.}) = [\mu_{X_{3H.}}^{oH} \wedge \mu_{X_{HE3H.}}^C] \vee [\mu_{X_{3H.}}^H \wedge \mu_{X_{HE3H.}}^C]$$

$$\mu^{c4}(X_{3H.}, X_{HE3H.}) = [\mu_{X_{3H.}}^{oH} \wedge \mu_{X_{HE3H.}}^H] \vee [\mu_{X_{3H.}}^{oH} \wedge \mu_{X_{HE3H.}}^{oH}] \vee$$

$$[\mu_{X_{3H.}}^H \wedge \mu_{X_{HE3H.}}^H] \vee [\mu_{X_{3H.}}^C \wedge \mu_{X_{HE3H.}}^C] \vee [\mu_{X_{3H.}}^H \wedge \mu_{X_{HE3H.}}^{oH}]$$

$$\mu^{c5}(X_{3H.}, X_{HE3H.}) = [\mu_{X_{3H.}}^C \wedge \mu_{X_{HE3H.}}^H] \vee [\mu_{X_{3H.}}^C \wedge \mu_{X_{HE3H.}}^{oH}]$$

$$\mu^{c6}(X_{3H.}, X_{HE3H.}) = [\mu_{X_{3H.}}^B \wedge \mu_{X_{HE3H.}}^H] \vee [\mu_{X_{3H.}}^B \wedge \mu_{X_{HE3H.}}^{oH}]$$

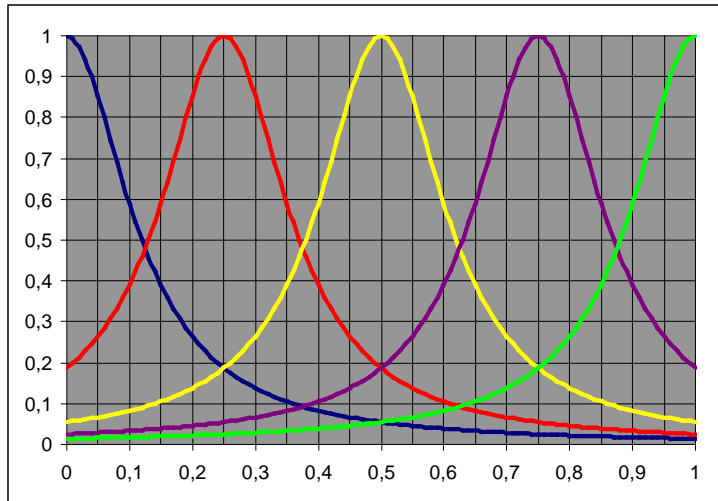
$$\mu^{c7}(X_{3H.}, X_{HE3H.}) = \mu_{X_{3H.}}^{oB} \wedge \mu_{X_{HE3H.}}^{oH}$$

$$f_{TP}(x; a, b, c, d) = \left\{ \begin{array}{ll} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & b \leq x \leq c \\ \frac{d-x}{d-c}, & c \leq x \leq d \\ 0, & d \leq x \end{array} \right.$$

Варианты реализации операции «свертки» (2)

Вариант 3. Система нечеткого логического вывода и настраивающий ее генетический алгоритм

Функции принадлежности:



$$\mu(x; a, b) = \frac{1}{1 + \left(\frac{x-b}{a}\right)^2}$$

Правила вывода:

$$\mu^{c1}(X_{3H.}, X_{HE3H.}, X_{HEOIP.}) = \{w_1 \cdot \mu_{X_{HE3H.}}^{oB}\} \vee \{w_2 \cdot [\mu_{X_{HE3H.}}^B \wedge \mu_{X_{3H.}}^{oH}]\}$$

$$\mu^{c2}(X_{3H.}, X_{HE3H.}, X_{HEOIP.}) = w_3 \cdot [\mu_{X_{HE3H.}}^B \wedge \mu_{X_{3H.}}^H]$$

$$\mu^{c3}(X_{3H.}, X_{HE3H.}, X_{HEOIP.}) = \{w_4 \cdot [\mu_{X_{HE3H.}}^C \wedge \mu_{X_{3H.}}^H]\} \vee \{w_5 \cdot [\mu_{X_{HE3H.}}^C \wedge \mu_{X_{3H.}}^{oH}]\} \vee \{w_6 \cdot [\mu_{X_{HE3H.}}^H \wedge \mu_{X_{3H.}}^{oH}]\}$$

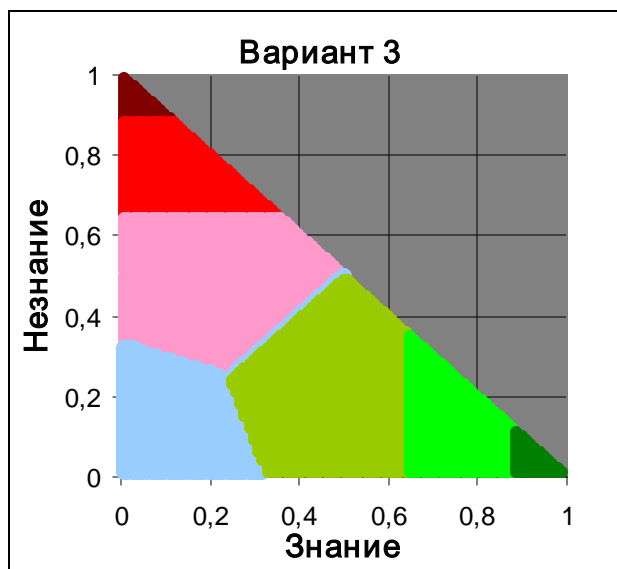
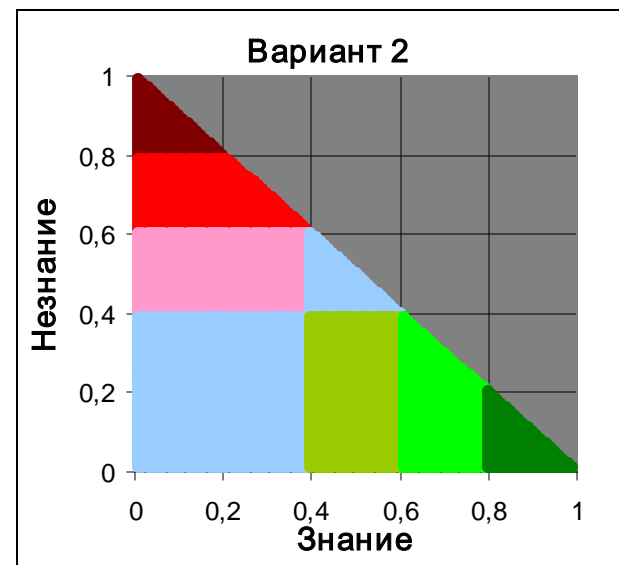
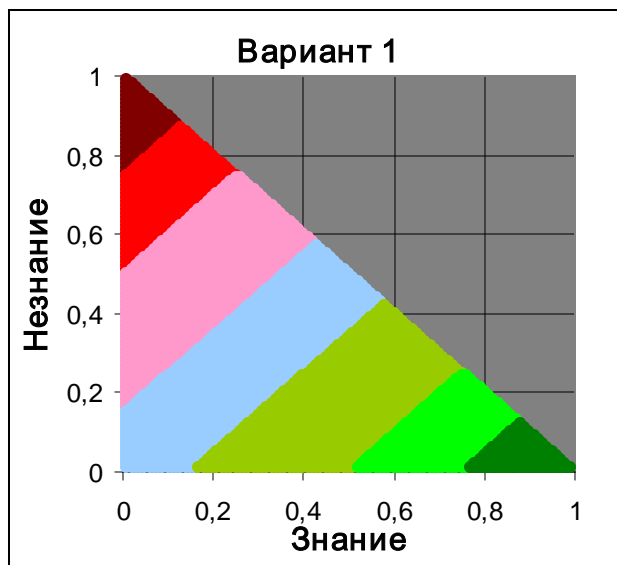
$$\mu^{c4}(X_{3H.}, X_{HE3H.}, X_{HEOIP.}) = w_7 \cdot \mu_{X_{HEOIP.}}^{oB}$$

$$\mu^{c5}(X_{3H.}, X_{HE3H.}, X_{HEOIP.}) = \{w_4 \cdot [\mu_{X_{3H.}}^C \wedge \mu_{X_{HE3H.}}^H]\} \vee \{w_5 \cdot [\mu_{X_{3H.}}^C \wedge \mu_{X_{HE3H.}}^{oH}]\} \vee \{w_6 \cdot [\mu_{X_{3H.}}^H \wedge \mu_{X_{HE3H.}}^{oH}]\}$$

$$\mu^{c6}(X_{3H.}, X_{HE3H.}, X_{HEOIP.}) = w_3 \cdot [\mu_{X_{3H.}}^B \wedge \mu_{X_{HE3H.}}^H]$$

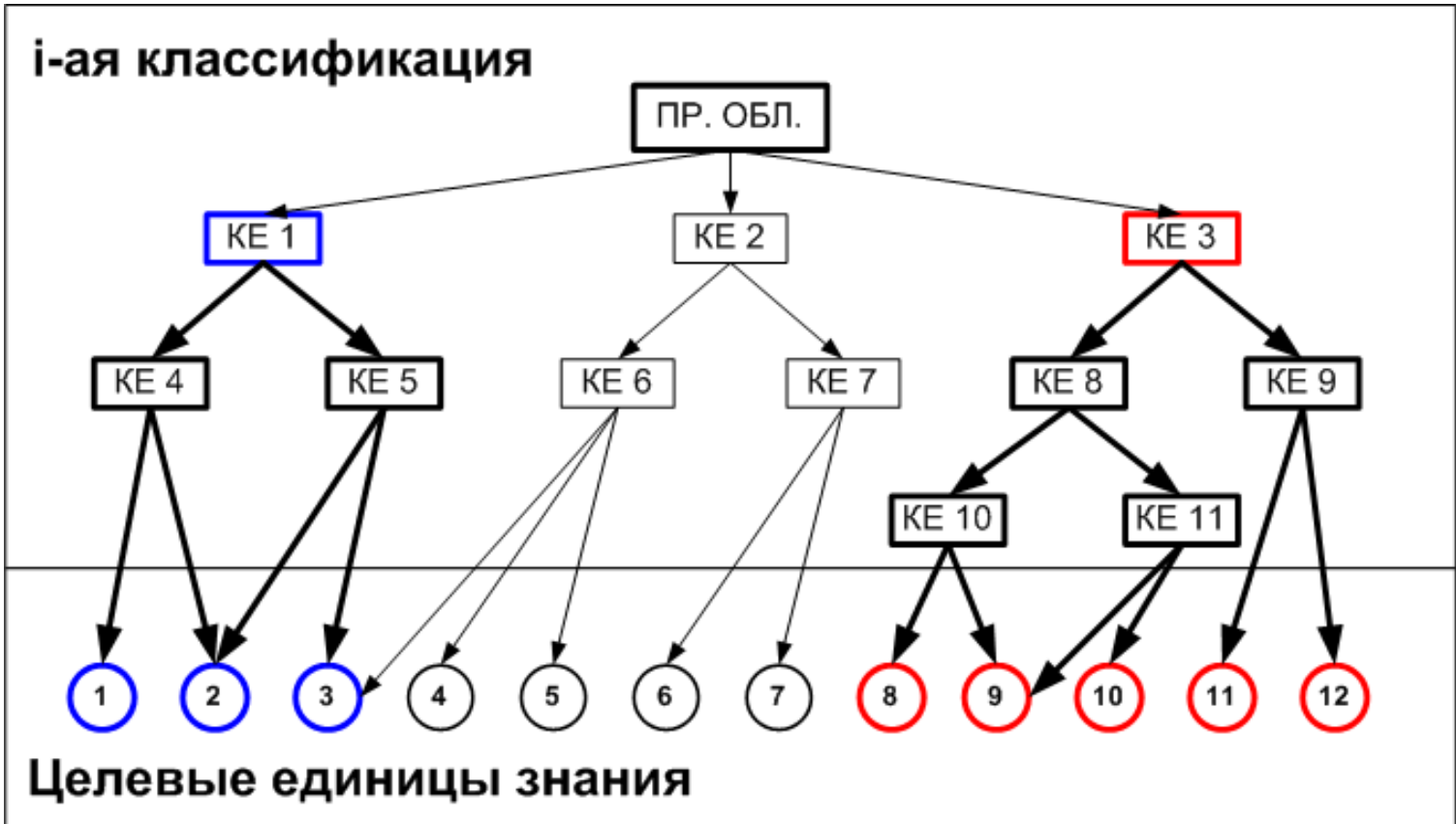
$$\mu^{c7}(X_{3H.}, X_{HE3H.}, X_{HEOIP.}) = \{w_1 \cdot \mu_{X_{3H.}}^{oB}\} \vee \{w_2 \cdot [\mu_{X_{3H.}}^B \wedge \mu_{X_{HE3H.}}^{oH}]\}$$

Варианты реализации операции «свертки» (3)



- Абсолютное незнание
- Определенно, незнание
- Скорее всего, незнание
- Неопределенность
- Скорее всего, знание
- Определенно, знание
- Абсолютное знание

Определение множества ЦЕЗ, соответствующих данной КЕ (примеры)



Форма для выполнения структурного анализа знаний

Система выявления и анализа знаний [ОСНОВЫ ИНФОРМАТИКИ] - [Модель персональных знаний]

Модель знаний Компоненты Анализ Диалог Справка

Модель персональных знаний

Классификация: Тематическая классификация (вариант 1) Показать Раскрыть все узлы

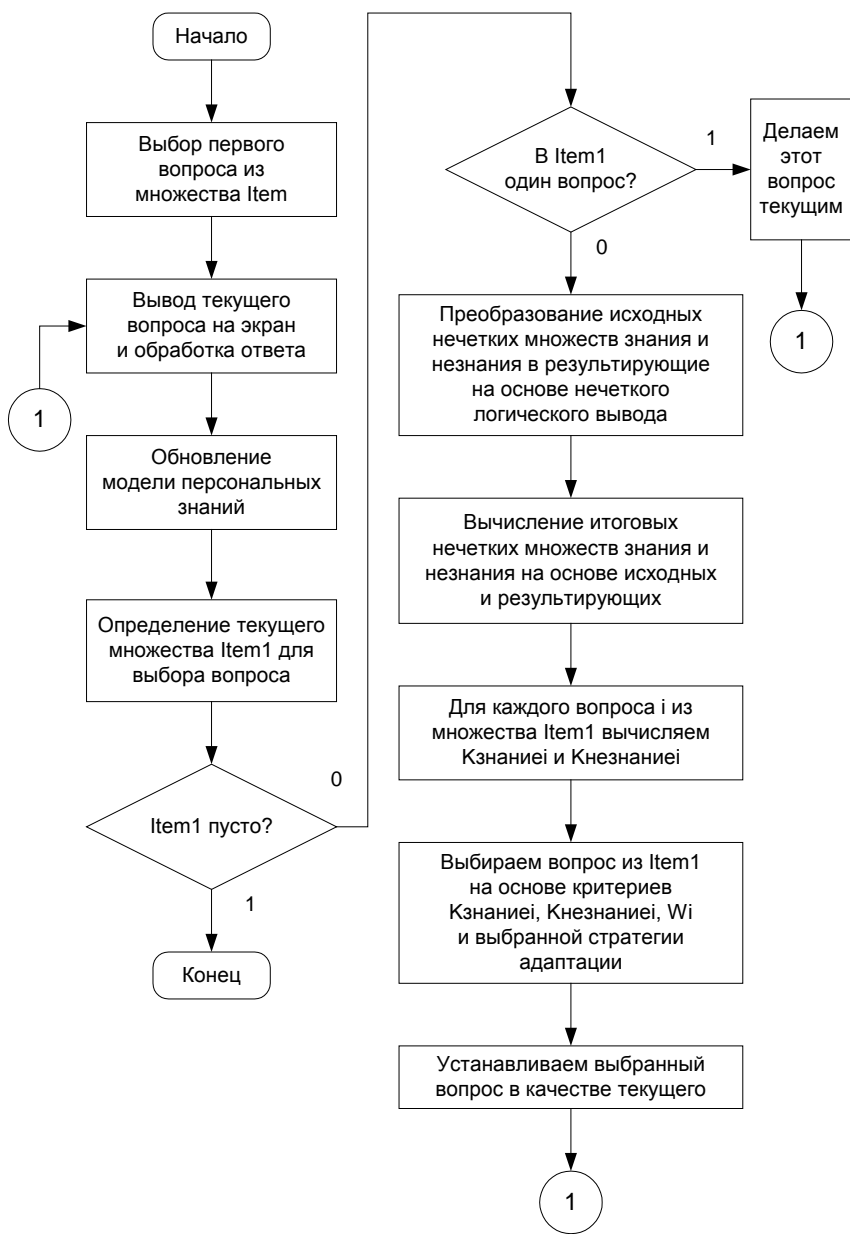
Показывать названия связей

Структура | Логические критерии | Соответствие эталонным образам знания

Название	Статус	% заданных вопросов
ИНФОРМАТИКА	скорее всего, знание	95
ОСНОВЫ ТЕОРИИ ИНФОРМАЦИИ	скорее всего, знание	95
Понятие информации	скорее всего, знание	90
Количество информации	неопределенность	100
Энтропия	скорее всего, знание	100
Единицы измерения информации	скорее всего, знание	100
Теория передачи информации	неопределенность	100
Системы счисления	определенно, знание	83
ЧТО ТАКОЕ КОМПЬЮТЕР	скорее всего, знание	67
Понятие компьютера	неопределенность	75
Классификация компьютеров	скорее всего, знание	50
АППАРАТНОЕ ОБЕСПЕЧЕНИЕ КОМПЬЮТЕРА	абсолютное знание	100
Устройства ввода информации	абсолютное знание	100
Устройства вывода информации	абсолютное знание	100
Процессоры	абсолютное знание	100
Запоминающие устройства	абсолютное знание	100
ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ КОМПЬЮТЕРА	определенно, знание	100
Операционные системы	определенно, знание	100
АЛГОРИТМИЗАЦИЯ И ПРОГРАММИРОВАНИЕ	скорее всего, знание	97
Основы алгоритмизации и программирования	определенно, знание	97
Структуры данных	абсолютное знание	84
Язык ассемблера	скорее всего, знание	100
Язык программирования Basic	абсолютное знание	100
Язык программирования Pascal	скорее всего, незнание	100
Язык программирования C	абсолютное незнание	100
Язык программирования C++	абсолютное незнание	100
Язык программирования Java	определенно, знание	100
Язык программирования Fortran	абсолютное знание	100
Язык программирования C#	скорее всего, знание	100
Языки программирования искусственного интеллекта	абсолютное знание	100
Объектно-ориентированное программирование	абсолютное незнание	100
ДИСКРЕТНАЯ МАТЕМАТИКА	скорее всего, знание	90
Теория множеств и отношений	абсолютное знание	100
Теория графов	определенно, знание	91
Формальные языки, автоматы и грамматики	скорее всего, незнание	83
Булевы функции	определенно, незнание	90
Математическая логика	скорее всего, знание	75
ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ	определенно, знание	93

Закреть

Рассуждения 4-го типа: механизм адаптации



$$Z_{3H.} = \{l \in L : f_L(l) = \text{"знание"}\} \quad Z_{HE3H.} = \{l \in L : f_L(l) = \text{"незнание"}\}$$

$$C_d^{нижн.} = \{c \in C : f_{CD}(c) = d \wedge (\neg \exists cp \in C : (c, cp) \in R_{CC})\}$$

$$CU_j = \{l \in L : \exists (c_j, l) \in R_{CLd}\} \quad 1 \leq j \leq |C_d^{нижн.}| \quad |C_d^{нижн.}| = K$$

$$QL_i = \{l \in f^{3H.}(q_i)\} \quad 1 \leq i \leq N$$

$$\mu_j^{3H.} = \frac{|Z_{3H.} \cap CU_j|}{|CU_j|} \quad \mu_j^{HE3H.} = \frac{|Z_{HE3H.} \cap CU_j|}{|CU_j|}$$

$$\theta_{исх.}^{3H.} = \{ \langle c_1, \mu_1^{3H.} \rangle, \langle c_2, \mu_2^{3H.} \rangle, \dots, \langle c_K, \mu_K^{3H.} \rangle \}$$

$$\theta_{исх.}^{HE3H.} = \{ \langle c_1, \mu_1^{HE3H.} \rangle, \langle c_2, \mu_2^{HE3H.} \rangle, \dots, \langle c_K, \mu_K^{HE3H.} \rangle \}$$

$$\theta_{рез.}^{3H.} = \theta_{исх.}^{3H.} \circ \tilde{R}_d^{3H.} \quad \theta_{рез.}^{HE3H.} = \theta_{исх.}^{HE3H.} \circ \tilde{R}_d^{HE3H.}$$

$$\theta_{рез.}^{3H.} = \{ \langle c_1, \beta_1^{3H.} \rangle, \langle c_2, \beta_2^{3H.} \rangle, \dots, \langle c_K, \beta_K^{3H.} \rangle \}$$

$$\theta_{рез.}^{HE3H.} = \{ \langle c_1, \beta_1^{HE3H.} \rangle, \langle c_2, \beta_2^{HE3H.} \rangle, \dots, \langle c_K, \beta_K^{HE3H.} \rangle \}$$

$$\theta_{умог.}^{3H.} = \{ \langle c_1, \lambda_1^{3H.} \rangle, \langle c_2, \lambda_2^{3H.} \rangle, \dots, \langle c_K, \lambda_K^{3H.} \rangle \}$$

$$\theta_{умог.}^{HE3H.} = \{ \langle c_1, \lambda_1^{HE3H.} \rangle, \langle c_2, \lambda_2^{HE3H.} \rangle, \dots, \langle c_K, \lambda_K^{HE3H.} \rangle \}$$

$$\alpha_i^j = \begin{cases} 1, & QL_i \cap CU_j \neq \emptyset \\ 0, & QL_i \cap CU_j = \emptyset \end{cases} \quad K_i^\Psi = \frac{\sum_{j:\alpha_i^j=1} \lambda_j^\Psi}{\text{Count}(j)}$$

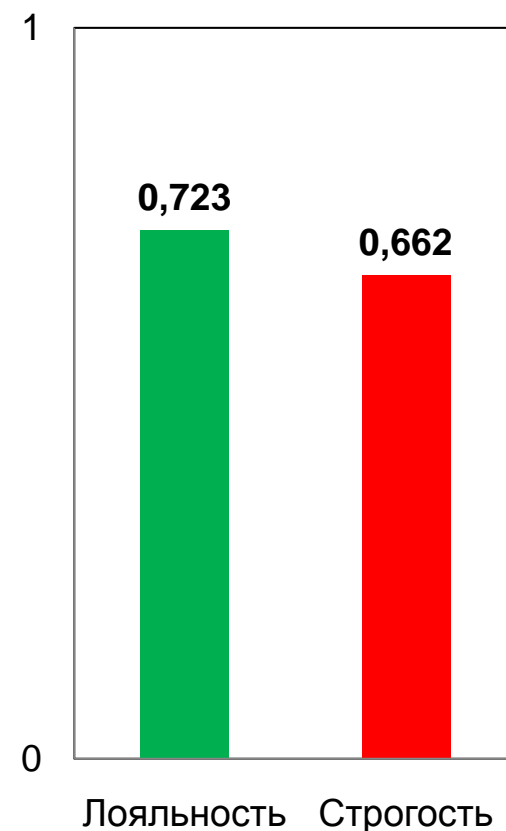
$$\Psi \in \{3H.; HE3H.\}$$

Критерии эффективности адаптации

$$K_{ML} = \frac{N^{АДАПТ+} - N^{СЛ+}}{N}$$

$$K_{МС} = \frac{N^{АДАПТ-} - N^{СЛ-}}{N}$$

№	ЭФФЕКТИВНОСТЬ АДАПТАЦИИ							
	Стратегия адаптации:							
	МАКСИМАЛЬНАЯ ЛОЯЛЬНОСТЬ				МАКСИМАЛЬНАЯ СТРОГОСТЬ			
	Использование нечеткого логического вывода для "расширения" областей знания и незнания (НЕТ, ДА):							
	НЕТ		ДА		НЕТ		ДА	
	Способ решения задачи многокритериального выбора (1, 2)							
	1	2	1	2	1	2	1	2
1	0,713	0,695	0,718	0,704	0,711	0,631	0,759	0,711
2	0,671	0,729	0,687	0,728	0,637	0,622	0,658	0,641
3	0,746	0,759	0,781	0,745	0,621	0,629	0,672	0,657
Ср. знач.	0,710	0,728	0,729	0,726	0,656	0,627	0,696	0,670
Ср. знач.	0,719		0,727		0,642		0,683	
Ср. знач.	0,723				0,662			

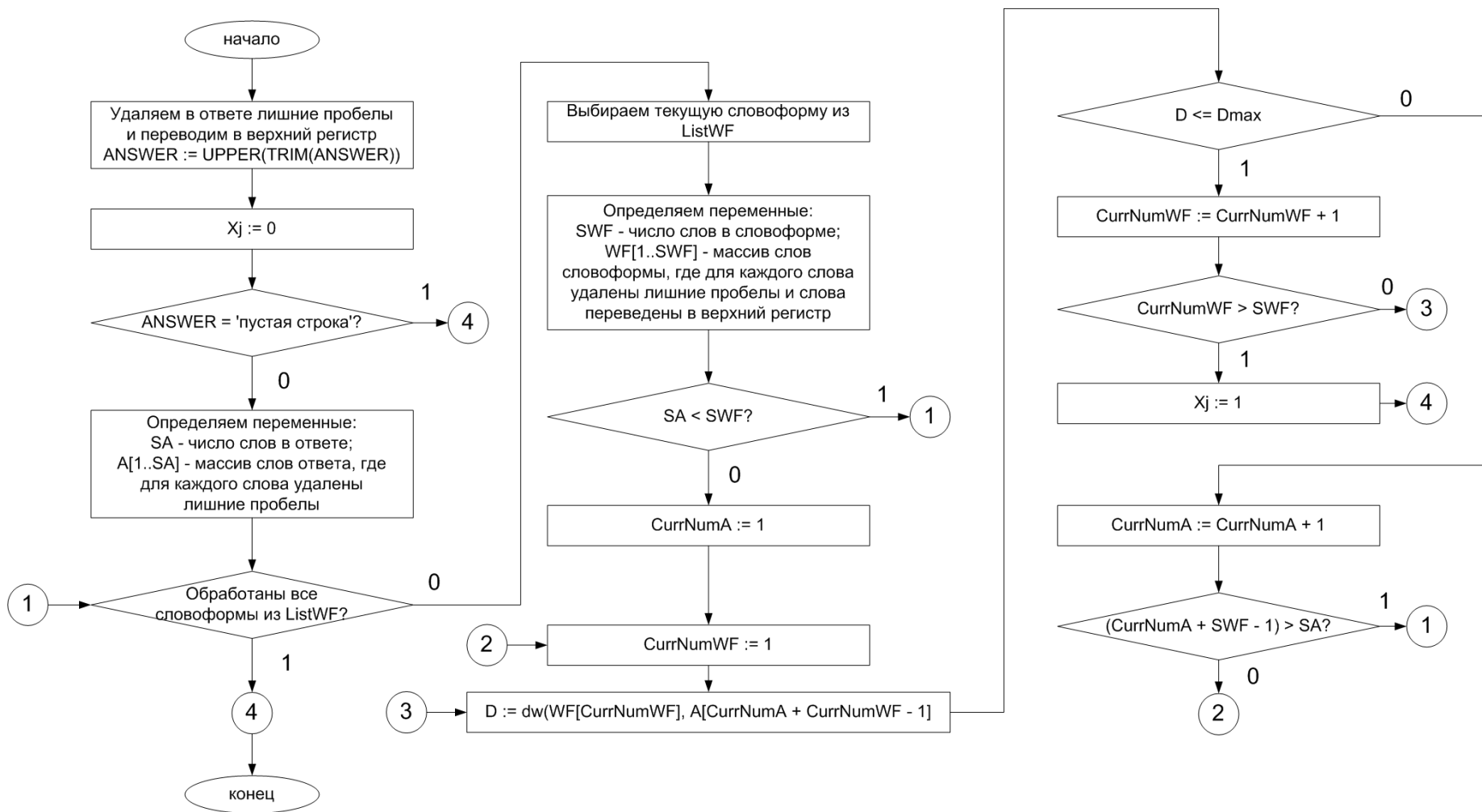


Также разработаны следующие модели и алгоритмы:

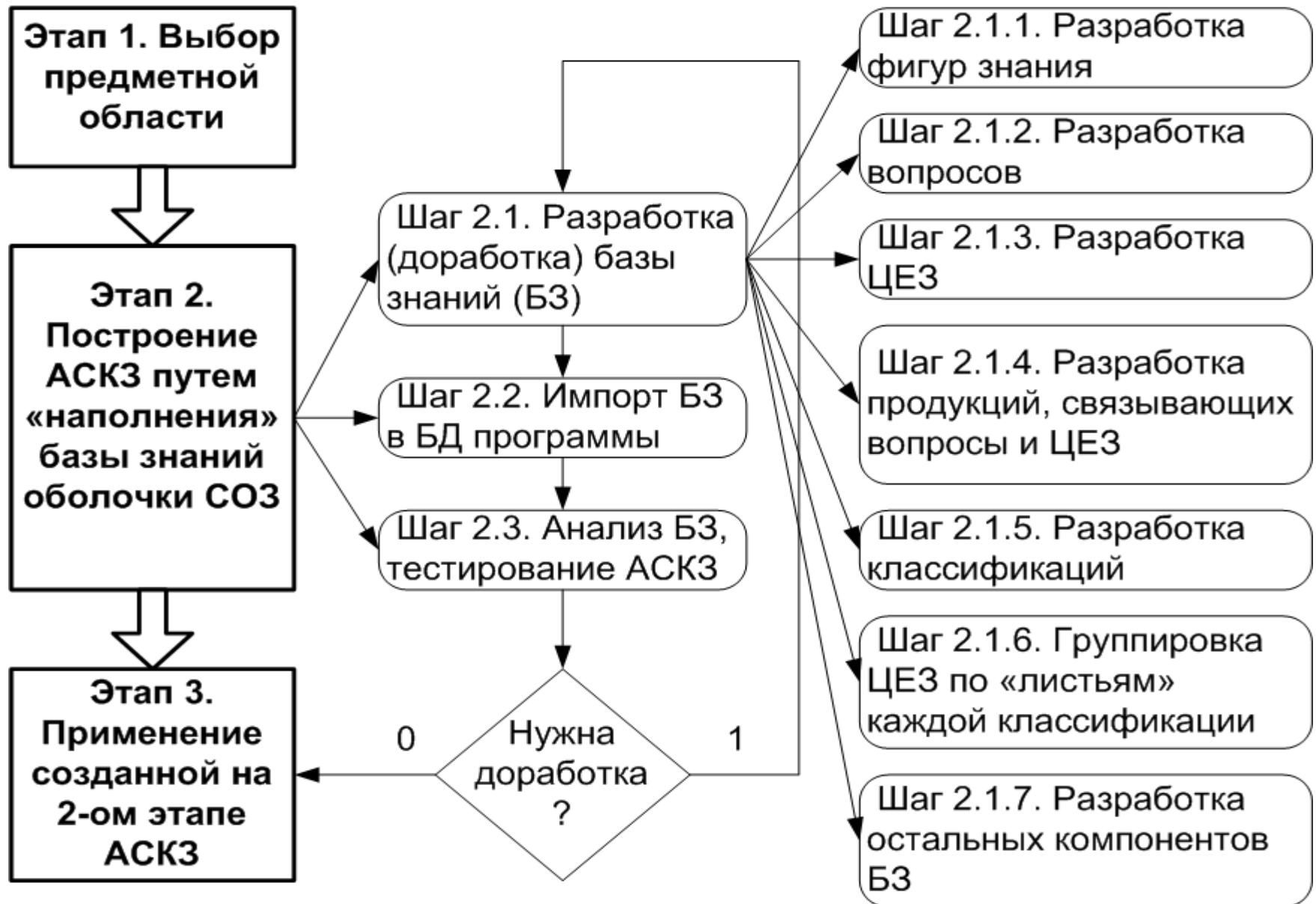
1. Алгоритм анализа ответа на вопрос с учетом опечаток – основан на вычислении взвешенного расстояния между строками с помощью алгоритма Вагнера-Фишера.
2. Модели отбора подмножеств вопросов. Предлагается 3 варианта решения задачи формирования подмножества вопросов: 1) Подмножество непосредственно задается экспертом; 2) Отбор подмножества происходит на основе некоторой классификации предметной области и подмножества КЕ этой классификации; 3) Отбор подмножества происходит в результате решения задачи оптимизации, которая формализована в виде задачи линейного программирования с булевыми переменными.
3. Модель вычисления различных интегральных числовых характеристик базы знаний о предметной области.
4. Модель уменьшения количества ЦЕЗ, имеющих статус «противоречие», на основе решения задачи многокритериальной порядковой экспертной классификации.
5. Инфологическая и даталогическая модели базы данных.
6. Алгоритмы импорта основного объема описания базы знаний предметной области в БД из файлов MS Word.

Алгоритм анализа ответа на вопрос

$$d_w(x_1, x_2) = \min(DelCount + InsCount + \sum_{i=1}^{SubsCount} w_i)$$



Методика АКЗ



Достоинства и недостатки предлагаемой методики АКЗ

Достоинства:

- Возможность построения АСКЗ для различных предметных областей.
- Возможность моделирования в системе предметной области как таковой, т.е. создания базы знаний, представляющей собой особый тип тезауруса, энциклопедии данной области.
- Более высокий уровень детализации описания декларативной компоненты знаний обучаемого (подобласть → понятия → отношения между понятиями → аспекты смысла отношений между понятиями).
- Возможность применения системы без ее апробации на достаточно больших репрезентативных выборках обучаемых.
- Использование системой открытой формы заданий для проверки знаний, не предполагающей реализации сложных процедур семантического анализа естественного языка.
- Возможность объяснить пользователю ход рассуждений, в результате которых была сформирована модель его знаний.
- Реализация в АСКЗ различных видов анализа персональных знаний пользователя.
- Ориентация системы на постоянное взаимодействие с пользователем в течение продолжительного периода времени, а не на один сеанс взаимодействия.

Недостатки:

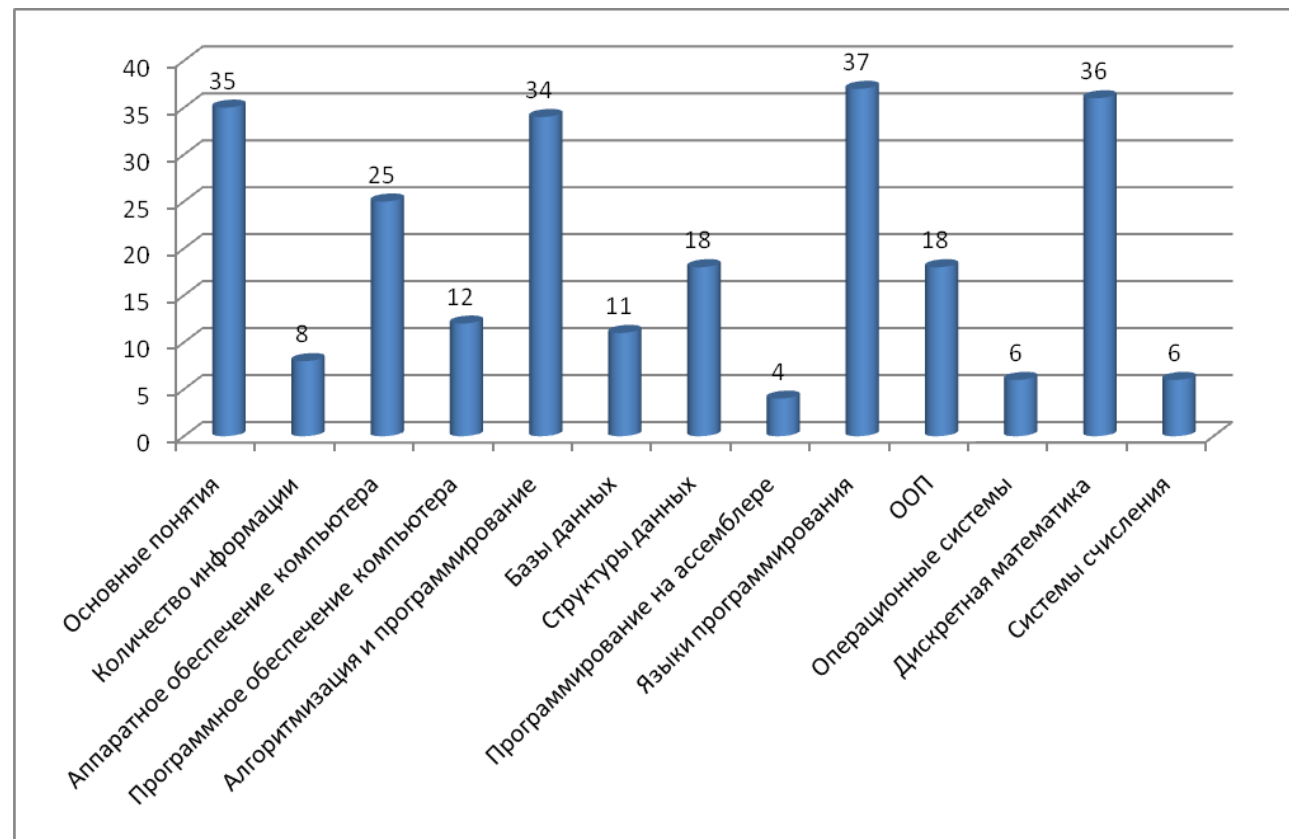
- Использование только одной формы заданий (вопросов).
- Проверка преимущественно декларативной, а не процедурной компоненты знаний обучаемого.
- Высокая трудоемкость разработки базы знаний предметной области.

База знаний предметной области «Основы информатики»

Значения некоторых параметров:

№	Название параметра	Значение
1	Число фигур знания (вопросов)	250
2	Число элементов модели знания	431
3	Число ЦЕЗ	494
4	Число связей «Фигура знания – ЦЕЗ»	1146
5	Среднее число вопросов на одну ЦЕЗ	1,295

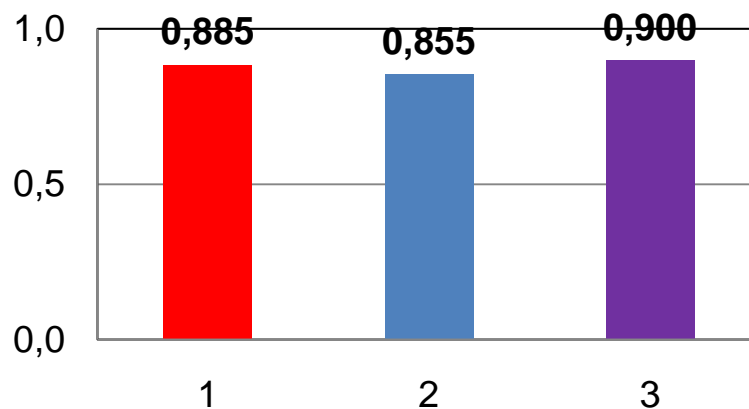
Распределение вопросов по когнитивным областям:



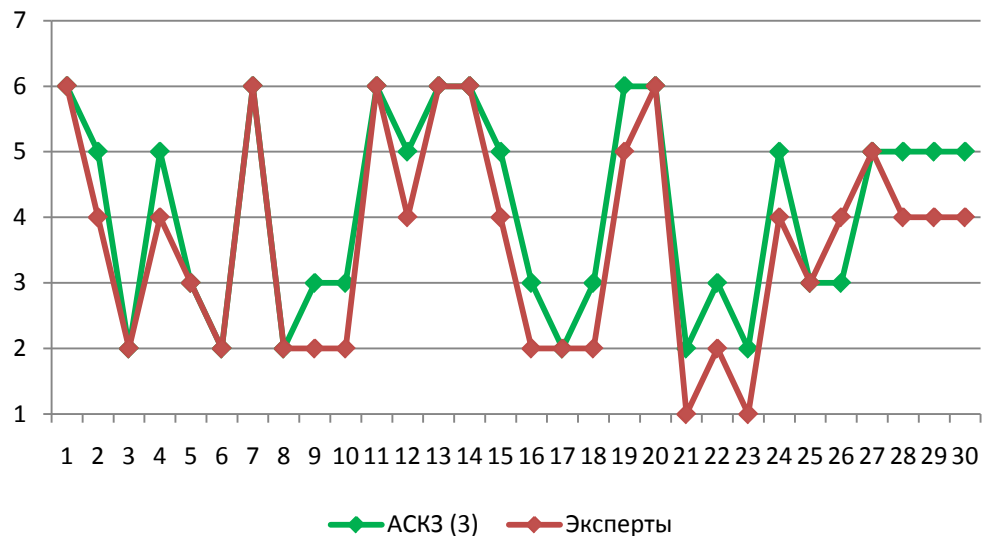
Экспериментальное исследование валидности выявления знаний

№	N(анк.)	Нпр.	C1	C2	C3	Э1	Э2	Э3	Эср.
1	1	14	5	6	6	6	6	6	6
2	2	12	5	6	5	5	4	4	4
3	3	3	2	2	2	2	2	2	2
4	6	10	4	4	5	4	4	4	4
5	8	8	4	4	3	3	3	2	3
6	9	5	3	2	2	2	2	2	2
7	13	17	6	7	6	6	7	6	6
8	16	2	2	2	2	2	2	2	2
9	17	6	3	3	3	3	2	2	2
10	18	6	3	3	3	3	2	2	2
11	22	14	5	6	6	6	6	6	6
12	26	12	5	6	5	5	4	4	4
13	28	17	7	7	6	6	7	6	6
14	32	16	6	7	6	6	6	6	6
15	35	10	4	5	5	4	4	4	4
16	44	5	3	3	3	2	2	2	2
17	48	4	3	2	2	2	2	2	2
18	51	4	3	2	3	2	2	2	2
19	53	13	5	6	6	5	5	4	5
20	55	15	6	6	6	6	6	6	6
21	56	1	1	1	2	2	1	1	1
22	57	7	3	3	3	3	2	2	2
23	64	1	2	2	2	2	1	1	1
24	66	9	5	5	5	4	4	3	4
25	67	7	3	2	3	3	3	2	3
26	69	8	4	4	3	4	4	3	4
27	102	13	5	6	5	5	6	5	5
28	201	11	5	5	5	4	4	4	4
29	204	9	4	4	5	4	4	3	4
30	302	11	4	4	5	4	4	4	4

Значения коэффициента ранговой корреляции Кендалла для 3-х вариантов реализации операции «свертки»:



Решения экспертов и АСКЗ для варианта 3:



Экспериментальное исследование механизма адаптации (1)

«Виртуальная» конфигурация:

$$X_i^{virt} = \{V_i^{250}, \tilde{R}_{d(virt)_i}^{3H.}, \tilde{R}_{d(virt)_i}^{HE3H.}\}$$

Критерии, потенциально влияющие на эффективность адаптации:

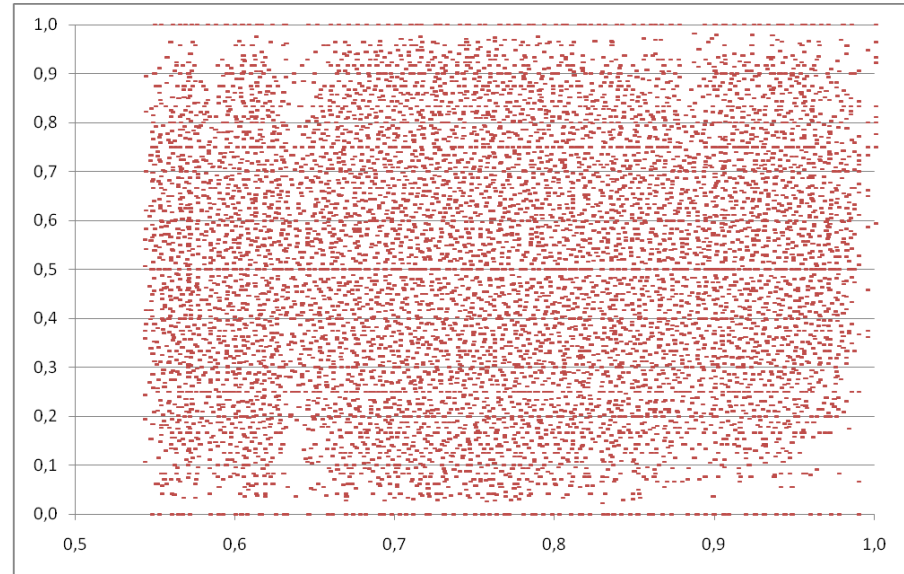
$$K_i^1 = \frac{\sum_{c_j \in C_d^{нижн.}} d_{ij}^*}{K} \quad K_i^2 = \frac{\sum_{\Psi \in \{3H., HE3H.\} \text{ IsCorrectRule}_i^\Psi(j,k)=True} \sum A_i^\Psi(j,k)}{\sum_{\Psi \in \{3H., HE3H.\} \text{ IsRule}_i^\Psi(j,k)=True} \sum A_i^\Psi(j,k)}$$

Обобщенная схема алгоритма случайной генерации «виртуальных» конфигураций:

```

for K1Value := 0 to 10 do
begin
  for KEConfIndex := 0 to (2^K - 1) do
  begin
    генерация  $\tilde{R}_{d(virt)_i}^{3H.}, \tilde{R}_{d(virt)_i}^{HE3H.}$ ;
    перевод KEConfIndex в двоичное представление;
    генерация  $X_i^{virt}$  на основе
    двоичного представления KEConfIndex;
    вычисление координат точки  $(K_i^1, K_i^2)$ ,
    соответствующей  $X_i^{virt}$ ;
  end
end
    
```

Пример множества точек (K_i^1, K_i^2) для 7387 «виртуальных» конфигураций:

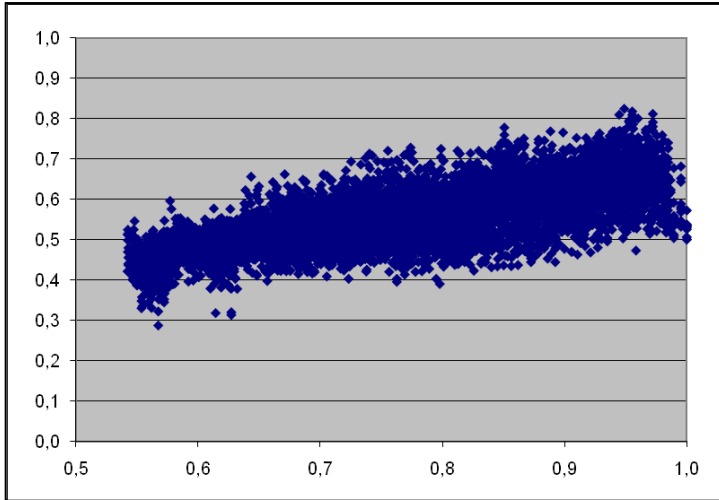


Общее число запусков алгоритма адаптации, выполненных в ходе проведения эксперимента: **884940.**

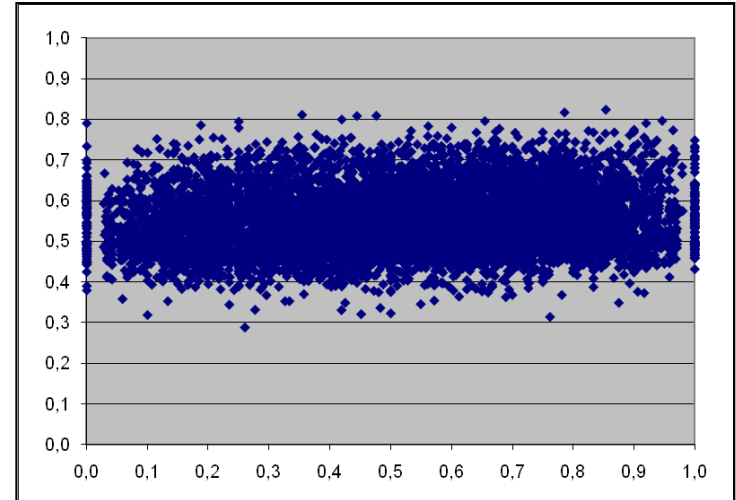
Общее число вопросов, заданных «виртуальным» субъектам в ходе проведения эксперимента: **22123500.**

Экспериментальное исследование механизма адаптации (2)

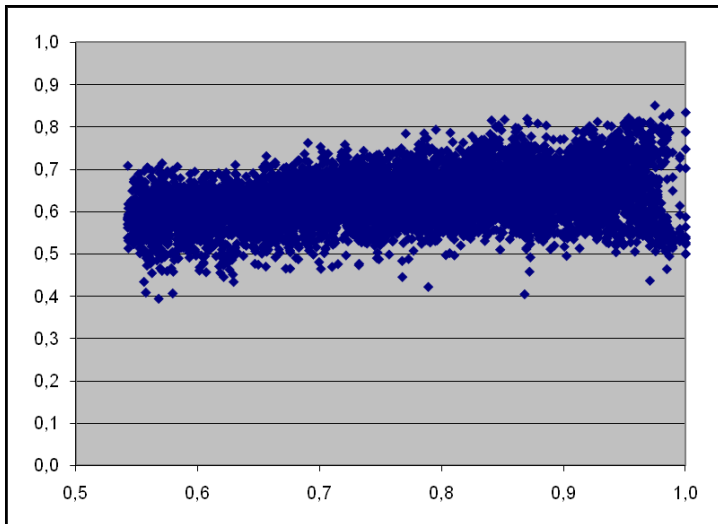
Зависимость $K_{ML} = f(K_i^1)$, 7352 конфигурации



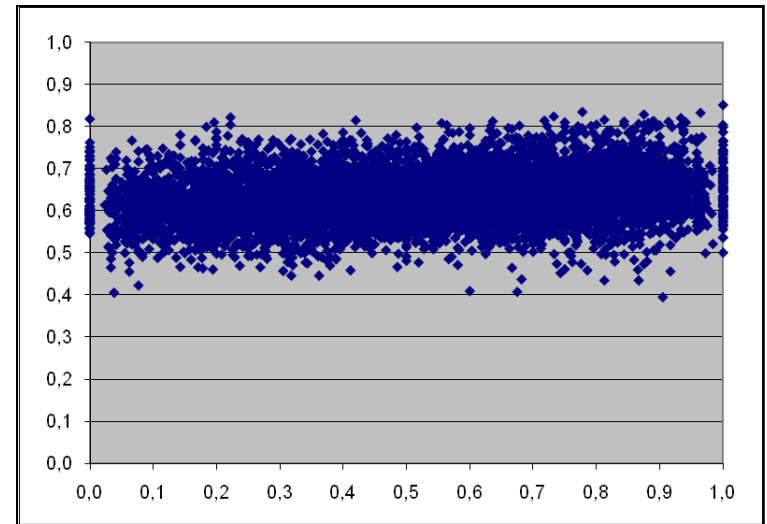
Зависимость $K_{ML} = f(K_i^2)$, 7352 конфигурации



Зависимость $K_{MC} = f(K_i^1)$, 7397 конфигураций



Зависимость $K_{MC} = f(K_i^2)$, 7397 конфигураций



Научная новизна

Научная новизна исследования заключается в следующем:

1. Разработана новая теоретико-множественная модель описания экспертных знаний о предметной области для решения задачи АКЗ, отличающаяся от известных использованием элементарных единиц вербального языкового знания.
2. Разработана новая теоретико-множественная модель описания персональных знаний обучаемого о предметной области для решения задачи АКЗ, отличающаяся от известных использованием целевых единиц знания, а также выделением трех уровней описания, соответствующих трем различным типам анализа персональных знаний обучаемого – структурному анализу, логическому анализу, и анализу соответствия эталонным образам знания.
3. Разработаны новые алгоритмы АКЗ, позволяющие формировать описание персональных знаний обучаемого, исходя из взаимодействия с ним, и основанные на использовании различных способов вывода. Отличительной особенностью этих алгоритмов является обработка знаний, формальное описание которых дано в соответствии с предлагаемыми теоретико-множественными моделями, а также применение четырехзначной логики аргументации.
4. На основе предлагаемых моделей и алгоритмов разработана новая методика АКЗ.

Практическая ценность работы

- Разработанный программный продукт, который является оболочкой системы, основанной на знаниях, позволяет создавать АСКЗ в различных предметных областях путем «наполнения» базы знаний системы.
- Разработана база знаний предметной области «Основы информатики», и АСКЗ на основе этой базы знаний.
- Материалы проведенного исследования были использованы в учебном процессе кафедры «Автоматизированные системы обработки информации и управления» МГТУ им. Баумана, в рамках дисциплин «Семиотика информационных технологий» и «Компьютерные технологии в науке и образовании».

Результаты работы

1. Проведены анализ и классификация существующих методов и моделей АКЗ, а также моделей представления знаний и моделей рассуждений, используемых для решения задачи АКЗ.
2. Разработана новая теоретико-множественная модель описания экспертных знаний о предметной области для решения задачи АКЗ.
3. Разработана новая теоретико-множественная модель описания персональных знаний обучаемого о предметной области для решения задачи АКЗ.
4. Разработаны новые алгоритмы АКЗ, основанные на использовании различных способов вывода на знаниях.
5. На основе предлагаемых моделей и алгоритмов разработана новая методика АКЗ.
6. Разработана и апробирована программная система, реализующая предлагаемую методику АКЗ.
7. Проведены экспериментальные исследования предлагаемой методики АКЗ, которые подтвердили валидность выявления знаний и эффективность предложенного механизма адаптации.



Спасибо за внимание!