УДК 681.518

Ю. О. Герман
Белорусский государственный технологический университет

ФОРМАЛИЗОВАННАЯ МОДЕЛЬ «ВИРТУАЛЬНОГО» УЧИТЕЛЯ

Приведено описание формализованной модели поведенческой базы знаний «виртуального» учителя в системе электронного обучения. Модель состоит из логических формул, включающих состояния «учителя» и сигналы, поступающие от «ученика». Задача вывода, решаемая на поведенческой базе знаний, требует найти новое состояние «учителя» в зависимости от предполагаемого текущего состояния «ученика» и самого «учителя». Рассмотрен вариант с противоречивой базой или нечеткой поведенческой базой знаний. Показано, как сформулировать общую математическую задачу, что позволяет с единых позиций строить поведение «виртуального» учителя в электронной обучающей системе. Задача вывода решения в противоречивой базе знаний решается с позиций отыскания максимального непротиворечивого ее подмножества, из которого выводима доказываемая формула.

Ключевые слова: база знаний, обучающая система, «виртуальный» учитель, вывод.

Yu. O. German
Belarusian State Technological University

A FORMAL MODEL OF “VIRTUAL” TUTOR

A formalized description of the behavioral knowledge base model of the virtual tutor in electronic learning system is given. The model consists of a number of logical formulas including the tutor’s states and the signals coming from the disciple. An inference problem solved on the basis of the behavioral knowledge base requires to find a new state of the tutor in dependence of his current state and the state of the disciple. The paper deals with the contradictory and fuzzy knowledge bases and suggests a new general problem formulation enabling one to build a tutor behavior in electronic learning system. The inference problem in contradictory system is solved on the basis of finding a maximum non-contradictory subset of the knowledge model providing sound proving for the theorem.

Key words: knowledge base, learning system, “virtual” tutor, inference.

Введение. Современное программное обеспечение для е-обучения представлено как простыми HTML-страницами, так и сложными системами управления обучением и учебным контентом (Learning Content Management Systems), использующимися в корпоративных компьютерных сетях. Сложные современные системы е-обучения включают (в том или ином виде и объеме) мультимедийный контент, поддержку интерактивного общения на естественном языке, развитую систему поиска информации, оценки знаний, средства обучения решению задач, техническую поддержку общения на разных уровнях и т. д.

В организации электронного обучения в настоящее время стандартом является SCORM (Shareable Content Object Reference Model) – международный стандарт для обмена учебными материалами на базе адаптированных спецификаций. Цель SCORM можно определить как независимость содержания электронного учебника от программы управления, что обеспечивает многократность использования учебных модулей, их переносимость и модификацию. В настоящее время рынок обучающих электронных систем представлен разными по сложности и возможностям адаптации к «личности» обучаемого программными системами. Современные тенденции развития рынка обучающих е-систем направлены в сторону универсализации и увеличения функциональности систем. Использование коммерческих систем управления электронным обучением не доступно большинству отечественных вузов по причине их высокой стоимости и необходимости продления лицензии на каждый учебный год. Системы с открытым исходным кодом позволяют реализовать тот же набор возможностей, что и коммерческие с существенно меньшими затратами и большей эффективностью. В конце прошлого века в связи с бурным развитием теоретико-прикладных аспектов искусственного интеллекта Брауном и Венгером было введено понятие «интеллектуальные обучающие системы» (ITS – Intelligent Tutoring System) [1, 2]. Ранее системы такого рода существовали под названием CAES (Computer-Aided Educational Systems), что не передавало роль искусственного интеллекта в той мере, в какой эта роль, несомненно, вышла на передний план в последнее время. Архитектура ITS состоит из трех модулей: модуля знаний (по предметной области), модуля «учителя» и модуля «ученика». Последние два задают поведение участников процесса обучения. Для описания поведения используются модели (знаний) на основе правил (rule-based models), а также модели на основе ограничений (constraint-based models).

Развитая система электронного on-line обучения строится как приложение, интегрирующее процессы «учителя» и «ученика», поведение которых организуется в соответствии с поведенческой базой знаний [3]. Основная задача, которая стоит перед разработчиками таких систем и рассматривается в настоящей работе, заключается в создании и развитии формальных средств для синтеза поведения «виртуального» учителя в различных контекстах процесса обучения. Для формального представления поведенческой базы знаний необходимо ввести множество сигналов, поступающих от «ученика», и множество состояний «учителя». Отслеживание процесса обучения сводится к определению следующих признаков (состояний «ученика» – A_i) [4]:

1) учащийся бессистемно просматривает материал (A_1);

2) учащийся возвращается к ранее пройденным фрагментам учебника (A_2);

3) учащийся остановился на некотором вопросе на долгое время (A_3);

4) процесс изучения завершен (инициируется учеником) (A_4);

5) ученик плохо отвечает на вопросы (A_5);

6) ученик не отвечает на сообщение (сообщение может быть нейтральным) (A_6);

7) ученик отказался от прохождения теста (A_7);

8) ученик прошел тест успешно (плохо) (A_8);

9) ученик слишком быстро просматривает страницы (A_9);

10) часто используется хэлпер (помощник) (A_{10});

11) выполнен переход на другую (не связанную) тему (A_{11});

12) ученик не пользуется хэлпером (A_{12});

13) ученик часто переходит к одной и той же странице (A_{13});

14) ученик проскакивает через страницы (A_{14});

15) ученик задерживается мало времени на более сложных страницах (A_{15});

16) ученик игнорирует инструкции «учителя» (A_{16});

17) ученик игнорирует предложение помощи или запрос (A_{17});

18) выбрана новая тема для изучения (A_{18});

19) истекло контрольное время (A_{19}).

Автомат «учитель» может находиться в следующих обобщенных состояниях [4]:

W_{obs} – наблюдение;

W_{contr} – контроль знаний;

W_{distr} – вмешательство;

W_{plan} – оценка состояния и планирование стратегии поведения.

В состоянии «Наблюдение» автомат «учитель» не вмешивается в процесс обучения, но отслеживает поведение ученика, а именно: что просматривает ученик, сколько времени, просматривает ли он материал в некоторой допустимой последовательности или хаотически и т. д. В состоянии «Контроль знаний» выполняется проверка усвоения материала в форме тестов или вопросов. Состояние «Вмешательство» предполагает, что учитель обнаружил отклонение от процесса обучения и требует от ученика выполнить определенные действия, например, вернуться к конкретному материалу, ответить на вопросы (пройти тест), сделать напоминание о том, что поведение ученика контролируется и оценивается и т. п. В состоянии «Оценка состояния и планирование стратегии поведения» выполняется коррекция состояния «учителя».

Перейдем к описанию поведенческой базы знаний системы с помощью формул логики.

Основная часть. Формулы поведенческой базы знаний могут быть представлены следующим (упрощенным) образом:

$$
\begin{gathered}
W_{obs}(t) \vee W_{distr}(t) \vee W_{contr}(t) \vee W_{plan}(t),\\
\neg W_{obs}(t) \vee \neg W_{distr}(t),\ \neg W_{obs}(t) \vee \neg W_{contr}(t),\\
\neg W_{obs}(t) \vee \neg W_{plan}(t),\ \neg W_{distr}(t) \vee \neg W_{contr}(t),\\
\neg W_{distr}(t) \vee \neg W_{plan}(t),\ \neg W_{contr}(t) \vee \neg W_{plan}(t),\\
W_{obs}(t+1) \vee W_{distr}(t+1) \vee W_{contr}(t+1) \vee\\
\vee W_{plan}(t+1) \vee Stop,\\
\neg W_{obs}(t+1) \vee \neg W_{distr}(t+1),\ \neg W_{obs}(t+1) \vee\\
\vee \neg W_{contr}(t),\ \neg W_{obs}(t+1) \vee \neg W_{plan}(t+1),\\
\neg W_{distr}(t+1) \vee \neg W_{contr}(t+1),\\
\neg W_{distr}(t+1) \vee \neg W_{plan}(t+1),\\
\neg W_{contr}(t+1) \vee \neg W_{plan}(t+1),\\
nil \to \neg A_1 \,\&\, \neg A_2 \,\&\, \neg A_3 \,\&\ldots\&\, \neg A_{19},\\
nil \vee A_1 \vee A_2 \vee A_3 \vee \ldots \vee A_{19}, \qquad (1)
\end{gathered}
$$

$$Stop \to \neg W_{obs}(t+1) \,\&\, \neg W_{distr}(t+1) \,\&$$
$$\&\, \neg W_{contr}(t+1) \,\&\, \neg W_{plan}(t+1),$$
$$A_4 \to Stop,\ A_{18} \to W_{contr}(t+1),$$
$$W_{obs}(t) \,\&\, A_{10} \to W_{distr}(t+1) \vee W_{contr}(t+1),$$
$$W_{obs}(t) \,\&\, (A_1 \vee A_2 \vee A_3 \vee A_{11} \vee A_{13}) \to W_{distr}(t+1),$$
$$W_{obs}(t) \,\&\, (A_6 \vee A_7) \to W_{obs}(t+1),$$
$$W_{obs}(t) \,\&\, A_{10} \to W_{obs}(t+1),$$
$$W_{obs}(t) \,\&\, A_{12} \to W_{obs}(t+1),$$
$$W_{obs}(t) \,\&\, (A_9 \vee A_{14} \vee A_{15}) \to W_{distr}(t+1) \vee$$
$$\vee\, W_{contr}(t+1),$$
$$W_{obs}(t) \,\&\, nil \to W_{obs}(t+1),$$
$$W_{distr}(t) \,\&\, nil \to W_{obs}(t+1),$$
$$W_{contr}(t) \,\&\, nil \to W_{obs}(t+1),$$
$$W_{plan}(t) \,\&\, nil \to W_{obs}(t+1),$$
$$W_{distr}(t) \,\&\, A_5 \to W_{obs}(t+1) \vee W_{contr}(t+1),$$
$$W_{distr}(t) \,\&\, (A_6 \vee A_{17}) \to W_{obs}(t+1),$$
$$W_{distr}(t) \,\&\, A_{19} \to W_{obs}(t+1) \vee W_{plan}(t+1),$$
$$W_{contr}(t) \,\&\, A_3 \to W_{plan}(t+1),$$
$$W_{contr}(t) \,\&\, (A_4 \vee A_8) \to W_{obs}(t+1),$$
$$W_{contr}(t) \,\&\, A_{19} \to Stop,$$
$$W_{plan}(t) \,\&\, (A_3 \vee A_{19}) \to W_{obs}(t+1).$$

В системе (1) *nil* соответствует отсутствию сигнала или нераспознанному сигналу. Итак, система «учитель − ученик» описывается как система взаимодействующих автоматов, причем поведение автомата «учитель» основано на поведенческой базе знаний, а поведение автомата «ученик» реализуется «учителем» гипотетически (модельно) на основании распознавания состояния «ученика» по сигналам (признакам). «Учитель» работает за «ученика» в том смысле, что принимает решения о его поведении по поступающим сигналам. Рассмотрим, как получить вывод решения в системе. В общем случае эта система может быть противоречивой или нечеткой. Пусть для каждого дизъюнкта D_k известна нечеткая мера p_k принадлежности его к множеству истинных формул (т. е. мера истинности). Пусть далее требуется найти корни $x_1, x_2, \ldots, x_n$ и проверить статистическую адекватность ему решения. Для каждого дизъюнкта D_k вводим булеву переменную $q_k \in 0{,}1$ и две импликации:

$$\begin{aligned} q_k &\to D_k, \\ D_k &\to q_k. \end{aligned} \qquad (2)$$

Потребуем

$$\textstyle\sum_k (q_k - p_k)^2 / p_k \to \min. \qquad (3)$$

Функционал (3) соответствует формуле для критерия χ^2. Он преобразуется к линейной целевой функции после возведения в квадрат выражения под знаком суммы:

$$\textstyle\sum_k (q_k - p_k)^2 / p_k = \sum_k (1 - 2p_k) q_k / p_k \to \min. \qquad (4)$$

Таким образом, получаем линейную систему неравенств с целевой линейной функцией и 0,1-переменными (дизъюнкту $\vee_i x_i^{\alpha_i}$ соответствует алгебраическое неравенство $\sum_i x_i^{\alpha_i} \geq 1$).

Эту задачу можно решить методом, изложенным в [5]. Статистическую адекватность решения следует проверить с помощью статистического критерия (в качестве иллюстрации похожей задачи с применением критерия χ^2 и подхода к ее решению см. [6]). Таким образом, задача построения выводов в нечеткой системе сводится к следующему:

1) построить целевую функцию (4);

2) заменить дизъюнкты алгебраическими неравенствами;

3) добавить в систему отрицание $\neg\varphi$ доказываемой формулы φ с мерой истинности $p_{\neg\varphi} = 1 - p_\varphi$. Найти корни.

Если решения нет, то формула доказана. Если решение есть, то проверить его статистическую адекватность (например, как показано в [4]). Если решение не «проходит» тест на адекватность, то считать φ доказанной. Если решение удовлетворяет статистическому критерию, то вывести φ нельзя. В случае противоречивой или четкой непротиворечивой системы функционал (4) преобразуется к виду

$$\textstyle\sum_k (q_k - p_k)^2 \to \min, \qquad (5)$$

и решение можно искать через определение максимального непротиворечивого подмножества дизъюнктов, из которого выводима доказываемая формула [6].

Обратимся к иллюстрации. Требуется вывести формулу $\varphi = \neg x_2$ в системе с четкими формулами ($\forall k\ p_k = 1.0$):

$$\begin{gathered} q_1 \leftrightarrow x_1 \vee \neg x_2 \vee x_3, \\ q_2 \leftrightarrow \neg x_1 \vee \neg x_2, \\ q_3 \leftrightarrow x_1 \vee \neg x_3, \\ q_4 \leftrightarrow \neg x_1 \vee \neg x_3, \\ q_5 \leftrightarrow x_2 \vee x_3, \\ x_2. \end{gathered} \qquad (6)$$

Целевая функция:

$$-q_1 - q_2 - q_3 - q_4 - q_5 \to \min. \qquad (7)$$

Эту систему можно решить, например, в Excel, заменив символы логических операций соответствующими символами алгебраических операций и использовав следующие эквивалентности:

$$\begin{gathered} A \to B = \neg A \vee B, \\ \neg A = 1 - A. \end{gathered}$$

Получаем:

$$
\begin{gathered}
-q_1+x_1-x_2+x_3\geq -1,\\
-x_1+q_1\geq 0,\\
x_2+q_1\geq 1,\\
-x_3+q_1\geq 0,\\
-q_2-x_1-x_2\geq -2,\\
x_1+q_2\geq 1,\\
x_2+q_2\geq 1,\\
-q_3+x_1-x_3\geq -1,\\
-x_1+q_3\geq 0,\\
x_3+q_3\geq 1,\\
-q_4-x_1-x_3\geq -2,\\
x_1+q_4\geq 1,\\
x_3+q_4\geq 1,\\
-q_5+x_2+x_3\geq 0,\\
-x_2+q_5\geq 0,\\
-x_3+q_5\geq 0,\\
-q_1-q_2-q_3-q_4-q_5\to \min.
\end{gathered}
$$

Решение в Excel дает: $x_1=0, x_2=1, x_3=0, q_1=0$, $q_2=q_3=q_4=q_5=1$. Можно сделать вывод о том, что максимальное по включению непротиворечивое множество содержит дизъюнкты q_2, q_3, q_4, q_5. В этом множестве выполняется формула x_2. Теперь следует присоединить к системе дизъюнкцию тех q_i, которые получили нулевое значение в найденном решении. В примере нужно присоединить к системе q_1 и выполнить повторный поиск решения. Новое решение будет таким: $x_1=0$, $x_2=1$, $x_3=1$, $q_1=1$, $q_2=1$, $q_3=0$, $q_4=q_5=1$. Присоединим к системе формулу q_3. К этому моменту система примет следующий вид:

$$
\begin{gathered}
-q_1+x_1-x_2+x_3\geq -1,\\
-x_1+q_1\geq 0,\\
x_2+q_1\geq 1,\\
-x_3+q_1\geq 0,\\
-q_2-x_1-x_2\geq -2,\\
x_1+q_2\geq 1,\\
x_2+q_2\geq 1,\\
-q_3+x_1-x_3\geq -1,\\
-x_1+q_3\geq 0,\\
x_3+q_3\geq 1,\\
-q_4-x_1-x_3\geq -2,\\
x_1+q_4\geq 1,\\
x_3+q_4\geq 1,\\
-q_5+x_2+x_3\geq 0,\\
-x_2+q_5\geq 0,\\
-x_3+q_5\geq 0,\\
q_1=1,\\
q_3=1,\\
-q_1-q_2-q_3-q_4-q_5\to \min.
\end{gathered}
$$

Решение: $x_1=1$, $x_2=1$, $x_3=1$, $q_1=1$, $q_2=0$, $q_3=1$, $q_4=0$, $q_5=1$. Присоединим дизъюнкт $q_2\vee q_4$. Система становится противоречивой. Процесс завершается. В ходе итераций добавлены следующие новые формулы:

$$
\begin{cases}
q_2\vee q_4,\\
q_1,\\
q_3.
\end{cases}
$$

Теорема. Пусть SYS^t непротиворечива, SYS^t & & $D^t\to\neg\varphi$, причем не имеет места ни $SYS^t\to$ $\to\neg\varphi$, ни $SYS^t\to\neg D^t$. Тогда $\{SYS^t, D^t\}$ непротиворечива.

Доказательство. Перепишем SYS^t & $D^t\to$ $\to\neg\varphi$ как $\neg SYS^t\vee\neg D^t\vee\neg\varphi$. Допустим, что $SYS^t\to D^t$. Но тогда сразу получим, что $SYS^t\to$ $\to\neg\varphi$, что противоречит условиям теоремы. Следовательно, неверно, что $SYS^t\to D^t$. Если допустить, что $\{SYS^t, D^t\}$ противоречива, то ввиду непротиворечивости SYS^t заключаем, что $SYS^t\to$ $\to\neg D^t$, но это исключено условиями теоремы.

Данная терема позволяет выяснить заключительный момент, состоящий в том, что полученная противоречивая система действительно содержит непротиворечивую часть, из которой выводима доказываемая формула. Согласно теореме выше, нужно показать непротиворечивость $\{SYS^t, D^t\}$. Применительно к разобранному примеру следует найти хотя бы одну выполняющую интерпретацию для заключительной системы без доказываемой формулы (удалив x_2). Таким решением, например, является следующее: $q_1=q_2=q_3=1$, $q_4=q_5=0$, $x_1=1$, $x_2=0$, $x_3=1$.

Итак, получен ответ на исходную задачу – интересующее нас максимальное по включению непротиворечивое подмножество дизъюнктов системы (6), из которого выводима формула $\neg x_2$, есть

$$
\begin{gathered}
x_1\vee\neg x_2\vee x_3,\\
\neg x_1\vee\neg x_2,\\
x_1\vee\neg x_3\,.
\end{gathered}
$$

Заключение. Достоинство представленного в статье подхода состоит в том, что он «сохраняет» стандартную резолюционную схему вывода и, таким образом, позволяет избежать известных аномалий нечеткого резолюционирования. В сравнении с методами нечеткого вывода типа Мамдани, Цукамото или Сугено данный подход не ограничивает степень общности представления формул. Это же касается неоднозначности представления нечеткой логической импликации (формулы «если − то»). Указанное свойство дает ключ к построению машины логического вывода в нечеткой (противоречивой) логике посредством аппроксимации нечетких значений значениями многозначной логики, достаточными для практических приложений.

В [6] представлен подход для сведения системы нечетких логических формул к задаче целочисленной оптимизации с 0,1-переменными. Ограничение этого подхода состоит в значительной вычислительной сложности процедуры целочисленной булевой оптимизации.

Таким образом, разработанный в данной статье подход обладает научной и практической новизной, имеет положительные особенности и может быть адаптирован для задач построения выводов в различных системах знаний электронных обучающих систем.

Литература

1. Wenger E. Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge. California: Morgan Kaufmann Publishers, 1987. 486 p.

2. Brown J. S., Burton R. R., Kleer de J. Pedagogical, natural language and knowledge engineering techniques in Sophie I, II, and III // Intelligent tutoring systems. New York: Academic Press, 1982. P. 227–282.

3. Гурин Н. И., Герман О. В., Герман Ю. О. Разработка активной обучающей среды для электронного учебника // Информатизация образования – 2012: педагогические основы разработки и использования электронных образовательных ресурсов: материалы Междунар. науч. конф., Минск, 24–27 окт. 2012 г. / Белорус. гос. ун-т; редкол.: В. В. Казаченок (отв. ред.) [и др.]. Минск, 2012. С. 95–99.

4. Гурин Н. И., Герман О. В., Герман Ю. О. Технология разработки компьютерных обучающих систем с функциями виртуального преподавателя // Труды БГТУ. 2011. № 6: Физ.-мат. науки и информатика. С. 146–149.

5. Герман Ю. О., Самко А. Р., Герман О. В. Эффективный в среднем алгоритм для задачи о покрытии с приложением к нечеткой классификации и нечеткому выводу // Доклады БГУИР. 2009. № 7. С. 93–100.

6. Герман О. В., Самко А. Р., Герман Ю. О. Система вывода для нечеткой логики на основе многозначных исчислений Я. Лукасевича // Труды БГТУ. 2010. № 6: Физ.-мат. науки и информатика. С. 190–192.

References

1. Wenger E. Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge. California, Morgan Kaufmann Publishers, 1987. 486 p.

2. Brown J. S., Burton R. R., Kleer de J. Pedagogical, natural language and knowledge engineering techniques in Sophie I, II, and III. *Intelligent tutoring systems*. New York, Academic Press, 1982. P. 227–282.

3. Gurin N. I., German O. V., German Yu. O. [Development of the active learning environment for electronic tutorial]. *Materialy Mezhdunarodnoy nauchnoy konferentsii (Informatizatsiya obrazovaniya – 2012: pedagogicheskiye osnovy razrabotki i ispol'zovaniya elektronnykh obrazovatel'nykh resursov* [Informatization of education – 2012: the pedagogical fundamentals for the development and application of digital educational resources]. Minsk, 2012, pp. 95–99 (In Russian).

4. Gurin N. I., German O. V., German Yu. O. Technology of creating the computer-based educational systems with a virtual teacher functions. *Trudy BGTU* [Proceedings of BSTU], 2011, no. 6: Physical-mathematical sciences and informatics, pp. 146–149 (In Russian).

5. German Yu. O., Samko A. R., German O. V. An efficient in average algorithm for the covering problems with application to a fuzzy classification and fuzzy inference. *Doklady BGUIR* [Reports of BSUIR], 2009, no. 7, pp. 93–100 (In Russian).

6. German O. V., Samko A. R., German Yu. O. A fuzzy logic inference system on the basis of multivalued Lukasiewicz' logics. *Trudy BGTU* [Proceedings of BSTU], 2010, no. 6: Physical-mathematical sciences and informatics, pp. 190–192 (In Russian).

Информация об авторе

Герман Юлия Олеговна – старший преподаватель кафедры информационных систем и технологий. Белорусский государственный технологический университет (220006, г. Минск, ул. Свердлова, 13а, Республика Беларусь). E-mail: MaksHladki@gmail.com

Information about the author

German Yuliya Olegovna – Senior Lecturer, the Department of Information Systems and Technologies. Belarusian State Technological University (13a, Sverdlova str., 220006, Minsk, Republic of Belarus). E-mail: MaksHladki@gmail.com

Поступила 12.12.2016