

РЕАЛИЗАЦИЯ ПРИНЦИПА ТРАНСПАРЕНТНОСТИ ПРИ ПОСТРОЕНИИ УЧЕБНОЙ ЭКСПЕРТНОЙ СИСТЕМЫ

© 2018 И. Ю. Пикалов¹, В. Б. Тарасюк², Е. И. Травкин³

¹ канд. пед. наук, доцент, доцент кафедры компьютерных технологий
и информатизации образования
e-mail: pikalov@kursksu.ru

² канд. техн. наук, доцент, доцент кафедры компьютерных технологий
и информатизации образования
e-mail: tarasuk.vladimir@mail.ru

³ канд. пед. наук, доцент кафедры компьютерных технологий
и информатизации образования
e-mail: etravkin@mail.ru

Курский государственный университет

В статье представлена учебная модель экспертной системы, реализованная в среде MS Excel, позволяющая детально проследить за процессом обучения экспертной системы. Приведены и детально представлены механизмы формирования вектора решений и рассчитываемых идентификаторов распознаваемых сущностей.

Ключевые слова: учебная модель, экспертная система, распознавание сущности, вектор решений, идентификатор сущности, обучение экспертной системы.

Современный уровень развития информационного общества характеризуется широким использованием интеллектуальных систем. Под таковыми понимаются технические или программные системы, способные решать задачи, традиционно считающиеся творческими, принадлежащие конкретной предметной области, знания о которой хранятся в памяти такой системы. Структура интеллектуальной системы включает три основных блока — базу знаний, механизм вывода решений и интеллектуальный интерфейс [Головицына 2008]. К этой же группе относятся системы распознавания образов. Многие из них опираются на модель экспертной системы. Под экспертной системой будем понимать программу (или группу программ) для компьютера, которая, используя заложенные в неё знания, предлагает решение поставленной задачи или вырабатывает рекомендации по её решению. [Пикалов, Тарасюк 2017]. Отличительной особенностью интеллектуальных систем является их способность к обучению или самообучению. Организация этого процесса может осуществляться различными способами и базироваться на многообразии стратегий и алгоритмов.

На начальном этапе изучения принципов реализации данной способности экспертных систем, нейронных сетей и других систем искусственного интеллекта возникают известные трудности: сложность понимания, скрытость процессов, отсутствие четких алгоритмов функционирования и построения таких систем. Это диктует необходимость реализации процесса обучения на элементарных моделях, базирующихся на простейших алгоритмах, способных наглядно демонстрировать инициализацию обучающего модуля экспертной системы, формирование базы знаний и механизма вывода решений.

В качестве основы для реализации принципа транспарентности воспользуемся моделью экспертной системы и алгоритмом построения обучающего модуля [Пикалов, Тарасюк 2017].

Немаловажное значение в реализации принципа транспарентности имеет адекватный выбор инструмента для построения учебной экспертной системы. Этот инструмент должен обеспечивать возможность построения экспертной системы непрограммирующим пользователям, визуализацию базы знаний и механизма принятия решений; удобный и интуитивно-понятный интерфейс. Наиболее популярным инструментом, соответствующим данным требованиям, является MS Excel. Именно он будет использован при построении транспарентной учебной системы.

Для простоты понимания излагаемой последовательности шагов рассмотрим данный процесс на конкретном примере распознавания сущностей, фиксируемых какой-либо системой наблюдения. В качестве исходных сущностей выбираем следующие объекты: человек, собака, автомобиль. Признаки, характеризующие данные сущности, условимся называть атрибутами и обозначим их переменными V_{ij} , где i — номер атрибута, j — номер сущности. При этом будет использоваться двухуровневое ранжирование. Наличие атрибута у сущности будет соответствовать единице, а его отсутствие — нулю. Установим следующий набор атрибутов для рассматриваемого примера: двуногое, колеса, четырехное, голова, хвост, выхлопная труба.

На начальном этапе строим таблицу исходных данных, отражающих значения атрибутов для распознаваемых сущностей, представленную на рисунке 1.

	A	B	C	D	E	F	G	H
1		Имя сущности	Двуногое	Колеса	Четырехное	Голова	Хвост	Выхлопная труба
2	1	Человек	1	0	0	1	0	0
3	2	Собака	0	0	1	1	1	0
4	3	Автомобиль	0	1	0	0	0	1

Рис. 1. Значения атрибутов для распознаваемых сущностей (последовательность 1)

Следующий этап формирования экспертной системы предполагает вычисление вектора решений R . Первичное значение вектора решений $R(1)$ определяется значениями атрибутов первой сущности. В данном случае

$$R(1) = (R_1, R_2, R_3, R_4, R_5, R_6) = (V_{11}, V_{21}, V_{31}, V_{41}, V_{51}, V_{61}) = (1, 0, 0, 1, 0, 0).$$

Размещение первичного значения вектора решений на рабочем листе книги показано на рисунке 2.

	A	B	C	D	E	F	G	H
		Имя	Другое	Колеса	Четырехное	Голова	Хвост	Выхлопная труба
1	№	сущности						
2	1	Человек	1	0	0	1	0	0
3	2	Собака	0	0	1	1	1	0
4	3	Автомобиль	0	1	0	0	0	1
5								
6		R(1)	1	0	0	1	0	0
7								

Рис. 2. Первичное значение вектора решений (последовательность 1)

Следующим этапом является вычисление идентификаторов сущностей (id_1 , id_2 , id_3) по следующей формуле:

$$id_j = \sum_{i=1}^6 (R_i \times V_{ij}).$$

В соответствии с синтаксисом MS Excel формула для каждой сущности принимает вид, представленный на рисунке 3.

	J	K	L	M	N	O	P	Q
1								
2		$R_1 \times V_{1j}$	$R_2 \times V_{2j}$	$R_3 \times V_{3j}$	$R_4 \times V_{4j}$	$R_5 \times V_{5j}$	$R_6 \times V_{6j}$	id_j
3								
4		$=\$C\$6*C2$	$=\$D\$6*D2$	$=\$E\$6*E2$	$=\$F\$6*F2$	$=\$G\$6*G2$	$=\$H\$6*H2$	$=CYMM(K4:P4)$
5		$=\$C\$6*C3$	$=\$D\$6*D3$	$=\$E\$6*E3$	$=\$F\$6*F3$	$=\$G\$6*G3$	$=\$H\$6*H3$	$=CYMM(K5:P5)$
6		$=\$C\$6*C4$	$=\$D\$6*D4$	$=\$E\$6*E4$	$=\$F\$6*F4$	$=\$G\$6*G4$	$=\$H\$6*H4$	$=CYMM(K6:P6)$

Рис. 3. Представление идентификаторов сущностей в соответствии с синтаксисом MS Excel (последовательность 1)

На основании расчетов получаем следующие значения id , представленные на рисунке 4: $id_1 = 2$, $id_2 = 1$, $id_3 = 0$.

	J	K	L	M	N	O	P	Q
1								
2		$R_1 \times V_{1j}$	$R_2 \times V_{2j}$	$R_3 \times V_{3j}$	$R_4 \times V_{4j}$	$R_5 \times V_{5j}$	$R_6 \times V_{6j}$	id_j
3								
4		1	0	0	1	0	0	2
5		0	0	0	1	0	0	1
6		0	0	0	0	0	0	0

Рис. 4. Расчет идентификаторов сущностей в MS Excel (последовательность 1)

Отсутствие совпадений id для разных сущностей показывает готовность системы к распознаванию, т.е. окончание сеанса обучения. Используемый в примере набор атрибутов и порядок сущностей позволил обучить систему за один проход, что возможно не всегда. Рассмотрим ещё один случай размещения сущностей в списке: автомобиль, собака, человек. Для измененной последовательности (последовательность 2) строим таблицу исходных данных, отражающих значения атрибутов для распознаваемых сущностей, представленную на рисунке 5.

	A	B	C	D	E	F	G	H
1	№	Имя сущности	Двуное	Колеса	Четырёхное	Голова	Хвост	Выхлопная труба
2	1	Автомобиль	0	1	0	0	0	1
3	2	Собака	0	0	1	1	1	0
4	3	Человек	1	0	0	1	0	0

Рис. 5. Значения атрибутов для распознаваемых сущностей (последовательность 2)

Для данной последовательности сущностей первичное значение вектора решений определяется следующим образом:

$$R(1) = (R_1, R_2, R_3, R_4, R_5, R_6) = (V_{11}, V_{21}, V_{31}, V_{41}, V_{51}, V_{61}) = (0, 1, 0, 0, 0, 1).$$

Размещение первичного значения вектора решений на рабочем листе книги для рассматриваемой последовательности показано на рисунке 6.

	L	M	N	O	P	Q	R	S	T	U	V	W
1												
2	Первый цикл обучения											
3												
4	R(1)	0	1	0	0	0	0	1		id:	2	
5										id:	0	
6	R(2)	1	1	0	1	0	1			id:	0	
7	R(3)	2	1	0	2	0	1			id:	2	
8										id:	4	
9	Второй цикл обучения											
10												
11										id:	2	
12										id:	2	
13	R(4)	2	1	1	3	1	1			id:	5	
14										id:	5	
15	R(5)	3	1	1	4	1	1			id:	5	
16										id:	7	
17	Третий цикл обучения											
18												
19										id:	2	
20										id:	6	
21										id:	7	
22												
23												

Рис. 6. Визуализация процесса обучения системы для второй последовательности сущностей (последовательность 2)

Используя формулу вычисления идентификаторов сущностей, производим их расчет. Рисунок 6, отражающий результаты расчетов, демонстрирует наличие совпадений значений идентификаторов при первичном определении вектора решений: $id_1 = 2$, $id_2 = 0$, $id_3 = 0$. Поэтому, следуя алгоритму [Пикалов, Тарасюк 2017], необходимо провести переопределение вектора решений. Подобная ситуация будет повторяться четыре раза. В итоге процесс обучения экспертной системы завершится на третьем цикле обучения (в этом цикле не происходило переопределение вектора решений и все рассчитанные идентификаторы получились разными) при значениях идентификаторов сущностей: $id_1 = 2$, $id_2 = 6$, $id_3 = 7$. Значение вектора решений при этом будет иметь следующий вид: $R(5) = (3, 1, 1, 4, 1, 1)$.

В соответствии с синтаксисом MS Excel расчеты вектора решений и идентификаторов сущностей представлены на рисунке 7.

	L	M	N	O	P	Q	R	S	T	U	V
1											
2	Первый цикл обучения										
3											
4	R(1)	=C2	=D2	=E2	=F2	=G2	=H2			id1	=СУММ(C2:H2*N4:S4)
5										id2	=СУММ(C3:H3*N4:S4)
6	R(2)	=N4+C4	=O4+D4	=P4+E4	=Q4+F4	=R4+G4	=S4+H4			id3	=СУММ(C4:H4*N4:S4)
7	R(3)	=N6+C4	=O6+D4	=P6+E4	=Q6+F4	=R6+G4	=S6+H4			id4	=СУММ(C4:H4*N6:S6)
8										id5	=СУММ(C4:H4*N7:S7)
9											
10	Второй цикл обучения										
11											
12										id1	=СУММ(C2:H2*N7:S7)
13	R(4)	=N7+C3	=O7+D3	=P7+E3	=Q7+F3	=R7+G3	=S7+H3			id2	=СУММ(C3:H3*N7:S7)
14										id3	=СУММ(C3:H3*N13:S13)
15	R(5)	=N13+C4	=O13+D4	=P13+E4	=Q13+F4	=R13+G4	=S13+H4			id4	=СУММ(C4:H4*N13:S13)
16										id5	=СУММ(C4:H4*N15:S15)
17											
18	Третий цикл обучения										
19											
20										id1	=СУММ(C2:H2*\$N\$15:\$S\$15)
21										id2	=СУММ(C3:H3*\$N\$15:\$S\$15)
22										id3	=СУММ(C4:H4*\$N\$15:\$S\$15)
23											

Рис. 7. Представление идентификаторов сущностей и переопределение вектора решений в соответствии с синтаксисом MS Excel (последовательность 2)

Следует иметь в виду, что расчет идентификаторов сущностей для второй последовательности проводился на основе возможностей MS Excel по работе с массивами.

Необходимо заметить, что принцип работы данной экспертной системы основан не на тотальном переборе атрибутов распознаваемых сущностей, а на выявлении некоторого обобщенного показателя, называемого вектором решений R, который играет роль базы знаний и в совокупности с идентификаторами сущностей способен распознавать предъявляемые объекты, не прибегая к исходным значениям атрибутов.

Библиографический список

Головицына М.В. Информационные технологии проектирования радиоэлектронных средств // БИНОМ. Лаборатория знаний, Интернет-университет информационных технологий — ИНТУИТ.ру, 2008.

Пикалов И. Ю., Тарасюк В. Б. Построение обучающего модуля экспертной системы определения принадлежности сущности к определенному классу // Auditorium. Электронный научный журнал Курского государственного университета. 2017. № 1 (13). URL: <http://auditorium.kursksu.ru/pdf/013-014.pdf> (дата обращения: 12.03.2018).