

*НОВЫЕ МЕТОД УСИЛЕНИЯ ИНТЕЛЛЕКТА И СПОСОБ ПРЕДСТАВЛЕНИЯ
ДЕРЕВА ПРИНЯТИЯ РЕШЕНИЙ, КОТОРЫЕ ПРИБЛИЖЕНЫ К
ЕСТЕСТВЕННОМУ ИНТЕЛЛЕКТУ*

О.Б. ПОПОВА

*Кубанский государственный технологический университет,
350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
электронная почта: popova_ob@mail.ru*

Актуальной задачей сегодня является адаптация современных методов усиления интеллекта, которые приближены к естественному интеллекту, и методов автоматического анализа данных к тем задачам, которые до сих пор решаются вручную. В статье предложено решение такой задачи, где рассмотрен новый способ представления дерева принятия решений, использующий структурирование знаний близкое к естественному интеллекту. Это стало возможным из-за исследований в области усиления естественного интеллекта, которые позволяют улучшить когнитивные способности естественного интеллекта внешними устройствами. Структурирование осуществляется при помощи разработанного автором бинарного дерева системы вопросов и ответов. Оно представляет собой одно дерево решений, которое несёт в себе все идеальные решения для всех известных ситуаций. Таким образом, оно не учитывает остальные не эффективные случаи. Тогда листами такого дерева решений будет множество «наиболее эффективных» решений. Новый метод усиления интеллекта и новый способ представления дерева принятия решений могут быть использованы в интеллектуальной системе поддержки принятия решений, которая позволяет усилить естественный интеллект учёного в поисковом исследовании.

Ключевые слова: метод усиления интеллекта, метод представления дерева принятия решений, Data Mining, естественный интеллект, бинарное дерево системы вопросов и ответов.

Сегодня стремительно развиваются методы сбора, обработки и хранения информации, которые позволяют собирать и хранить огромные массивы данных. Объёмы таких хранимых данных велики, поэтому стали так актуальными методы автоматического исследования данных Data Mining. Среди которых есть широко используемый метод – деревья решений. Он прост в понимании и интерпретации, не требует подготовки данных, является надёжным методом и позволяет работать с большим объёмом информации без специальных подготовительных процедур.

Но ряд недостатков данного метода не позволяет решить некоторые современные задачи, которые до сих пор решаются вручную, или не решаются с требуемой эффективностью и точностью. Среди них задача поискового

исследования, связанная с поиском нужного метода решения исследуемой задачи из заданной области научного исследования.

Перечислим основные недостатки данного метода решения [1]:

- задача получения оптимального дерева решений даже для простой задачи является NP-полной, где единственно оптимальное решение существует лишь локально в каждом узле, поэтому современные алгоритмы получения деревьев решений не могут обеспечить оптимальности всего дерева целиком;

- из-за возникновения проблемы «переобучения» приходится использовать метод «регулирования глубины дерева»;

- если модель описывает концепт сложным путём, то дерево решений будет непомерно большим, то необходимо использовать дополнительные подходы со своими алгоритмами, чтобы решить данную проблему;

- для разных данных необходимо использовать методики, распределяющие особым образом информационный вес атрибутов;

- на сегодняшний день не существует эвристических правил, которые имели бы большую практическую ценность, так как многие из них применимы в каких-то частных случаях;

- все существующие правила, необходимые для построения дерева решений, рассчитаны для работы с числовыми множествами, так как это позволяет применить далее критериальный метод решения или использовать метод бинарных отношений для получения значения функции выбора, но они совершенно не подходят к не числовым множествам;

- одно дерево решений составляется для всех случаев одной ситуации, поэтому требуется составить столько деревьев решений, сколько существует ситуаций – это не позволяет оптимально решать задачи и получать идеальное решение.

Данные недостатки связаны с традиционным способом представления дерева принятия решений. Так, набор данных, записывается следующим образом: $(x, Y) = (x_1, x_2, x_3 \dots x_k, Y)$, где Y – целевая переменная, которую необходимо

проанализировать, классифицировать и обобщить. Эта переменная зависит от вектора \bar{x} , который состоит из входных переменных - $x_1, x_2, x_3, \dots, x_k$. Понятно, что решением для данной ситуации будет тот случай со своим набором входных данных, который даёт наиболее подходящее решение Y . Среди представленных данных – это может быть верным решением, но не идеальным, так как такой набор данных соответствует определённому уровню опыта, который может быть не полным и требовать дополнительного обучения.

Автор предлагает новый способ представления дерева принятия решений, который приближен к естественному интеллекту, и использует разработанный им новый метод усиления интеллекта.

Для улучшения эффективности метода было принято составлять всего одно дерево решений, которое несёт в себе все идеальные решения для всех известных ситуаций. Таким образом, оно не учитывает остальные не продуктивные, не эффективные, не оптимальные и так далее случаи.

Тогда листьями такого дерева решений будет множество «наиболее эффективных» решений $(x_i, Y_i) = (x_{1i}, x_{2i}, x_{3i} \dots x_{ki}, Y_i)$, где \bar{x}_i - это множество наиболее верных когнитивных представлений о всех ситуациях, которые дают единственно верные и «наиболее подходящие» решения. Каждому верному когнитивному представлению о ситуации соответствует своё «наиболее эффективное» решение, например, $(x_2, Y_2) = (x_{12}, x_{22}, x_{32} \dots x_{k2}, Y_2)$. Таким образом, все пути по дереву от корня до листа – это самые точные когнитивные представления о всех ситуациях, где число решений равно числу известных ситуаций и количеству возможных путей от корня до каждого листа дерева. Внесение новых элементов в дерево возможно лишь при появлении новых «идеальных» методов решения совершенно новых ситуаций.

Такое представление дерева решений отличается от известного ранее тем, что для каждого дерева решений существовало множество когнитивных представлений об одной ситуации \bar{x}_i , среди которых только одно подходило и становилось решением целевой функции Y . В процессе анализа этой ситуации

происходило обучение и дополнение дерева решений новыми более точными решениями.

Личная когнитивная «сила» представления ситуации зависит от многих факторов, таких как опыт, объём знаний, которые получаются в процессе обучения, «сила» естественного интеллекта, которые стали исследоваться только недавно, поэтому данный подход не был известен ранее. Недавно появились методы усиления естественного интеллекта, а так же внешние способы усиления естественного интеллекта, которые сейчас набирают силу и популярность, но до сих пор не интегрированы в Data Mining и интеллектуальные системы поддержки принятия решений и другие информационные технологии.

Поэтому при составлении модели ситуации её описывали разными случаями когнитивного восприятия от самого худшего до самого лучшего, где самое лучшее когнитивное восприятие соответствовало личному субъективному опыту, которое на самом деле не являлось объективно лучшим и требовало различных алгоритмов обучения. Это позволяло описать модель в рамках стандартного представления об искусственном интеллекте, который не был приближен к естественному интеллекту, но позволял решать хорошо формализованные задачи известными математическими методами. Здесь данные о ситуации представляются массивами, бинарными отношениями, графами, нейронными сетями, которые требуют различных алгоритмов работы с большим объёмом данных.

Так, любая ситуация может быть описана разными людьми определённым множеством \bar{x}_i , где возможна различная комбинация входных элементов, которые описывают эту ситуацию. Это комбинация может отличаться значениями, разной последовательностью входных элементов, их количеством. Ещё возможны случаи, когда некоторые входные данные будут разбиты на составные части, которые так же могут отличаться друг от друга своей последовательностью и значениями.

Опишем одну такую ситуацию:

$$\vec{x}_i = \begin{cases} (x_5, x_1, \dots, x_{12}) \rightarrow \text{'худшее когнитивное представление'}; \\ (x_5, x_2, x_1, x_{12}, x_3, x_{14}, \dots, x_{12}), \text{ где } x_1 = x_{1_1} + x_{1_2} + x_{1_3} + x_{1_4}; \\ (x_1, x_5, x_2, x_{12}, x_{1_2}, x_9, x_{1_3}, x_{1_4}, \dots, x_4) \\ \dots \\ (x_7, x_3, x_2, x_{15}, x_1, x_{11}, \dots, x_{10}) \rightarrow \text{'лучшее когнитивное представление'}. \end{cases}$$

Если бы с такой ситуацией столкнулся бы естественный интеллект, то из множества случаев одной ситуации, он запомнил бы одну, которая соответствует самому лучшему когнитивному представлению. Так же он бы запомнил одно правильное решение, соответствующее данной ситуации, а при встрече с аналогичными ситуациями он применил бы верное решение. В течение жизни количество верных решений растёт, позволяя говорить об опыте, в котором фиксируются в основном лучшие когнитивные представления.

Естественно, что наиболее опытным экспертом можно считать того, чей естественный интеллект запечатлел наибольшее количество верных решений для различных ситуаций, которые он может в нужный момент использовать, не ошибаясь. Они могут представлять особую ценность, если человек ищет решение, а не выход из испорченной им ситуации, после принятия неверного решения. В этом случае требуется помнить весь опыт и все способы решения – хорошие и плохие.

В новом способе представления дерева решений используется другой взгляд на моделирование ситуаций. В дереве решений используются когнитивные представления о различных ситуациях наиболее опытного эксперта, где каждая ситуация будет решена сразу же «наиболее подходящим» для данной ситуации способом. Это позволит сделать дерево решений оптимальным и наиболее доступным для интерпретации.

Ещё этот способ позволяет достаточно хорошо визуализировать все знания, доступные наиболее опытному эксперту, сразу же на одном дереве. Это соответствует наиболее продуктивному способу структурирования информации и представления знаний естественным интеллектом [2], который известен и применяется как методика, улучшающая когнитивные способности

естественного интеллекта, позволяющая точно структурировать полученные им знания.

В предлагаемом способе представления дерева решений используются свои правила получения корня, промежуточных узлов и листьев дерева [3-5], которые отличаются от известных.

Они отражают алгоритм получения дерева принятия решений [6].

Структура дерева решений позволяет не только хранить знания, но и искать решение, перемещаясь по дереву, отвечая на вопросы, которые помогают составить верное когнитивное представление о текущей ситуации. В таком дереве однотипные ситуации представлены соответствующим набором свойств (\bar{x}), то есть своей последовательностью вопросов, которые в результате дадут один верный ответ – «наиболее подходящее» решение.

Предлагаемое дерево отличается от известных тем, что в промежуточных узлах дерева располагаются вопросы (атрибуты), в которых сформулированы сложно сочинённые свойства (не обязательно числовые), отделяющие группу решений от общего множества всех решений. В этом дереве не используется понятие «ветка» и не проставляется информационный вес атрибутов, так как изменился общий подход в представлении дерева решений. На каждый вопрос существует два ответа «да» и «нет», выбирая один из которых происходит переход на следующий уровень дерева, то есть каждый узел дерева при разбиении имеет только два потомка. Такое разбиение свойственно естественному структурированию.

Такое дерево относится к бинарным деревьям. Из известных алгоритмов построения деревьев решений наиболее близким к предлагаемому является алгоритм CART (Classification and Regression Tree). Он используется для получения дихотомической классификационной модели и решения задач классификации и регрессии.

Предлагаемые новые метод усиления интеллекта и способ представления дерева принятия решений, которые приближены к естественному интеллекту, были опробованы на решении задачи выбора «наиболее подходящего» метода

оптимизации решаемой задачи из множества известных методов оптимизации. Полученные результаты были опубликованы в ряде статей [7-11], монографий [6,12], так же был получен патент на изобретение [13] и свидетельство на программу [14].

Они могут быть так же использованы для поискового исследования, которое позволяет определить метод решения исследуемой задачи из заданной области научного исследования. Где дерево решений – это все известные методы решения научных задач из заданной научной области, применяемые для разных однотипных ситуаций.

ЛИТЕРАТУРА

1. Википедия – свободная энциклопедия: Дерево принятия решений [Электронный ресурс] – Режим доступа: http://ru.wikipedia.org/wiki/Дерево_принятия_решений, 09.01.2016

2. Урок 3. Структурирование информации [Электронный ресурс] – Режим доступа: <http://4brain.ru/memory/strukturirovanie.php>, 09.01.2016.

3. Попова О.Б., Попов Б.К., Ключко В.И. Получение корня бинарного дерева системы вопросов и ответов // Современные проблемы науки и образования. – 2013. – № 3; URL: <http://www.science-education.ru/109-9146>.

4. Попова О.Б., Попов Б.К., Ключко В.И. Правила получения элементов бинарного дерева системы вопросов и ответов // Фундаментальные исследования. – 2013. – №6-1. – С. 55-59; [Электронный ресурс] – Режим доступа: URL: <http://www.fundamental-research.ru/ru/article/view?id=31413>

5. Popova O., Popov B., Karandey V., Evseeva M. Intelligence amplification via language of choice description as a mathematical object (binary tree of question-answer system) // Procedia – Social and Behavioral Sciences. – 2015. – V. 214. – С. 897–905; URL: <http://www.sciencedirect.com/science/article/pii/S1877042815061030>.

6. Бинарное дерево выбора знания из области знания, используя систему вопросов и ответов. Теория и практика: монография / О.Б. Попова, Б.К. Попов, В.И. Ключко; ФБГОУ ВПО «Кубан. гос. технол. ун-т». – Краснодар: Издательский Дом – Юг, 2013 – 166 с.

7. Попова О.Б. Структура технической системы процесса выбора метода оптимизации. Деп. в ВИНТИ № 243-В2012 от 25.05.2012

8. Попова О.Б., Попов Б.К. Эквивалентная замена процесса выбора знания из области знаний на техническую систему вопросов и ответов // Фундаментальные

исследования. – 2012. – №11-5. – С. 1201-1205; URL: <http://www.fundamental-research.ru/ru/article/view?id=30734>

9. Попова О.Б., Попов Б.К. Замена реальной системы (процесс выбора метода оптимизации) на техническую систему (программа-советчик «Оптимэль») // Современные проблемы науки и образования. – 2012. – №5. – С. 132;

10. Попова О.Б. Информационный поиск по научной проблеме - сокращение времени выбора метода оптимизации решаемой задачи и его обработка. Деп. в ВИНТИ № 31-В2010 от 26.05.2010

11. Попова О.Б. Сокращение времени выбора метода оптимизации решаемой задачи. Деп. в ВИНТИ № 712-В2009 от 23.11.2009

12. Попова О.Б., Попов Б.К., Ключко В.И. Системный анализ процесса выбора метода оптимизации информационной системы: монография. – Краснодар: ООО «Издательский Дом-Юг», ФГБОУ ВПО «КубГТУ», 2012. – 130 с.

13. Попова О.Б., Попов Б.К. Интеллектуальная информационная система выбора «Оптимэль». Патент на изобретение RUS № 2564641 от 27.05.2014.

14. Попова О.Б., Попов Б.К. «Оптимэль». Свидетельство о государственной регистрации программы №2012615868 от 27.06.2012.

REFERENCES

1. Vikipediya – svobodnaya entsiklopediya: Derevo prinyatiya resheniy [Elektronnyy resurs] – Rezhim dostupa: http://ru.wikipedia.org/wiki/Derevo_prinyatiya_resheniy, 09.01.2016

2. Urok 3. Strukturirovanie informatsii [Elektronnyy resurs] – Rezhim dostupa: <http://4brain.ru/memory/strukturirovanie.php>, 09.01.2016.

3. Popova O.B., Popov B.K., Klyuchko V.I. Poluchenie kornya binarnogo dereva sistemy voprosov i otvetov // Sovremennye problemy nauki i obrazovaniya. – 2013. – № 3; URL: <http://www.science-education.ru/109-9146>.

4. Popova O.B., Popov B.K., Klyuchko V.I. Pravila polucheniya elementov binarnogo dereva sistemy voprosov i otvetov // Fundamentalnye issledovaniya. – 2013. – №6-1. – S. 55-59; [Elektronnyy resurs] – Rezhim dostupa: URL: <http://www.fundamental-research.ru/ru/article/view?id=31413>

5. Popova O., Popov B., Karandey V., Evseeva M. Intelligence amplification via language of choice description as a mathematical object (binary tree of question-answer system) // Procedia – Social and Behavioral Sciences. – 2015. – V. 214. – S. 897–905; URL: <http://www.sciencedirect.com/science/article/pii/S1877042815061030>.

6. Binarnoe derevo vybora znaniya iz oblasti znaniya, ispolzuya sistemu voprosov i otvetov. Teoriya i praktika: monografiya / O.B. Popova, B.K. Popov, V.I. Klyuchko; FBGOU VPO «Kuban. gos. tekhnol. un-t».-Krasnodar: Izdatelskiy Dom - Yug, 2013-166 s.

7. Popova O.B. Struktura tekhnicheskoy sistemy protsessa vybora metoda optimizatsii. Dep. v VINITI № 243-V2012 ot 25.05.2012

8. Popova O.B., Popov B.K. Ekvivalentnaya zamena protsessa vybora znaniya iz oblasti znaniy na tekhnicheskuyu sistemu voprosov i otvetov // Fundamentalnye issledovaniya. – 2012. – №11-5. – S. 1201-1205; URL: <http://www.fundamental-research.ru/ru/article/view?id=30734>

9. Popova O.B., Popov B.K. Zamena realnoy sistemy (protsess vybora metoda optimizatsii) na tekhnicheskuyu sistemu (programma-sovetchik «Optimel») // Sovremennyye problemy nauki i obrazovaniya. – 2012. – №5. – S. 132;

10. Popova O.B. Informatsionnyy poisk po nauchnoy probleme - sokrashchenie vremeni vybora metoda optimizatsii reshaemoy zadachi i ego obrabotka. Dep. v VINITI № 31-V2010 ot 26.05.2010

11. Popova O.B. Sokrashchenie vremeni vybora metoda optimizatsii reshaemoy zadachi. Dep. v VINITI № 712-V2009 ot 23.11.2009

12. Popova O.B., Popov B.K., Klyuchko V.I. Sistemnyy analiz protsessa vybora metoda optimizatsii informatsionnoy sistemy: monografiya. – Krasnodar: OOO «Izdatelskiy Dom-Yug», FGBOU VPO «KubGTU», 2012. – 130 s.

13. Popova O.B., Popov B.K. Intellektualnaya informatsionnaya sistema vybora «Optimel». Patent na izobreteniye RUS № 2564641 ot 27.05.2014.

14. Popova O.B., Popov B.K. «Optimel». Svidetelstvo o gosudarstvennoy registratsii programmy №2012615868 ot 27.06.2012.

*THE NEW METHOD OF THE INTELLIGENCE AMPLIFICATION AND THE WAY
OF PRESENTATION DECISION TREES THAT ARE CLOSE
TO THE NATURAL INTELLIGENCE*

O.B. POPOVA

*Kuban State Technological University,
2, Moskovskaya st., Krasnodar, Russian Federation, 350072,
e-mail: popova_ob@mail.ru*

Actual problem today is to adapt modern techniques amplification intelligence, which are close to the natural intelligence, and automated methods of data analysis to the tasks that are still solved by handmade. The article provides a solution to this problem, which is considered a new way of presenting the decision tree using the structuring of knowledge closely to the natural intelligence. This was made possible because of the research in the field of

amplification natural intelligence that can improve the cognitive abilities of natural intelligence by external devices. Structuring is carried out using the developed by the author of a binary tree of questions and answers. It represents one decision tree that carries all the ideal solution for all known situations. Thus, it does not account the rest are not effective cases. Then leaf of the decision tree is a set of «best» solutions. The new method of the intelligence amplification and the new way of representing a decision tree can be used in intelligent decision support system that allows you to strengthen natural intelligence of the scientist in the pilot study.

Key words: the method of the intelligence amplification, the method of presenting a decision tree, Data Mining, natural intelligence, the binary tree of the system of questions and answers.