УДК 004.93'1; 004.932; 002.6-027.21; 002.6:001.8

МНОГОУРОВНЕВЫЕ АЛГОРИТМЫ ДЛЯ ЗАДАЧ ПРИНЯТИЯ РЕШЕНИЙ ПРЕЦЕДЕНТНОГО ТИПА

В. В. КРАСНОПРОШИН[1)]***, В. А. ОБРАЗЦОВ***[1)]

[1)]*Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь*

Рассматривается специальный класс задач принятия решений прецедентного типа, которые часто возникают в слабо формализованных предметных областях. Для решения таких задач, как правило, применяются эвристические алгоритмы, которые не могут быть строго обоснованы. Показано, что данный класс задач сводится к стандартной задаче распознавания образов с обучением. Это позволяет вместо эвристических алгоритмов использовать многоуровневые модели, которые дают возможность повысить точность решения, а в некоторых случаях обосновать его правильность. Приведен анализ различных вариантов построения многоуровневых моделей. Предложен многоуровневый алгоритм для задачи принятия решений, основанный на структурировании информации.

Ключевые слова: многоуровневый алгоритм; задача принятия решений; прецедентный тип; распознавание образов с обучением; модели корректировки; модели на основе структурирования информации.

Благодарность. Работа выполнена при поддержке Белорусского республиканского фонда фундаментальных исследований (грант № Ф21АРМ-005).

MULTILEVEL ALGORITHMS FOR PRECEDENT-TYPE DECISION-MAKING PROBLEMS

V. V. KRASNOPROSHIN[a]***, V. A. OBRAZTSOV***[a]

[a]*Belarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus*
Corresponding author: V. A. Obraztsov (obraztsov@bsu.by)

In this paper, we considers a special class of precedent-type decision-making problems, which often arise in weakly formalised subject areas. To solve such problems, as a rule, heuristic algorithms are used, which cannot be strictly justified. It is shown that this class of problems can be reduced to a standard problem of pattern recognition with learning. Instead of heuristic algorithms, this allows to use multilevel models that make it possible to improve the accuracy of the solution, and in some cases to justify its correctness. An analysis of different variants for constructing multilevel models is given. A multilevel algorithm for the decision-making problem based on the structuring of information is proposed.

Keywords: multilevel algorithms; decision-making problem; precedent-type; pattern recognition with learning; correction models; models based on information structuring.

Acknowledgements. This work was supported by the Belarusian Republican Foundation for Fundamental Research (grant No. F21ARM-005).

Образец цитирования:
Краснопрошин ВВ, Образцов ВА. Многоуровневые алгоритмы для задач принятия решений прецедентного типа. *Журнал Белорусского государственного университета. Математика. Информатика.* 2023;3:82–91.
EDN: QJCRMQ

For citation:
Krasnoproshin VV, Obraztsov VA. Multilevel algorithms for precedent-type decision-making problems. *Journal of the Belarusian State University. Mathematics and Informatics.* 2023; 3:82–91. Russian.
EDN: QJCRMQ

Авторы:
Виктор Владимирович Краснопрошин – доктор технических наук, профессор; профессор кафедры информационных систем управления факультета прикладной математики и информатики.
Владимир Алексеевич Образцов – кандидат физико-математических наук; доцент кафедры информационных систем управления факультета прикладной математики и информатики.

Authors:
Viktor V. Krasnoproshin, doctor of science (engineering), full professor; professor at the department of information management systems, faculty of applied mathematics and computer science.
krasnoproshin@bsu.by
Vladimir A. Obraztsov, PhD (physics and mathematics); associate professor at the department of information management systems, faculty of applied mathematics and computer science.
obraztsov@bsu.by

Анализ проблемы и постановка задачи

Проблема принятия решений играет важную роль в жизни человека. В результате развития формальных методов ее решения сложилась отдельная дисциплина – теория принятия решений (ТПР) [1]. Становление ТПР определялось, с одной стороны, развитием вычислительной техники и математических методов, а с другой – возникновением новых задач в сфере жизнедеятельности человека. Появление задач в порой слабо формализованных предметных областях требует модификации уже существующих и разработки новых подходов. В последнее время это часто связано с проблемами измеримости и структуризации информации, использования экспертных знаний и многими другими аспектами ТПР. Решение подобных проблем невозможно без привлечения современных научных теорий, таких как теория искусственного интеллекта, теория распознавания образов (ТРО) и пр. Симбиоз ТПР и ТРО как раз и представляет собой одно из возможных направлений в развитии новых моделей и методов принятия решений.

Универсальной постановки для задач принятия решений не существует. Построение решений осуществляется стандартным для математики образом. Если обозначить через X множество возможных решений, то процесс построения решений реализуется в множестве X^2. Для организации процесса в множестве X^2 необходимо определить тип отношения на этом множестве: либо частичный порядок, либо эквивалентность. Множества могут быть заданы либо аналитически (с использованием принципа свертки), либо по прецедентности (или по примерам).

Таким образом, конкретизируя тип отношения на множестве X^2, а также способы задания информации и принятия решений, можно получить большое разнообразие задач принятия решений.

Ограничимся случаем, когда на множестве X^2 задано отношение эквивалентности. Следствием этого [2] является разбиение множества решений X на непересекающиеся подмножества, включающие сходные между собой решения. Существенную роль при этом играет способ задания информации о множестве X. Практически важным является случай, когда такая информация задана неполностью (частично). В этой ситуации говорят о задании информации по прецедентности (или по примерам). Приходим к варианту задачи, который можно назвать задачей выбора или в формальном смысле задачей вычисления свойств объектов $x \in X$. Сам же способ принятия решений в такой последовательности просто ассоциируется с вычисленным свойством [3].

В данной ситуации задача принятия решений сводится к вычислению (определению) свойств анализируемой информации и на формальном уровне может быть сведена к одной постановке. Это задачи исчисления высказываний, предикатов [4], задачи логической диагностики [5] и, наконец, просто плохо формализованные задачи ТРО [6].

Во всех перечисленных задачах имеются множества, разбитые на подмножества, которые заданы примерами объектов – носителями определенного свойства. Таким образом, приходим к классическому варианту задачи распознавания образов с обучением: множество объектов X произвольной природы разбито на некоторое, возможно и бесконечное, число подмножеств (классов) $X_1, \ldots, X_l$. Информация о подмножествах $X_1, \ldots, X_l$ задана с помощью конечной выборки X^0, содержащей объекты из каждого класса X_i. Требуется, пользуясь только выборкой X^0, указать алгоритм A (может быть, наилучший в некотором смысле), определенный на всем множестве X, результат работы которого для каждого $x \in X$ можно интерпретировать в терминах принадлежности классам X_i.

Задача распознавания в приведенной постановке является индуктивной по построению, и все алгоритмы, используемые для ее решения, будут эвристическими. Вместе с тем в задачах принятия решений даже в условиях индуктивности необходимо, чтобы результат был в максимальной степени строгим. Для этого в ТРО разработана методология, характерная именно для индуктивных задач. Ее суть заключается в следующем: вначале с помощью эвристических алгоритмов получается возможное решение, которое затем улучшается допустимыми математическими средствами. Это позволяет как минимум существенно упростить требования к алгоритмам, а как максимум повысить качество финального решения.

На рис. 1 приведена общая схема построения многоуровневых моделей распознавания для случая, когда число уровней равно двум. Обобщение для большего числа уровней очевидно.

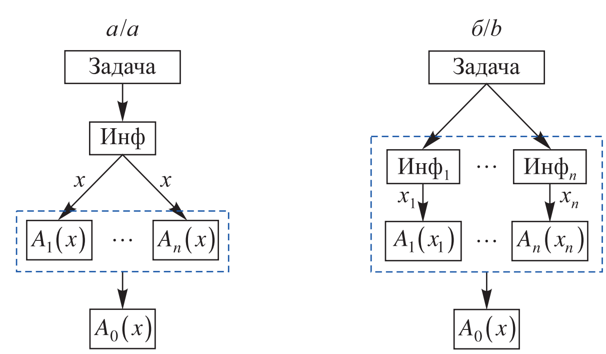

Рис. 1. Общая схема построения двухуровневых моделей распознавания (Инф – информация): *а* – вариант 1; *б* – вариант 2

Fig. 1. General scheme for constructing two-level recognition models: *a* – variant 1; *b* – variant 2

Принципы построения моделей в представленных двух вариантах различаются. В варианте 1 (см. рис. 1, *а*) строится суперпозиция алгоритмов $A_1, \ldots, A_n$ и некоторых отображений, результатом чего является алгоритм A_0. Все алгоритмы работают с одним входным объектом x. В варианте 2 (см. рис. 1, *б*) информация о задаче структурируется (декомпозируется). Каждый алгоритм A_i работает со своим входным объектом x_i, а алгоритм A_0 применяется для синтеза результата на входном объекте x, который является суперпозицией объектов $x_1, \ldots, x_n$.

Предполагается рассмотреть варианты построения многоуровневых (двухуровневых) алгоритмов распознавания в контексте задачи принятия решений.

Многоуровневые модели корректировки

Активное использование варианта 1 (см. рис. 1, *а*) в ТРО было инициировано Ю. И. Журавлевым [7]. Смысл его предложения был очень простым. Поскольку эвристические алгоритмы являются плохо управляемыми, то, получив с их помощью на первом шаге хоть какое-то решение, на следующем шаге можно скорректировать его так, чтобы итоговое решение было лучше любого из исходных решений, полученных эвристическими алгоритмами. Тем самым управление решением передавалось на второй либо последующие шаги, а требования к эвристическим алгоритмам существенно понижались. Также было замечено, что любой алгоритм распознавания можно рассматривать как суперпозицию распознающего оператора и решающего правила. Вследствие этого исследования таких схем стали развиваться в двух направлениях. В рамках первого направления предполагалось, что набором алгоритмов $A_1, \ldots, A_n$ доставляется финальный результат, который формируется на базе пространства $\mathbb{B}_2 = \{0, 1\}$ либо $\mathbb{B}_3 = \{0, 1, 2\}$. Данное направление стало называться логической корректировкой, так как для построения алгоритма A_0 использовалась алгебра логики [8]. В рамках второго направления в качестве алгоритмов $A_1, \ldots, A_n$ применялись распознающие операторы, которые в общем случае формируют результат на базе пространства действительных чисел $\mathbb{R}$. Благодаря этому для построения алгоритма A_0 на уровне распознающих операторов можно использовать обычные алгебраические операции, а далее остается достроить суперпозицию с решающим правилом. Такое направление называется алгебраической корректировкой.

Уточним постановку задачи распознавания, а также введем некоторые дополнительные обозначения. Уточнение постановки связано со свойством алгоритма «может быть, наилучший в некотором смысле». Дело в том, что в задаче имеется только конечная выборка объектов $X^0 \subset X$. При условии, что множество X разбито на подмножества $X_1, \ldots, X_l$, информация $P(x) = (P_1(x), \ldots, P_l(x))$ считается заданной для всех $x \in X^0$. Здесь $P_i(x) \in \{0, 1\} \wedge (P_i(x) = 1 \Leftrightarrow x \in X_i)$ – обычный предикат, характеризующий принадлежность классам X_i. Так как система предикатов $P(x)$ определена только для объектов $x \in X^0$, то и наилучший в любом смысле алгоритм можно определить лишь на этих объектах. В итоге выборка объектов $X^0 \subset X$ обычно разбивается на две выборки – обучающую ($X_S^0 \subset X$, $\left|X_S^0\right| = m$, $\forall i \in \{1, \ldots, l\}$, $X_i \cap X_S^0 \neq \varnothing$) и контрольную ($X_C^0 \subset X$, $\left|X_C^0\right| = q$, $\forall i \in \{1, \ldots, l\}$, $X_i \cap X_C^0 \neq \varnothing$). Эти выборки удовлетворяют условию $X_S^0 \cap X_C^0 = \varnothing$, $X_S^0 \cup X_C^0 = X^0$. Теперь свойство «может быть, наилучший в некотором смысле» подразумевает, что алгоритм A должен доставлять результат, наилучший для выборки X_C^0. С учетом этого уточним постановку задачи распознавания следующим образом: требуется, пользуясь только выборкой X_S^0, указать алгоритм A, являющийся наилучшим для контрольной выборки X_C^0 и определенный на всем множестве X, результат работы которого для каждого $x \in X$ можно интерпретировать в терминах принадлежности классам X_i.

Задачу, которая решается в такой постановке, обозначим через $Z = \left(X_S^0, X_C^0\right)$. Нетрудно заметить, что каждой выборке $X_C^0 = \{x_1, \ldots, x_q\}$ можно поставить в соответствие матрицу $I\left(X_C^0\right) = \left[P(x_1), \ldots, P(x_q)\right] \in \mathbb{B}_2^{ql}$. Назовем ее информационной матрицей для задачи Z. Теперь алгоритм распознавания можно определить как отображение вида $\forall x \in X,\ A : x \times X_S^0 \to \left(P_1^A(x), \ldots, P_l^A(x)\right)$, где $P_i^A(x) \in \mathbb{B}_2 \vee \mathbb{B}_3$. Но независимо от выбора пространства результатов в последнем случае выборке X_C^0 сопоставим матрицу $I^A\left(X_C^0\right) = \left[P^A(x_1), \ldots, P^A(x_q)\right] \in \mathbb{B}_2^{ql} \vee \mathbb{B}_3^{ql}$. Здесь через $P^A(x) = \left(P_1^A(x), \ldots, P_l^A(x)\right)$ обозначена система предикатов, полученная с помощью алгоритма A. В свою очередь, матрица $I^A\left(X_C^0\right)$ может быть названа алгоритмической матрицей для задачи Z. Перейдем непосредственно к рассмотрению результатов различных видов корректировки.

Логическая корректировка. Зафиксируем задачу Z и набор эвристических алгоритмов $A_1, \ldots, A_n$ ($n \in \mathbb{N}$). Результатом решения задачи Z алгоритмом A_i будет матрица $I^{A_i}\left(X_C^0\right) \in \mathbb{B}_2^{ql} \vee \mathbb{B}_3^{ql}$, которая в силу эвристичности алгоритма A_i может не совпадать с матрицей $I\left(X_C^0\right)$. Предположим, что любая оптимизация алгоритма A_i, которая привела бы к уменьшению указанного несовпадения, либо невозможна, либо избыточно трудозатратна. В таких условиях уместно ввести некоторую функцию $f:\left(\mathbb{B}_3^{ql}\right)^n \to \mathbb{B}_3^{ql}$. Если при этом потребовать, чтобы f была инвариантна относительно размерности задачи (величин q, l, числа алгоритмов n), а также строилась только с помощью логических операций (дизъюнкции ($\vee$), конъюнкции ($\wedge$), отрицания ($\neg$)), то вполне уместно говорить, что с помощью f можно корректировать результаты $A_1\left(X_C^0\right), \ldots, A_n\left(X_C^0\right)$. Соответствующий алгоритм A_f называется логическим корректором (на рис. 1 он обозначен через A_0). Заметим также, что аналогичную функцию f можно ввести и для пространства $\mathbb{B}_2^{ql}$. Во всех случаях приходим к следующей задаче: необходимо построить логический корректор A_f, который улучшает результаты работы эвристических алгоритмов, и при этом снизить сложность его реализации.

Результаты решения такой задачи в полном объеме можно найти в работе [9]. Ниже излагается только их суть и дается краткий анализ.

Первый результат касается построения оптимального по качеству логического корректора. Для этого из матриц $A_1\left(X_C^0\right), \ldots, A_n\left(X_C^0\right)$ строится специальная матрица T, столбец i которой соответствует последовательности наборов строк матрицы $A_i\left(X_C^0\right)$. Для оценки качества алгоритма вводится функционал качества $\varphi:\mathbb{B}_3^{ql} \times \mathbb{B}_2^{ql} \to \mathbb{R}$, который вычисляется покомпонентно следующим образом: $\varphi(0,0)=\varphi(1,1)=0$, $\varphi(0,1)=\varphi_{01}$, $\varphi(1,0)=\varphi_{10}$, $\varphi(2,0)=\varphi_{20}$, $\varphi(2,1)=\varphi_{21}$. Корректирующую функцию f предлагается строить как набор строк матрицы T. Доказано, что в этом наборе с помощью простых замен можно осуществить конструктивное построение f^*, для которого $\varphi\left(I^{A_{f^*}}\left(X_C^0\right), I\left(X_C^0\right)\right) \to \min$. Также показано, что сложность построения алгоритма A_{f^*} не превосходит $O\left(q^2 l^2\right)$.

Второй результат касается реализации функций f на базе формул трехзначной логики. Для этого вначале определяются функции, аналогичные дизъюнктивным нормальным формам (ДНФ) двузначной логики. Для построения f используется матрица T. В этом случае f можно рассматривать как не всюду определенную функцию трехзначной логики на решетке E_n^3. В итоге задачу построения корректора можно свести к продолжению f на E_n^3. Последнюю задачу предлагается рассматривать в классе нормальных канонических форм [10]. Построение искомой функции f осуществляется в несколько этапов. Вначале с помощью матрицы T строятся ДНФ, соответствующие различным областям истинности. Затем для продолжения на решетке E_n^3 осуществляется специальная операция склеивания построенных ДНФ. И наконец, для полученной функции строится сокращенная (тупиковая) ДНФ, что позволяет существенно снизить сложность итогового алгоритма A_f.

Необходимо отметить еще два обстоятельства. Во-первых, любое теоретическое исследование в такой области, как ТРО, было бы неполноценным без подкрепления практическими результатами. В этом смысле исследование логической корректировки можно считать вполне полноценным. Описанные выше алгоритмы A_f использовались при решении практических задач. Полученные при этом результаты подтвердили возможность применения таких алгоритмов на практике. Более подробно с данными выводами можно ознакомиться в работе [11]. Во-вторых (и это, может быть, самое главное), алгебры $\mathfrak{B}_2 = \left\langle \mathbb{B}_2^{ql}, \wedge, \vee, \neg \right\rangle$ и $\mathfrak{B}_3 = \left\langle \mathbb{B}_3^{ql}, \wedge, \vee, \neg \right\rangle$ не являются конечно представимыми и, как следствие, не содержат базис. По этой причине логические корректоры не обладают всеми необходимыми свойствами, обеспечивающими существование алгоритмов, точных для задачи Z. В лучшем случае удается построить оптимальный алгоритм, что не позволяет ничего сказать о разрешимости задачи в глобальном смысле.

Алгебраическая корректировка. Смысл алгебраической корректировки также очень простой. Алгоритм $A: x \times X_S^0 \to \left(P_1^A(x), \ldots, P_l^A(x)\right)$, который плохо поддается корректировке логическими функциями f, заменяется на новый. Для этого вводится более «богатое» по сравнению с $\mathbb{B}_3^l$ ($\mathbb{B}_2^l$) пространство Y (как правило, $Y = \mathbb{R}^l$), но с одним ограничением: должно существовать отображение $c: Y \to \mathbb{B}_3^l$ ($\mathbb{B}_2^l$) (называемое обычно решающим правилом), которое является непротиворечивым и допускает существование корректных алгоритмов таких, что $I^A\left(X_C^0\right) = I\left(X_C^0\right)$. Отображение $B: x \times X_S^0 \to Y$, являющееся частью исходного алгоритма A, называется обычно распознающим оператором. В работе [7] показано, что всякий алгоритм A допускает представление в виде суперпозиции распознающего оператора B и корректного решающего правила c (для этого отображение c должно быть просто сюръективным).

Введем в пространстве Y операторные корректоры $g: Y \to Y$, подходящая суперпозиция которых с наборами распознающих операторов и образует искомые алгебры. В работах [7; 8] было показано, что для большинства известных эвристических моделей построение корректных алгоритмов может быть осуществлено в рамках линейной алгебры. Правда, к таким моделям предъявлялись достаточно серьезные требования: существование набора операторов B, которые на контрольной выборке размерности q образуют базис пространства матриц $(Y)^q \overset{\text{def}}{=} \mathbb{R}^{ql}$. Таким образом, в рамках данного направления можно сформулировать следующую задачу: необходимо указать отличные от линейных методы построения операторных корректоров g, обеспечивающих существование (а еще лучше построение в явном виде) корректных алгоритмов в модели A_g, а также определить, как это сделать наиболее простым с вычислительной точки зрения способом и, главное, что это дает для разрешимости задачи Z.

Более подробное описание соответствующих результатов можно найти в работе [11]. Ниже приводится их краткий обзор. Итак, начнем с того, что необходимо решить задачу Z, причем указать в модели A_g корректный алгоритм такой, что $I^A\left(X_C^0\right) = I\left(X_C^0\right)$. Зафиксируем решающее правило c в пространстве $\mathbb{R}^{ql}$ и построим область $R_c(Z)$ такую, что $\forall b \in \mathbb{B}_2^{ql},\ \exists r \in R_c(Z): c(r) = b$. Последнее условие гарантируется сюръективностью c, а то, что оно приводится только для $\mathbb{B}_2^{ql}$, не ограничивает общности изложенных ниже результатов. Теперь можно сформулировать первый результат: для корректности модели A_g на Z необходимо и достаточно, чтобы $A_g\left(X_C^0\right) \cap R_c(Z) \neq \varnothing$. Этот простой критерий корректности дает возможность свести исходную задачу к исследованию области $R_c(Z)$. Данная область образует выпуклое подмножество в пространстве $\mathbb{R}^{ql}$, что в классе линейных пороговых решающих правил c позволяет рассматривать ее как решение системы нестрогих линейных неравенств. Множество таких решений может быть охарактеризовано в терминах отделимости специальным образом построенных подпространств. Для их описания потребуется следующее обозначение. Разобьем множество индексов $I = \{1, \ldots, q\} \times \{1, \ldots, l\}$ на два ($t \in \{0, 1\}$) подмножества: $M_t = \left\{(i, j) \middle| (i, j) \in I, P_j(x_i) = t\right\}$. Такая возможность существует в силу введенных выше ограничений. Теперь критерий корректности можно сделать более конструктивным: для корректности модели A_g на Z необходимо и достаточно, чтобы

$$\exists B \in g(B_1, \ldots, B_m): \left(\min_{(i,\, j) \in M_1} \left\{ \left\langle B\left(X_C^0\right), E_{ij} \right\rangle \right\} > c_0 \geq \max_{(i,\, j) \in M_0} \left\{ \left\langle B\left(X_C^0\right), E_{ij} \right\rangle \right\} \right),$$

где $g(B_1, \ldots, B_m)$ – суперпозиция отображений g и набора распознающих операторов $B_1, \ldots, B_m$ ($m \in \mathbb{N}$); E_{ij} – канонический базис в пространстве $\mathbb{R}^{ql}$; $\left\langle B\left(X_C^0\right), E_{ij} \right\rangle$ – скалярное произведение матриц в пространстве $\mathbb{R}^{ql}$; c_0 – параметр линейного порогового решающего правила c.

Введем два вида алгебр: $\mathfrak{A}_1 = \left\langle \mathbb{R}^{ql}, +, \circ \right\rangle$ (имеет тип $\langle 2, 1\rangle$; символ $\circ$ обозначает унарную операцию умножения матрицы на скаляр) и $\mathfrak{A}_2 = \left\langle \mathbb{R}^{ql}, +, \times, \circ \right\rangle$ (имеет тип $\langle 2, 2, 1\rangle$). Обе алгебры являются коммутативными относительно введенных операций, т. е. операции выполняются покомпонентно. Дело в том, что в общем случае $q \neq l$, и поэтому определить иначе коммутативные алгебры невозможно. Нетрудно заметить, что алгебры $\mathfrak{A}_1$ и $\mathfrak{A}_2$ представляют собой модели для построения искомой суперпозиции. Для алгебры $\mathfrak{A}_1$, линейной по построению, без труда можно определить условия полноты (так как она является конечно представимой и содержит базис) и корректности. Фактически эти условия повторяют полученные в работе [8] результаты. Что касается алгебры $\mathfrak{A}_2$, то это уже полиномиальная модель. Для нее могут быть получены более интересные результаты. Полнота для алгебры $\mathfrak{A}_2$ также связана с существованием распознающих операторов B, обеспечивающих возможность алгебраического несовпадения для каждого элемента из множества I, а вот корректность определяется условием, позволяющим отличать в алгебраическом смысле элементы из множества M_1. Кроме того, показано, что в алгебре $\mathfrak{A}_3 = \left\langle \mathbb{R}^{ql}, +, \circ^{-1}, \circ \right\rangle$ (имеет тип $\langle 2, 1, 1\rangle$; символ $\circ^{-1}$ обозначает операцию коммутативного обращения матриц) условия корректности могут быть получены в линейном замыкании [11].

Еще один вид операторных корректоров можно построить, основываясь на изоморфизме пространства $\mathbb{R}^{ql}$ и тензорного произведения $\mathbb{R}^q \otimes \mathbb{R}^l$. Для этого можно использовать билинейное отображение $h: \mathbb{R}^q \times \mathbb{R}^l \to \mathbb{R}^{ql}$ (билинейным оно названо, так как является линейным по каждой из двух переменных – классам и объектам контрольной выборки). Детали построения таких корректоров изложены в работах [9; 12]. Полученную в итоге билинейную модель обозначим через $\mathfrak{A}_4$ (в данном случае не совсем уместно использовать алгебраическую нотацию, так как здесь скорее можно вести речь о суперпозиции алгебр). Для нее основной результат сформулируем следующим образом: в $\mathfrak{A}_4$ корректный алгоритм A можно построить, если существует распознающий оператор B такой, что строки матрицы $B\left(X_C^0\right)$ не совпадают для всех объектов из X_C^0, принадлежащих различным классам. Для модели $\mathfrak{A}_4$ также определено условие полноты и предложен явный вид корректного алгоритма A.

Полученные для алгебраической корректировки результаты позволили расширить представление о разрешимости задачи распознавания в целом. Ранее при рассмотрении задачи оптимизации для логической корректировки вводились специальные функционалы качества φ. Теперь сделаем небольшое уточнение для данных функционалов в терминах предикатов $P(x)$ и $P^A(x)$. Для этого зафиксируем некоторый алгоритм A и предположим, что ему можно сопоставить подмножества $X_i^A = \left\{x \in X : P_i^A(x) = 1\right\}$. Почти очевидно, что $X_i \neq X_i^A$ для всех $i \in \{1, \ldots, l\}$, но тем не менее $X^A = \bigcup_{i=1}^{l} X_i^A = \bigcup_{i=1}^{l} X_i = X$. Тогда любая монотонная функция $\varphi_A : X \times X^A \to [0, 1]$, удовлетворяющая условиям $\varphi_A\left(X, X^A\right) = 1$, если $X = X^A$, и $\varphi_A\left(X, X^A\right) = 0$, если $X \cap X^A = \varnothing$, представляет собой функционал качества алгоритма A. Очевидно, что задача Z может быть названа разрешимой, если существует такой алгоритм A, для которого $\varphi_A\left(X, X^A\right) = 1$. В задаче распознавания все множество X является недоступным, имеется только конечная выборка объектов $X^0 \subset X$. Более того, качество алгоритма A можно оценивать только на контрольной выборке X_C^0. Исходя из этого, если говорить о разрешимости задачи Z, то интерес представляет связь между $\varphi_A\left(X, X^A\right)$ и $\varphi_A\left(X_C^0, X_C^{0A}\right)$. Такая связь также почти очевидна: $\forall A \left(\varphi_A\left(X, X^A\right) < 1 \Leftrightarrow \exists X_C^0, \varphi_A\left(X_C^0, X_C^{0A}\right) < 1\right)$. Непосредственным следствием данной эквивалентности будет такое утверждение: корректность алгоритма A является необходимым условием разрешимости задачи. С использованием этой зависимости можно получить целый ряд достаточных условий разрешимости, а также условия на структуру контрольной выборки X_C^0, метод корректировки для построения пространства X.

Сформулируем теперь некоторые выводы из приведенного описания способов решения задачи распознавания. Первый вывод: для логической и алгебраической корректировок достаточно двухуровневой схемы. Для логической корректировки это обусловливается тем, что соответствующие алгебры не являются конечно представимыми и, как следствие, не содержат базис. Для алгебраической корректировки дело обстоит несколько иначе. В данном случае может быть сформулирован в некотором смысле предельный результат, который стал следствием развития такого подхода. Суть этого результата очень проста: если объекты контрольной выборки не совпадают, а эвристический алгоритм не является невырожденным, то в билинейном замыкании единственного распознающего оператора всегда может быть конструктивно построен корректный (точный для заданной выборки) алгоритм. Отсюда почти очевидно такое утверждение: если корректный алгоритм невозможно построить в двухуровневой модели, то его невозможно построить в принципе для рассматриваемой задачи. Второй вывод касается постановочной сути задачи. В процессе решения любой задачи приходится ограничиваться рассмотрением только корректных алгоритмов, что обеспечивает лишь необходимые условия разрешимости задачи распознавания на всем множестве допустимых объектов. Любые другие предположения относительно структуры информации, свойств алгоритмов и прочих параметров носят исключительно теоретический характер и не оказывают влияния на индуктивность задачи. И наконец, третий вывод имеет в большей степени практический характер. За пределами контрольной выборки X_C^0 алгебраический корректор A_0 вел себя не совсем адекватно. Когда речь шла о нелинейных замыканиях (степенные корректоры, билинейные операторы [11]), то соответствующие алгоритмы отказывались от распознавания даже при небольших отклонениях от объектов контрольной выборки. Можно предположить, что при использовании линейных корректоров [7] алгоритмы были бы лучше, но для них, к сожалению, нет явного вида алгоритмов. Последнее обстоятельство затрудняет практическую проверку линейных корректоров.

Многоуровневые модели, основанные на структурировании информации

Перейдем теперь к обсуждению варианта 2, представленного на рис. 1, *б*. В этом варианте подразумевается, что решение задачи осуществляется через структурирование информации. Изложенные ниже результаты можно разделить на два направления. В первом из них структурируется выборка объектов X^0, а во втором – признаковое пространство, в котором определено множество X. Сразу отметим, что в первом случае решение полностью соответствует схеме, приведенной на рис. 1, *б*. Во втором случае требуется незначительное уточнение схемы, которое представлено на рис. 2. Стоит упомянуть, что само по себе структурирование информации играет несущественную роль, обычно оно осуществляется для достижения одной из следующих целей. Первая цель – это уменьшение сложности решения исходной задачи через ее разбиение на подзадачи. Такое разбиение обычно называется декомпозицией. Вторая цель, которая здесь также достижима, – это синтез алгоритма как результат комбинирования алгоритмов, решающих отдельные подзадачи в рамках построенной структуры. Эти подзадачи двойственны и могут рассматриваться в рамках единой методологии. Ниже приводится интерпретация результатов в контексте достижения второй цели.

Синтез алгоритмов на основе структурирования выборок. Полное описание результатов, представленных далее, можно найти в работе [13]. Отметим только, что там рассмотрение ведется исходя из достижения первой цели, т. е. речь идет о декомпозиции задачи распознавания. Пусть задана некоторая задача $Z = \left(X_S^0,\ X_C^0\right)$. Предположим, что выборки X_S^0 и X_C^0 можно разбить на конечное число $t \in \mathbb{N}$ подмножеств: $X_S^0 = \left\{X_S^{0,1},\ \ldots,\ X_S^{0,t}\right\}$ и $X_C^0 = \left\{X_C^{0,1},\ \ldots,\ X_C^{0,t}\right\}$. Допустим, что между этими выборками можно установить соответствие, которое представим в виде матрицы

$$\begin{array}{cc} \begin{array}{ccc} X_C^{0,1} & \cdots & X_C^{0,t} \end{array} & \\ \begin{bmatrix} \left(X_S^{0,1},\ X_C^{0,1}\right) & \cdots & \left(X_S^{0,1},\ X_C^{0,t}\right) \\ \vdots & \ddots & \vdots \\ \left(X_S^{0,t},\ X_C^{0,1}\right) & \cdots & \left(X_S^{0,t},\ X_C^{0,t}\right) \end{bmatrix} & \begin{array}{c} X_S^{0,1} \\ \vdots \\ X_S^{0,t} \end{array} \end{array}.$$

Каждый элемент данной матрицы, соответствующий некоторым столбцу $X_C^{0,j}$ и строке $X_S^{0,i}$, можно рассматривать как подзадачу $Z_{ij} = \left(X_S^{0,i},\ X_C^{0,j}\right)$. Полное описание всего множества таких задач получается в том случае, если имеется набор подзадач $Z_{ii},\ i \in \{1,\ \ldots,\ l\}$. По нему же восстанавливается исходная задача Z. Таким образом, при описании свойств, которыми должно обладать разбиение, можно ограничиться свойствами набора из t подзадач. Показано, что при выполнении условий $X_S^{0,i} \cap X_S^{0,j} = \varnothing$, $X_C^{0,i} \cap X_C^{0,j} = \varnothing\ \forall i \neq j$ для большинства известных моделей распознавания решения подзадач Z_{ii} соответствующими алгоритмами A_i совпадут с решением задачи Z алгоритмом $A = A_1 \oplus \ldots \oplus A_t$ (символ $\oplus$ обозначает операцию прямого суммирования), при этом существует условие, которому каждый алгоритм A_i должен удовлетворять на подзадачах Z_{ui} при $u \neq i$, и оно является необходимым и достаточным.

Условия, о которых идет речь выше, хотя и просты, но не очень конструктивны, так как в большей степени носят существовательный характер. По этой причине проверить их выполнение на практике непросто. Чтобы избежать подобных затруднений, определены еще два достаточных условия, при выполнении которых на наборе задач искомый алгоритм может быть получен также прямым суммированием алгоритмов A_i. Для формулировки первого условия введем понятие линейной оболочки $l\left(X_S^{0,i} \bigcup X_C^{0,i}\right)$ в пространстве X. Оказывается, если подзадачи $Z_{ii},\ i \in \{1,\ \ldots,\ l\}$, удовлетворяют условию $\forall i,\ j \in \{1,\ \ldots,\ t\}$, $l\left(X_S^{0,i} \bigcup X_C^{0,i}\right) \cap l\left(X_S^{0,j} \bigcup X_C^{0,j}\right) = \varnothing$ при $i \neq j$, то результат будет аналогичен сформулированному выше в тех же самых моделях и с теми же последствиями для синтеза алгоритмов. Для определения второго достаточного условия введем понятие характеристического вектора. Пусть $R_i \subset \mathbb{R}$ $(i = \{1,\ \ldots,\ n\})$, $R_1 \times \ldots \times R_n \subset \mathbb{R}^n$. Отображение $\gamma_{(R_1,\ \ldots,\ R_n)}(x) = \left(\gamma_{R_1}(x_1),\ \ldots,\ \gamma_{R_n}(x_n)\right)$, определенное для всех $x = (x_1,\ \ldots,\ x_n) \in \mathbb{R}^n$ на основании правила $\left(\gamma_{R_i}(x) = 1,\ x \in R_i\right) \wedge \left(\gamma_{R_i}(x) = 0,\ x \notin R_i\right)$, назовем характеристическим вектором. По аналогии с линейной оболочкой можно определить характеристическую оболочку: $\gamma_{(R_i,\ \ldots,\ R_n)}\left(X_S^{0,i} \bigcup X_C^{0,i}\right)$. В этом случае легко доказать, что если для всех подзадач выполняется условие $\forall i,\ j \in \{1,\ \ldots,\ t\}$, $\gamma_{(R_1,\ \ldots,\ R_n)}\left(X_S^{0,i} \bigcup X_C^{0,i}\right) \cap \gamma_{(R_1,\ \ldots,\ R_n)}\left(X_S^{0,j} \bigcup X_C^{0,j}\right) = \varnothing$ при $i \neq j$, то результат снова будет аналогичен сформулированному выше.

Таким образом, для реализации алгоритмов по варианту 2 потребуется построить совокупность подзадач $Z_{ii},\ i \in \{1,\ \ldots,\ l\}$, проверить выполнение одного из трех условий (указанных выше), а затем синтезировать алгоритм в виде $A = A_1 \oplus \ldots \oplus A_t$. При выполнении всех условий решение (и даже корректность) сохраняется. Но во всех случаях алгоритмы остаются эвристическими.

Синтез алгоритмов на основе структурирования признакового пространства. Необходимость в подобных алгоритмах часто возникает при решении задач медицинской диагностики [14]. Для таких задач часть информации о диагнозе, как правило, берется из литературы в виде логических построений (правил), а часть информации – из личного опыта врача. В последнем случае информация обычно не формализована и может быть представлена в виде примеров (прецедентов). Таким образом, имеем дело с вариантом 2. Для установления диагноза, когда информация представлена по правилам, используется алгоритм A_1. Традиционно в качестве такого алгоритма применяется метод (алгоритм) резолюций Робинсона, который реализует логический вывод. В свою очередь, для диагностики по примерам целесообразно использовать подходящий алгоритм распознавания A_2. Но здесь могут возникнуть проблемы. Для одного и того же пациента алгоритм A_1 по части описания, соответствующей логической информации, может предложить одно решение, а алгоритм A_2 по другой части описания может предложить иное решение. Возникает вопрос, нельзя ли объединить решения $A_1(x)$ и $A_2(x)$, чтобы получить интегрированный результат $A_0(x)$.

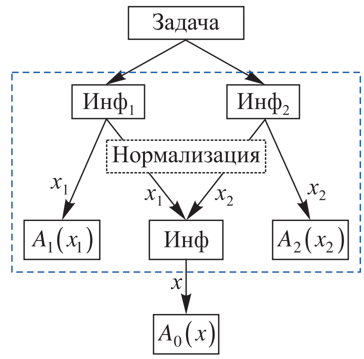

Рис. 2. Схема для задачи со структуризацией признакового пространства

Fig. 2. Scheme for a problem with structuring feature space

Предлагаемый подход к решению указанной проблемы основан на модификации двухуровневой модели в варианте 2. Схематично это решение представлено на рис. 2.

Возможность такого объединения информации возникает по причине того, что в предметной области часто можно выполнить бинеаризацию признакового пространства. В данном случае правила представляются в виде формул алгебры логики, а прецедентная информация рассматривается как векторы в пространстве $\mathbb{B}_2^n$ ($n \in \mathbb{N}$). По этой причине предлагается вначале построить специальный алгоритм распознавания, ориентированный на работу в таком пространстве. Для его построения можно использовать меру прецедентности $\mu : \mathbb{B}_2^n \times \mathbb{B}_2^n \to [-1, 1]$ как функцию, удовлетворяющую следующим условиям:

$$\forall x_1, x_2 \in \mathbb{B}_2^n \begin{cases} \mu(x_1, x_2) = 1 \Leftrightarrow x_1 = x_2, \\ \mu(x_1, x_2) = \mu(x_2, x_1), \\ \mu(x_1, x_2) + \mu(x_1, \bar{x}_2) = 0, \end{cases}$$

где $\bar{x}_2$ – логическое отрицание объекта x_2. Введем обозначения $I \overset{\text{def}}{=} \{1, \ldots, n\}$, $L \overset{\text{def}}{=} \{1, \ldots, l\}$ и некоторую матрицу $\|a_{ij}\| \in \mathbb{R}^{ln}$. Будем полагать, что элементы данной матрицы удовлетворяют следующим условиям: $\forall i, j,\ a_{ij} \geq 0$; $\forall i, \sum\limits_j a_{ij} > 0$. И наконец, так как речь идет о задаче распознавания, будем считать, что заданы все ее компоненты в виде выборок X^0, X_S^0, X_C^0. В результате проведенных рассуждений алгоритм A_2 можно рассматривать как отображение $\forall x \in X,\ A : x \times X_S^0 \to \left(P_1^A(x), \ldots, P_l^A(x)\right)$ и описать в виде последовательности следующих шагов.

Шаг 1. Фиксируем некоторый объект $x \in X$.

Шаг 2. Для каждого $i \in L$ и для всех $x_v \in X_{S,\,i}^0$ вычисляем

$$\mu(x, x_v) = \left(\sum_{j=1}^{n} a_{ij}\right)^{-1} \times \left(\sum_{j=1}^{n} (-1)^u \times a_{ij}\right),$$

где

$$u = \begin{cases} 1, \text{ если } x_j \neq x_{vj}, \\ 2, \text{ если иначе.} \end{cases}$$

Шаг 3. Если все $x_v \in X_{S,\,i}^0$ исчерпаны, то вычисляем

$$P_i^A(x) = \max_{x_v \in X_{S,\,i}^0} \{\mu(x, x_v)\}.$$

Если номера классов $i \in L$ не исчерпаны, то выбираем очередной класс $i \in L$ и возвращаемся к шагу 2. В противном случае алгоритм заканчивает работу.

Нетрудно заметить, что описанный алгоритм A однозначно связан с параметрами $\|a_{ij}\|$, выбор и способ подсчета которых определяется в основном внеалгоритмическими соображениями: выбором пространства описания объектов из множества X, желанием придать числам $a_{ij} \in \mathbb{R}$ и величинам $P_i^A(x)$ некоторую содержательную интерпретацию, способом формализации понятия прецедентности объектов и т. п. Для лучшего понимания задачи и анализа исходной информации величины $P_i^A(x)$ желательно интерпретировать в терминах предметной области, поэтому выбор и способ подсчета параметров $\|a_{ij}\|$ имеют существенное значение. Хотя стоит заметить, что работа алгоритма A от этого зависит несильно. К примеру, их значения можно положить равными единице, что для задачи распознавания непринципиально.

Опишем одну из возможных схем определения параметров $\|a_{ij}\|$.

Шаг 1. Фиксируем номер класса $i \in L$.

Шаг 2. Для всех признаков $j \in \{1, \ldots, n\}$ вычисляем

$$b_{ij} = (m_i)^{-1} \left(\sum_{x_\nu \in X_{S,\,i}^0} x_{\nu j} \right),$$

где $x_{\nu j}$ – значение признака j в векторе $x_\nu \in X_{S,\,i}^0$.

Шаг 3. Выполняем шаги 1 и 2 до тех пор, пока все номера классов и все признаки в каждом классе не будут исчерпаны. Затем переходим к шагу 4.

Шаг 4. Для всех признаков $j \in I$ и классов $i \in L$ вычисляем

$$b_j = (l)^{-1} \left(\sum_{i=1}^{l} b_{ij} \right), \; a_{ij} = \left| b_{ij} - b_j \right|.$$

Отметим, что введение такого алгоритма A_2 позволило получить лишь часть решения. Для получения решения на всем объеме информации предлагается информацию из правил преобразовать к объектному представлению. Нетрудно показать, что в этом случае применимым является алгоритм A_2, при этом результаты его использования совпадают с результатами применения алгоритма резолюций к исходному множеству правил. Предлагается также провести процедуру нормализации, в ходе которой полученное множество объектов объединяется с выборкой X^0 (X_S^0). Для объединенной выборки доказано, что нормализация не меняет результатов, которые могли быть получены в случае обратного перехода к исходной информации алгоритмами $A_1(x)$ и $A_2(x)$.

В данном случае синтез заключается в доопределении алгоритма A_2 на множество объектов, полученных в ходе нормализации. С использованием построенной схемы были проведены тесты для данных, содержащихся в различных репозитариях машинного обучения. Кроме того, успешно решены практические задачи из области ортопедии и спортивной травматологии.

Заключение

В работе рассмотрен специальный класс задач принятия решений прецедентного типа, для решения которых, как правило, применяются эвристические алгоритмы. Показано, что данный класс задач сводится к стандартной задаче распознавания образов с обучением. Это позволяет вместо эвристических алгоритмов использовать многоуровневые модели, которые дают возможность повысить точность решения, а в некоторых случаях обосновать его правильность. Приведен анализ различных вариантов построения многоуровневых моделей.

Библиографические ссылки

1. Таха ХА. *Введение в исследование операций.* 7-е издание. Минько АА, переводчик. Москва: Издательский дом «Вильямс»; 2007. 912 с.

2. Мальцев АИ. *Алгебраические системы.* Москва: Наука; 1970. 392 с. (Современная алгебра).

3. Gödel K. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik.* 1931;38:173–198. DOI: 10.1007/BF01700692.

4. Laurière J-L. *Intelligence artificielle. Résolution de problèmes par l'homme et la machine.* Paris: Eyrolles; 1987. XI, 473 p.

5. Barwise J, editor. *Handbook of mathematical logic.* Amsterdam: North-Holland; 1977. XII, 1165 p. (Studies in logic and the foundations of mathematics; volume 90).

6. Nilsson NJ. *Learning machines: foundations of trainable pattern-classifying systems.* New York: McGraw-Hill; 1965. XI, 137 p. (McGraw-Hill series in systems science).

7. Журавлев ЮИ. Об алгебраическом подходе к решению задач распознавания или классификации. В: Яблонский СВ, редактор. *Проблемы кибернетики. Выпуск 33.* Москва: Физматгиз; 1978. с. 5–68.

8. Журавлев ЮИ. Экстремальные алгоритмы в алгебре над некорректными алгоритмами. *Доклады Академии наук СССР.* 1977;237(3):509–512.

9. Krasnoproshin VV, Obraztsov VA. The choice of algorithms to solve the pattern recognition problem. *Pattern Recognition and Image Analysis.* 1996;6(3):526–535.

10. Краснопрошин ВВ. Об оптимальном корректоре совокупности алгоритмов распознавания. *Журнал вычислительной математики и математической физики.* 1979;19(1):204–215.

11. Zhuravlev YuI, Ablameiko SV, Biryukov AS, Dokukin AA, Krasnoproshin VV, Obraztsov VA, et al. Algorithms for algebraic and logical correction and their applications. *Pattern Recognition and Image Analysis.* 2010;20(2):105–117. DOI: 10.1134/S10546618 1002001X.

12. Krasnoproshin VV, Obraztsov VA. Problems of solvability and choice of algorithms for decision making by precedence. *Pattern Recognition and Image Analysis.* 2006;16(2):155–169. DOI: 10.1134/S1054661806020027.

13. Краснопрошин В, Образцов В. Сложные задачи распознавания образов и возможности их решения. В: Марков К, Рязанов В, Иванова К, Митов И, редакторы. *Classification, forecasting, data mining.* София: ИТЕА; 2009. с. 69–75 (Information science and computing; number 8).

14. Krasnoproshin VV, Obraztsov VA, Popok SA, Vissia H. Decision-making in sports traumatology. In: Peris-Ortiz M, Álvarez-García J, Del Río-Rama M. *Sports management as an emerging economic activity: trends and best practices.* Cham: Springer; 2017. p. 207–219. DOI: 10.1007/978-3-319-63907-9_13.

References

1. Taha HA. *Operations research: an introduction.* 7th edition. Upper Saddle River: Prentice Hall; 2003. XVII, 830 p.

Russian edition: Taha HA. *Vvedenie v issledovanie operatsii.* 7th edition. Min'ko AA, translator. Moscow: Williams Publishing House; 2007. 912 p.

2. Mal'tsev AI. *Algebraicheskie sistemy* [Algebraic systems]. Moscow: Nauka; 1970. 392 p. (Sovremennaya algebra). Russian.

3. Gödel K. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik.* 1931;38:173–198. DOI: 10.1007/BF01700692.

4. Laurière J-L. *Intelligence artificielle. Résolution de problèmes par l'homme et la machine.* Paris: Eyrolles; 1987. XI, 473 p.

5. Barwise J, editor. *Handbook of mathematical logic.* Amsterdam: North-Holland; 1977. XII, 1165 p. (Studies in logic and the foundations of mathematics; volume 90).

6. Nilsson NJ. *Learning machines: foundations of trainable pattern-classifying systems.* New York: McGraw-Hill; 1965. XI, 137 p. (McGraw-Hill series in systems science).

7. Zhuravlev YuI. [An algebraic approach to solving recognition or classification problems]. In: Yablonskii SV, editor. *Problemy kibernetiki. Vypusk 33* [Problems of cybernetics. Issue 33]. Moscow: Fizmatgiz; 1978. p. 5–68. Russian.

8. Zhuravlev YuI. [Extremal algorithms in algebra over incorrect algorithms]. *Doklady Akademii nauk SSSR.* 1977;237(3):509–512. Russian.

9. Krasnoproshin VV, Obraztsov VA. The choice of algorithms to solve the pattern recognition problem. *Pattern Recognition and Image Analysis.* 1996;6(3):526–535.

10. Krasnoproshin VV. [On the optimal corrector of a set of recognition algorithms]. *Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki.* 1979;19(1):204–215. Russian.

11. Zhuravlev YuI, Ablameiko SV, Biryukov AS, Dokukin AA, Krasnoproshin VV, Obraztsov VA, et al. Algorithms for algebraic and logical correction and their applications. *Pattern Recognition and Image Analysis.* 2010;20(2):105–117. DOI: 10.1134/S10546618 1002001X.

12. Krasnoproshin VV, Obraztsov VA. Problems of solvability and choice of algorithms for decision making by precedence. *Pattern Recognition and Image Analysis.* 2006;16(2):155–169. DOI: 10.1134/S1054661806020027.

13. Krasnoproshin V, Obraztsov V. [Complex problems of pattern recognition and the possibilities of their solution]. In: Markov K, Ryazanov V, Ivanova K, Mitov I, editors. *Classification, forecasting, data mining.* Sofia: ITHEA; 2009. p. 69–75 (Information science and computing; number 8). Russian.

14. Krasnoproshin VV, Obraztsov VA, Popok SA, Vissia H. Decision-making in sports traumatology. In: Peris-Ortiz M, Álvarez-García J, Del Río-Rama M. *Sports management as an emerging economic activity: trends and best practices.* Cham: Springer; 2017. p. 207–219. DOI: 10.1007/978-3-319-63907-9_13.

Получена 29.06.2023 / исправлена 01.11.2023 / принята 03.11.2023.
Received 29.06.2023 / revised 01.11.2023 / accepted 03.11.2023.