

УДК 621.372-501.12.001

В.П.Шибицкий

СТРУКТУРНЫЙ МЕТОД АНАЛИЗА ИНФОРМАЦИОННЫХ СИСТЕМ ОБУЧЕНИЯ С НЕЧЕТКО ЗАДАННОЙ ЦЕЛЕВОЙ ФУНКЦИЕЙ

Рассмотрены проблемные вопросы структурного анализа информационных процессов в слабоформализованной среде обучения. Предложена методика формирования управляющих воздействий в эргатической системе обучения.

Одним из аспектов проблемы активизации информационной среды в эргатической системе обучения является формализация процессов представления, декларирования и приобретения знаний оператором.

Развитие информационных технологий и рост производительности компьютерных систем предопределили совершенствование структурных форм представления информации [1,2,3], для которых характерен высокий уровень абстрагирования. Представление знаний как методология моделирования и формализации концептуальных знаний, ориентированная на компьютерную обработку, является одним из важнейших разделов инженерии знаний. Проанализируем модели представления знаний, используемые для построения эргатических систем обучения.

Логическая модель используется для представления знаний в системе логики предикатов первого порядка и формирования заключения с помощью силлогизма. Такая модель позволяет реализовать систему формально точных определений и выводов.

В продукционных моделях знания представлены совокупностью правил вида “если – то”. Эти модели бывают с прямыми и обратными выводами и включают три компонента: базу правил, состоящую из набора продукций (правил вывода); базу данных, содержащую множество фактов; интерпретатор для получения логического вывода на основании этих знаний. База правил и база данных образуют базу знаний, а интерпретатор соответствует механизму логического вывода.

Семантические сети представляют собой систему знаний в виде целостного образа сети, узлы которой соответствуют смысловым понятиям в форме объектов, а дуги сетевого графа - отношениям между объектами. Изначально семантическая сеть была задумана в психологии как модель представления структуры долговременной памяти. Однако с увеличением размеров сети существенно увеличивается время на обработку элементов знаний. Шагом на пути формализации представления знаний явилось создание фреймовой теории.

Фрейм - это структурная единица представления знаний, запомненная в прошлом, детали которой при необходимости могут быть изменены согласно текущей ситуации. Фреймовая система - это иерархическая структура фреймов, каждый из которых состоит из произвольного числа слотов, представленных определенной структурой данных, кратко описываемой в виде схемы “объект-атрибут-значение”.

Развитие теории интеллектуальных систем [4] и методов программирования привело к созданию нового объектного типа данных в языках программирования: C++, Object Pascal, Delphi.

Объектно-ориентированный способ представления дидактических знаний позволяет построить иерархическую модель предметной области путем объединения объектов в классы и выявления между ними отношений (табл.1).

Таблица 1

Объектно-ориентированный способ представления знаний

Принцип	Сущность	Примечание
Инкапсуляция	Моделирование динамики состояний объекта за счет организации доступа к данным и методам	Доступ к полям и методам объекта сосредоточен внутри класса
Наследование	Организация системы классов в виде иерархии объектов	Класс-потомок связан отношением с классом-предком и наследует поля данных и методы предка
Полиморфизм	Переопределение методов обработки полей данных объекта-предка	Если реализация метода класса-предка не устраивает разработчика, то можно переопределить этот метод, сохраняя при этом доступ к классу-предку

Информационный объект $x_i \in X$ - это абстрактная структурная форма представления смысловой единицы знаний (рис. 1, а и б), которая в отличие от фрейма включает формальные процедуры (методы) обработки данных объекта.

Динамическое поведение информационного объекта в системе обусловлено:

- отношениями в заданной структуре информационных объектов;
- состоянием объекта в результате выполнения процедур;
- временем активного состояния объекта в процессе обучения.

Статические характеристики информационного объекта (рис. 1, а) определяются по значениям следующих параметров:

- c - собственной информационной емкости;
- n - степени сложности или насыщения объекта новыми понятиями;
- w - относительного веса объекта в структуре дидактических знаний.

Задача декларирования знаний сводится к формированию последовательности объектов (рис. 1, а) в дискретные моменты времени. В системе обучения возникает необходимость обобщения дискретной последовательности объектов в абстрактную структурную форму, удовлетворяющую дидактическим принципам связности и полноты знаний в непрерывном процессе (рис. 1, б) усвоения их оператором.

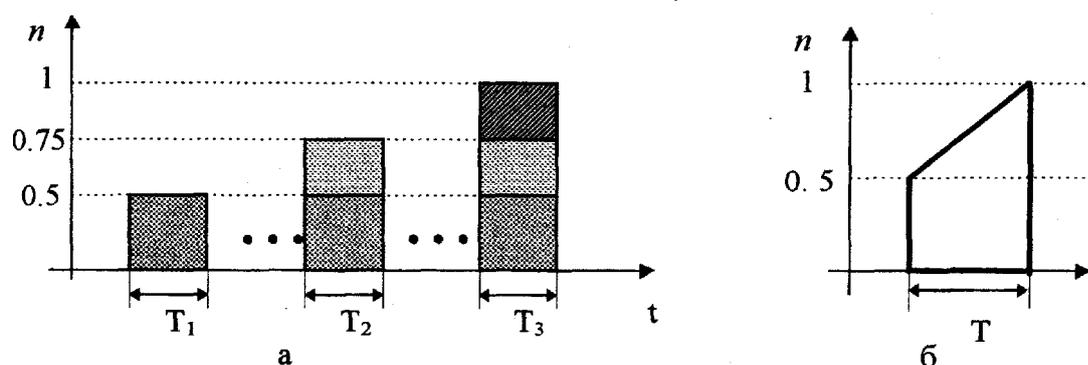


Рис. 1

Теоретико-множественная модель структуры (рис.2, а и б) формально объединяет множество объектов с учетом иерархии подмножеств X_1, X_2, X_3 множества X . Эта модель позволяет оценить связность объектов и формировать их классы [2].

Классом будем считать подмножество $X_i \subset X$ с элементами, удовлетворяющими критерию эквивалентности $\xi_i \in \{\xi_i\}$; для класса X_i операция $X_i \cap (X \setminus X_i) = \emptyset$.

Проблемным вопросом анализа и синтеза структур дидактических знаний является разработка модели нечеткой структуры на основе понятий множества и отношения.

Нечеткую модель структуры (рис.2, а) дидактических знаний представим в виде бинарного отношения на множестве $X_x = X \times X$ и множестве инцидентов V (рис. 2, б).

Определение. Нечеткой моделью H структуры дидактических знаний назовем бинарное отношение

$$H \subset X_x^* \times V$$

с областью определения на нечетко заданном множестве X_x^* и такое, что для элементов $\{x_i, x_j\}$, $i \neq j$ декартового произведения X_x нечеткого подмножества $X_x^* \subset X_x$ определена функция принадлежности:

$$\mu_{X_x^*}(x_i, x_j) = \begin{cases} 1, & \text{если для } x_i, x_j \text{ отношение } v_i \in V \text{ найдено;} \\ 0, & \text{если для } x_i, x_j \text{ отношение } v_i \in V \text{ не найдено.} \end{cases}$$

Функция принадлежности элементов базового множества X_x нечеткому подмножеству $X_x^* \subset X_x$ позволяет сформировать подмножество пар объектов, в котором любой объект может иметь определенное отношение хотя бы с одним из других объектов.

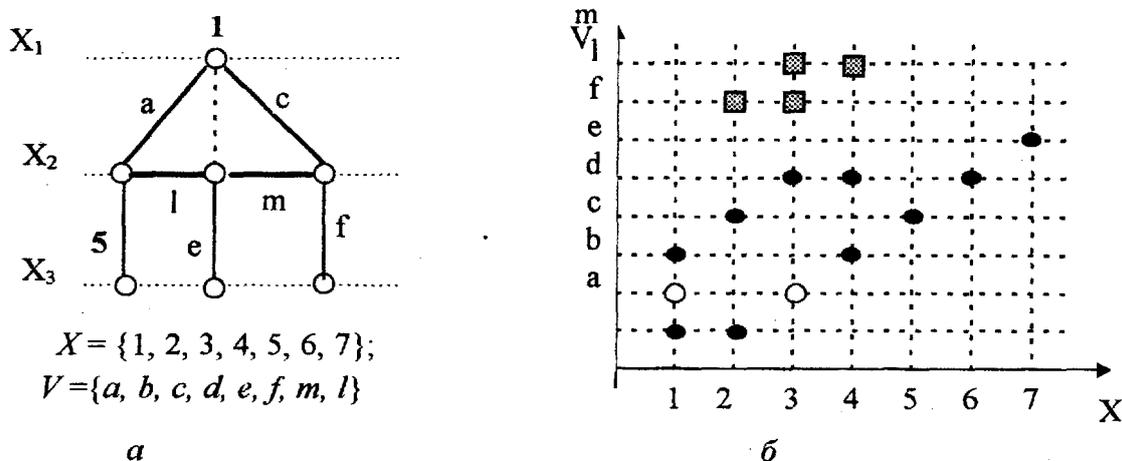


Рис.2

Для управления процессом обучения необходимо соблюдение принципа полноты дидактических знаний. Для его реализации требуется фактическое обеспечение всех составляющих компонент (умений) профессиональной деятельности. Полноту множества X знаний, обеспечивающих компоненты деятельности $k_i \in K$, оцениваем по результату деятельности. Удовлетворительный результат деятельности означает, что компоненту деятельности (умению) $k_i = P(x_i)$ обязательно соответствует элемент множества знаний.

Отношение U частичного соответствия компонента деятельности $k_i = U(x_i)$ с элементом знаний предполагает наличие смысловой связи $\{k_i, x_i\}$, но в этом бинарном отношении $k_i \in K$ использует только часть $x_i \in X$ (рис. 3). В случае более сложного отношения M для множеств X и K компонента деятельности определяется как функция $k_i = M(x_1, \dots, x_i, \dots, x_n)$.

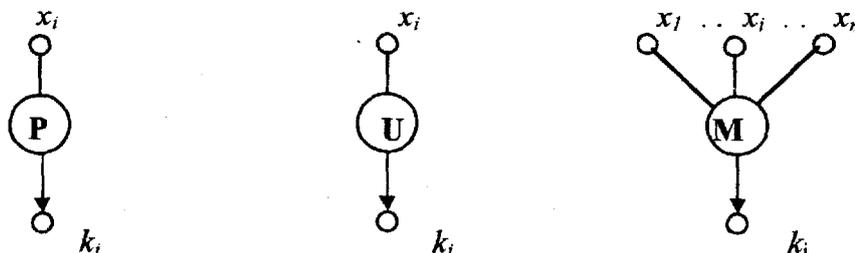


Рис.3

Систему дидактических знаний можно представить в виде множества S , определяемого структурой H , множеством отношений Q , функцией распределения f , и обозначить: $S = (H, f, Q)$. Будем считать, что f - взаимно однозначное отображение Q на V , представляющее собой множество упорядоченных пар (v, q) . Отображение f есть функция распределения элементов $q_i \in Q$ по элементам $v_i \in V$. Аргументы H, f, Q находятся во взаимном отношении, которое для фиксированных H, f, Q определяет одну и только одну систему S дидактических знаний.

На основании предложенной модели структуры дидактических знаний рассмотрим процедуры формирования информационного потока $P \subset H$ с целью управления процессом обучения операторов эргатических систем.

Определение. Элементом $p_i \in P$ потока дидактических знаний назовем последовательность объектов x_i, x_j с экспертно заданной или вычисленной степенью эквивалентности $q_i \in Q$, которая удовлетворяет функции $p_i = q_i(x_i, x_j)$.

Пусть в системе обучения определено базовое множество X и имеется объект управления (оператор), для которого необходимо сформировать подмножество $X_i \subset X$, удовлетворяющее критерию $\xi_i \in \{\xi_i\}$. Для классификации множества X используем функцию принадлежности $\mu_{X_i}(x_i)$ элементов $x_i \in X$ подмножеству X_i .

Экспертно оценим или вычислим степень эквивалентности $q_i \in Q$ информационных объектов в подмножестве X_i и определим значение:

$$\xi_i = \sum_{i=1}^n q_i,$$

где n - число бинарных отношений элементов в X_i .

Условие существования множества Q как покрытия X_i позволяет сформировать информационный поток $P \subset H$ как собственное подмножество дидактических знаний системы. Алгоритм формирования информационного потока P состоит из трех процедур.

Процедура 1. В нечеткой структуре H дидактических знаний и заданного множества $X_i \subset X$ декларируем объекты $x_i \in X$ и формируем базовое множество X_x из элементов $\{x_i, x_j\}$ декартового произведения $X \times X$.

Функция принадлежности $\mu_{X_i}(\{x_i, x_j\})$ элементов $\{x_i, x_j\}$ базового множества X_x нечеткому подмножеству X_x^* фиксирует отношения $q_i \in Q$, отражающие информационную связь объектов x_i, x_j заданного множества $X_i \subset X$.

Процедура 2. Сформированный поток P с элементами $p_i = q_i(x_i, x_j)$ должен совпадать с исходной системой дидактических знаний на экспертно заданном множестве $X_i \subset X$ объектов:

$$\bigcup_{i=1}^n (x_i, x_j)_i = X_i,$$

где n - число элементов потока.

Совпадение сформированного множества P с исходной системой знаний на экспертно заданном множестве X_i предполагает учет дидактического критерия полноты знаний и определяет классификацию множества X .

Далее требуется введение процедуры проверки множества P на информационную связность с целью выявления отношения эквивалентности объектов в заданном множестве X_i и вычисления критерия $\xi_i \in \{\xi_i\}$.

Приведенные ранее процедуры анализа множественной модели дидактических знаний используют бинарное отношение на множестве X , которое является необходимым, но не достаточным условием существования потока P .

Процедура 3. Для экспертно заданного источника x_i в множестве X проверяется отношение $q_i \in Q$ один x_i к одному x_j или один x_i к подмножеству $X_i \subset X$ с элементами, для которых задано бинарное отношение $q_i \in Q$ с источником $x_i \in X$. При этом подмножество $X_i \subset X$ определяет уровень иерархии представления знаний для заданного x_i .

Далее процедура повторяется для источников из множества X_i и проверяются отношения $q_i \in Q$ с подмножеством $X \setminus X_i$ с целью формирования подмножества $X_j \subset X \setminus X_i$. Процедура считается выполненной, если бинарное отношение определено на подмножествах X_i, X_j базового множества информационных объектов.

Структурная сложность знаний обусловлена разнообразием взаимосвязей между элементами содержания обучения и является главным критерием проблемного обучения. Можно выделить следующие структуры дидактических знаний:

- последовательные структуры знаний, предполагающие усвоение темы только в случае усвоения всех входящих в ее состав элементов;
- параллельные структуры, допускающие усвоение темы в случае усвоения хотя бы одного элемента знаний, входящих в ее состав;
- комбинированные структуры, представляющие собой совокупность последовательных и параллельных структур.

В качестве тестовых примеров анализа структур знаний и формирования классов объектов выбираем структуры (табл.2) различной сложности со степенями эквивалентности объектов из множества $Q = \{-1, 0, 1\}$.

Результаты исследования алгоритмов по критериям структурной сложности, определяемой числом классов структур дидактических знаний, и, как следствие, затратами машинного времени приведены в табл.3.

Таблица 2

Тестовые структуры

Номер структуры	Номер элемента потока								
	1	2	3	4	5	6	7	8	9
1	1,2 -1	2,3 0	3,4 0	4,6 -1	1,5 0	3,5 1	4,5 0	6,5 0	3,5 0
2	1,2 -1	2,3 0	3,4 0	4,6 -1	1,5 0	3,5 1	4,5 0	6,5 0	
3	1,2 -1	2,3 0	3,5 -1	1,4 0	2,4 0	2,4 1	3,4 0	5,4 0	
4	1,2 -1	2,3 0	3,4 -1	1,4 0	2,4 1	4,5 0			
5	1,2 -1	2,4 -1	1,3 0	2,3 0	4,3 0	4,3 1			
6	1,2 -1	2,3 0	3,5 -1	1,4 0	2,4 0	4,5 1			

Таблица 3

Классы эквивалентных структур

Номер структуры	Критерий сложности			Класс структуры и критерий эквивалентности				t, с
	n	m	d	-2	-1	0	1	
1	6	9	44	13	19	10	2	0,22
2	6	8	29	7	12	8	2	0,22
3	5	8	31	8	13	8	2	0,19
4	5	6	11	3	4	3	1	0,03
5	4	6	13	3	5	4	1	0,09
6	5	6	11	2	4	4	1	0,19
7	6	11	192	47	48	45	52	0,31
8	8	16	2304	603	561	559	581	0,6
9	10	21	27648	7287	6237	6523	7604	11,37

Управление в системе обучения представляет собой динамический процесс, направленный на достижение цели. Явно заданная цель обучения может быть представлена в виде логического условия, выраженного в терминах модели обучаемого (МО). Под моделью обучаемого будем понимать знания обучающей системы об операторе, используемые для организации процесса обучения. Для достижения цели строится план обучения, представляющий собой последовательность учебных воздействий. Таким образом, модель обучаемого - это модель процесса диалога, которую создает система, интерпретируя способности и уровень понимания обучаемым предметной области.

В системах обучения используются три типа МО: разностная, пертурбационная, оверлейная. Разностная модель строится на основе различий между знаниями эксперта и обучаемого, которые накапливаются в модели пользователя в процессе ведения диалога с системой. В данном случае система обучения должна включать подсистему для объяснения ошибок оператора и формировать контрольные задания с учетом причин этих ошибок, для того чтобы убедиться, что обучаемый усвоил материал.

Для работы с ошибками обучаемых применяются так называемые модели ошибок (bug models или Buggy models), позволяющие определить и отразить причины неверного поведения обучаемых. Наиболее исследованный вид таких моделей - это perturbation models.

Оверлейная модель строится исходя из предположения, что знания оператора и знания системы имеют аналогичную структуру, и первые являются подмножеством последних. Оверлейная модель позволяет не включать в модель оператора модули, которые не были затронуты в теме диалога, что сокращает область и время поиска знаний, необходимых для решения поставленной задачи.

Требуемый уровень знаний - конечная цель Z обучения, является нечетким подмножеством базового множества X знаний системы, где для $x \in X$ задана функция принадлежности $\mu_Z(x)$ (рис. 4). Процесс достижения дидактической цели и формирование нечеткого множества Z определяется достижением подцелей $T_j \in T$.

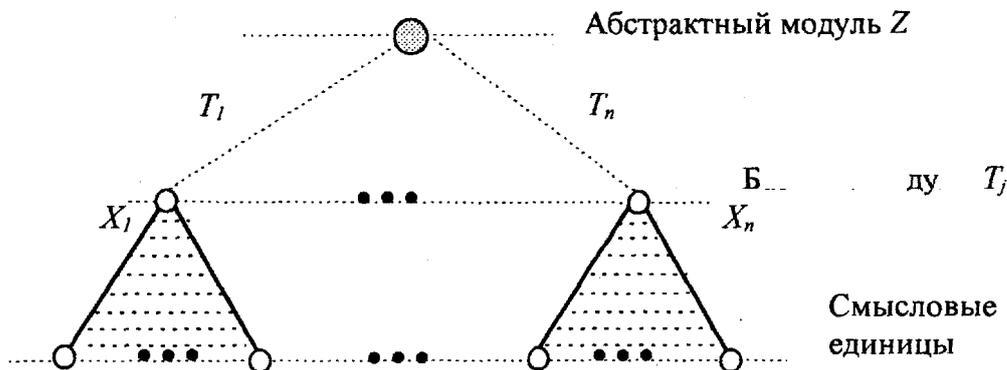


Рис.4

Множество T базовых модулей знаний образует класс целей $T_j \in T$ формирования обучающих воздействий на более низком уровне структурирования знаний на элементы x_1, \dots, x_k , не требующие дальнейшего разбиения, при этом $x_i \in T$, $T \subset X$. Экспертно задавая отношения объектов дидактических знаний на различных уровнях иерархии и проводя тестирование операторов на нижнем уровне иерархии знаний, получаем возможность построить модель обучения и формально оценить уровень обученности оператора в нечеткой структуре дидактических знаний.

В паре объектов $\{x_i, x_j\}_i \subset X_x$ предок-потомок задан порядок отношения наследования $p_i \in P$ элемента $x_i \in X_i$ знаний. Для иерархической структуры H справедливо:

$$\bigcap_{k=1}^m H(x_i, x_j)_k = \emptyset, \quad \bigcup_{k=1}^m (x_i, x_j)_k = X_i.$$

Отсюда следует, что для каждой пары $\{x_i, x_j\}$ элементов дидактических знаний фиксируется только одно отношение $v_i \in V \Leftrightarrow P$. Структуру дидактических знаний можно формально определить через множество пар элементов знаний соседних уровней иерархии, обладающих свойством наследования знаний. Оценивание отношения $p_i \in P$ наследования знаний верхнего уровня иерархии и вычисление значения оценки выполняется с помощью теста на уровне процесса усвоения элементов знаний.

Иерархической структурой назовем бинарное отношение S и такое, что для любых пар y, z элементов знаний справедливо: $S(y) \cap S(z) = \emptyset, S(y), S(z) \subset P$.

Это условие отражает тот факт, что в реальной системе знаний один и тот же параметр оценки наследования может быть присоединен только к одной паре (x_i, x_j) . Формально это условие фиксирует из всех подмножеств такое, которое является математическим аналогом системы знаний S с множеством X элементов и множеством P отношения наследования знаний.

Рассмотрим пример достижения целей обучения, когда умения взаимно однозначно соответствуют элементам знаний, и экспертно оценим данные, полученные в результате тестирования на нижнем уровне структуры знаний.

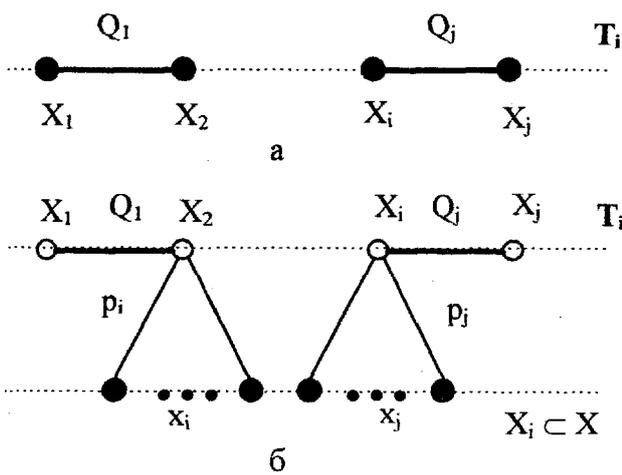


Рис. 5

В результате экспертная структура (рис. 5, а) настраивается на уровень знаний оператора (рис. 5, б) с новым структурным критерием эквивалентности для подмножества информационных объектов.

Предположим, что система обучения декларирует поток знаний в виде класса $X_i \subset X$ с заданными структурными свойствами (рис. 5, а). Для оценивания процесса усвоения потока дидактических знаний эргатическая система обучения располагает системой тестов. Если после анализа результатов тестирования принимается решение о недостаточном усвоении модуля X_i , то в этом случае система обучения декларирует информационный поток дидактических знаний на более низком уровне (рис. 5, б) с учетом экспертно заданных параметров наследования знаний $X_i \subset X$.

Список литературы

1. Шибизкий В.П., Шибизкая Н.Н., Савчук Н.Н. Структурный анализ интеллектуальных систем управления с декомпозицией и распараллеливанием вычислительных процессов // Вісник КМУЦА. – К. : КМУЦА, 1999. – № 2. – С.179-187.
2. Волков А.А. Моделирование систем графами // Вісник КМУЦА. – К. : КМУЦА, 1998. – № 1. – С.268-274.
3. Минасв Ю.М., Філімонова О.Ю. Операції нечіткої математики в складі інформаційних технологій на платформі Windows // Вісник КМУЦА. – К. : КМУЦА, 1998. – № 1. – С.245-252.
4. Нечеткие множества и теория возможностей. Последние достижения / Пер. с англ; Под ред. Р.Р.Ягера.-М.: Радио и связь, 1986. – 408 с.