

ПРОБЛЕМА ИЗВЛЕЧЕНИЯ ЗНАНИЙ С ИСПОЛЬЗОВАНИЕМ РАССУЖДЕНИЙ НА ОСНОВЕ ПРЕЦЕДЕНТОВ

А. В. КРЫЛОВ

*Санкт-Петербургский институт информатики и автоматизации РАН,
199178, Санкт-Петербург, Россия
E-mail: KrAlex98@yandex.ru*

Рассматриваются методы представления знаний в прецедентной базе знаний, а также поиска и извлечения прецедентов, схожих с текущей ситуацией, для решения задач проактивного управления с системами поддержки принятия решений. Анализ методов извлечения прецедентов из базы показал, что для систем поддержки принятия решений, имеющих большие базы прецедентов, наиболее рациональным представляется вариант комплексирования методов, основу которого составляет метод ближайших соседей, с использованием остальных методов в этом случае осуществляется предварительная фильтрация прецедентов в базе.

Ключевые слова: *экспертные системы, системы поддержки принятия решения, прецедентный подход, методы извлечения прецедентов, сложные организационно-технические объекты*

Введение. На современном этапе развития объекты управления принимают все более сложный, неоднозначный характер. При этом они постоянно модернизируются, становятся более универсальными и надежными, усложняется их организационная, техническая и особенно информационная составляющие, что в свою очередь, приводит и к усложнению процесса управления этими объектами [1].

Большое количество информации как о самих сложных организационно-технических объектах, так и об окружающей их среде затрудняет работу операторов соответствующей системы управления: они в состоянии мгновенно оценить все аспекты сложившейся обстановки и принять правильное, обоснованное решение в короткие сроки, особенно с учетом дальнейшего развития ситуации. При повышении требований к качеству управленческих решений этот процесс усложняется — чем выше ответственность за принимаемые решения, тем более качественной должна быть оценка. На помощь операторам в этом случае приходят системы поддержки принятия решения (СППР), позволяющие лицам, принимающим решение, осуществить полный и объективный анализ динамически изменяющейся обстановки, а также прогноз ее развития, особенно при решении слабоструктурированных задач.

Одно из направлений развития систем автоматизации процессов выработки управленческих решений связано с моделированием процессов человеческого мышления — искусственным интеллектом. Мощным инструментом в теории искусственного интеллекта считаются экспертные системы (ЭС), которые представляют класс компьютерных информационных систем, основанных на работе со знаниями [1—3].

Экспертные системы. Ядром экспертной системы является база знаний (БЗ) о предметной области. База знаний в экспертных системах содержит правила вида „если..., то...“; механизм вывода в них является интерпретатором базы знаний (или решателем).

Особенностями ЭС являются следующие факторы:

— благодаря разделению базы знаний и программных управляющих структур разработчики экспертной системы могут сосредоточиться на накоплении и организации знаний, а не на деталях компьютерной реализации;

— разделение знаний и системы управления позволяет вносить изменения в одну часть базы знаний без появления противоречий в другой;

— разделение знаний и управляющих элементов программы позволяет использовать в различных системах одни те же модули управления и интерфейса пользователя.

В начале XXI в. получили развитие экспертные системы 5-го поколения, обладающие большими ресурсами для хранения информации в своей базе знаний и способные хранить не только логические правила, но и результаты работы лиц, принимающих решение, с учетом информации об успешности результатов их деятельности. На основе этой информации была сформирована „база прецедентов“ (БП), а экспертные системы, использующие в составе базы знаний еще и базу прецедентов, называются „прецедентные экспертные системы“ [4].

Прецедентный подход. В общем понимании „*прецедент*“ можно определить как единичное знание о предыдущем опыте. При этом какую именно информацию содержит прецедент, зависит от предметной области и целей использования прецедента.

В этом случае могут применяться *рассуждения, базирующиеся на прецедентах* (Case-Based Reasoning — CBR), лежащие в основе метода формирования умозаключений, опирающегося не на логический вывод, диктуемый исходными предпосылками (логические рассуждения), а на поиск и анализ случаев формирования подобных умозаключений в прошлом [5].

Рассуждения на основе прецедентов базируются на накопленном опыте и последующей адаптации решения известной задачи к решению новой. Прецедентный подход позволяет упростить процесс принятия решений в условиях временных ограничений и при наличии различного рода неопределенности в исходных данных и экспертных знаниях, а также в случае возникновения различных нештатных (аномальных) ситуаций.

Прецедентный подход получил широкое распространение. Например, в сфере медицинской диагностики решение, предлагаемое экспертной системой, основывается не только на анатомии и физиологии и теоретических моделях диагнозов, но и на изучении истории болезней и врачебного опыта, приобретенного при лечении пациентов (системы CASEY, PROTOS [6]). Не менее широко прецедентный подход используется и в юриспруденции, где важным компонентом правовых рассуждений является идентификация прецедента для принятия решений в конкретном случае (системы Rissland, Rissland&Ashly [6]).

В общем случае модель представления прецедента включает:

- описание ситуации;
- решение для данной ситуации;
- результат применения решения [7].

Из описания модели представления прецедента возникают и возможности данного метода по прогнозированию результатов принятия решения.

Методы рассуждения на основе прецедентов включают четыре основных этапа, образующие так называемый CBR-цикл (рис. 1) [8].

Процесс использования прецедентов при решении новой задачи, согласно рис. 1, реализуется следующим образом.

1. При возникновении определенной задачи программным средством на основе методов поиска решений в базе прецедентов осуществляется поиск и извлечение наиболее соответствующего (подобного) сложившейся ситуации прецедента (или прецедентов) из базы. В зависимости от предметной области и используемого метода поиска применяются различные метрики сравнения текущей ситуации с ситуациями из БП.

2. Осуществляется повторное использование извлеченного прецедента для решения текущей задачи (проблемы) при наличии сходства.

3. При частичном сходстве осуществляется пересмотр и адаптация решения по выбранному прецеденту (или набору прецедентов) в соответствии с условиями текущей задачи.

4. Исходя из новых условий и адаптированного к ним решения по имеющемуся прецеденту (прецедентам) формируется новый прецедент, который сохраняется в базе.

При использовании прецедентного метода в ЭС возникают некоторые проблемы интерпретации данных и знаний, хранящихся в базе прецедентов. Наибольшие затруднения вызывают этап извлечения прецедентов и этап адаптации прецедента (прецедентов) к текущим условиям обстановки.

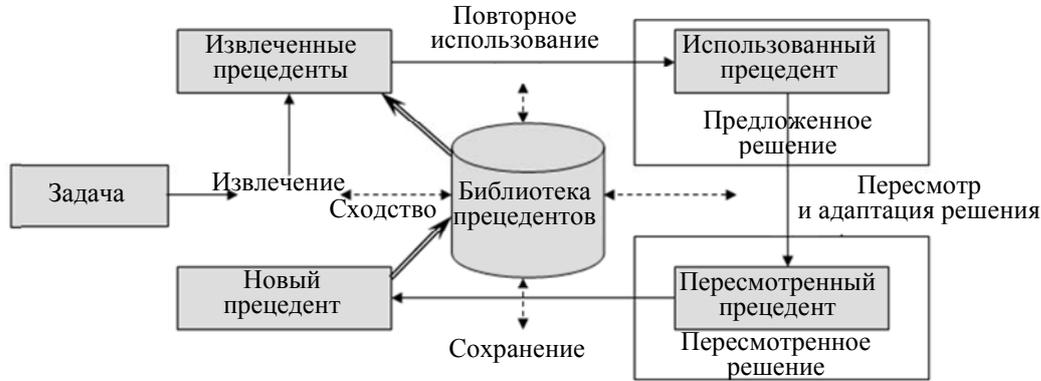


Рис. 1

Методы поиска решений в базе прецедентов. Проблема выбора подходящего прецедента является одной из самых важных в этих системах. Естественно поиск прецедента должен быть организован согласно цели его использования, что, в свою очередь, требует хороших знаний о предметной области.

Существует целый ряд методов поиска решений на основе прецедентов и их модификаций [8], например: метод ближайшего соседа; метод извлечения прецедентов на основе деревьев решений; метод извлечения прецедентов на основе знаний; метод извлечения с учетом применимости прецедента. Рассмотрим их более подробно.

1. *Метод ближайшего соседа* (Nearest Neighbor — NN) — это самый распространенный метод сравнения и извлечения прецедентов, позволяющий довольно легко определить степень сходства текущей проблемной ситуации с прецедентом из базы. Для определения степени сходства на множестве параметров, используемых для описания прецедентов и текущей ситуации, вводится определенная метрика. Далее в соответствии с выбранной метрикой определяется расстояние от целевой точки, соответствующей текущей проблемной ситуации, до точек, соответствующих прецедентам из БП, и выбирается точка, ближайшая к целевой [9].

Метод определения ближайшего соседа (ближайших соседей) также широко применяется для решения задач классификации, кластеризации и распознавания образов. Объяснения предлагаемой классификации представляются весьма логичными: объект u отнесен к классу U , потому что к этому же классу относится близкий объект обучающей выборки x_i . Такая „прецедентная“ логика хорошо понятна экспертам во многих предметных областях.

Основными преимуществами данного метода являются простота реализации и универсальность в смысле независимости от специфики конкретной проблемной области. К существенным недостаткам метода можно отнести сложность выбора метрики для определения степени сходства и прямую зависимость требуемых вычислительных ресурсов от размера БП, а также низкую эффективность при работе с неполными исходными данными.

2. *Метод извлечения прецедентов на основе деревьев решений* предполагает поиск требуемых прецедентов путем сравнения ситуаций на каждом из разветвлений дерева, вершинами которого являются прецеденты из БП. Каждая вершина дерева указывает, по какой ее ветви следует искать решение. Выбор ветви осуществляется на основе информации о текущей проблемной ситуации.

Для решения данной задачи необходимо добраться до концевой вершины, соответствующей одному или нескольким прецедентам. Если концевая вершина связана с некоторым

подмножеством прецедентов, то для выбора наиболее подходящего из них может использоваться метод ближайшего соседа. Такой подход рекомендуется применять для больших БП, так как основная часть работы по извлечению прецедентов выполняется заранее на этапе построения дерева решений, что значительно сокращает время поиска решения.

3. *Метод извлечения прецедентов на основе знаний* позволяет учесть знания экспертов в конкретной предметной области (с учетом коэффициентов значимости параметров прецедентов, выявленных зависимостей и закономерностей и др.). Этот метод может успешно применяться совместно с другими методами извлечения прецедентов для обеспечения меньшего поля выборки, особенно когда БП имеет большие размеры и предметная область является открытой и динамической [9].

4. *Метод извлечения с учетом применимости прецедентов*. В большинстве систем, использующих механизмы рассуждений на основе прецедентов, предполагается, что наиболее схожие с текущей проблемной ситуацией прецеденты являются наиболее применимыми [10]. Однако это не всегда так. В основе метода извлечения с учетом применимости (адаптируемости) прецедентов лежит принцип, что извлечение прецедентов базируется не только на их сходстве с текущей проблемной ситуацией, но и на оценке результативности применения найденного прецедента к текущей ситуации. Таким образом, на выбор прецедентов влияет возможность их применения в конкретной ситуации. В некоторых системах эта проблема решается путем сохранения прецедентов вместе с комментариями по их применению. Использование такого подхода, с заранее исключенной частью заведомо неперспективных прецедентов, позволяет сделать поиск решения более эффективным.

Каждый из рассмотренных методов имеет свои преимущества. Принимая во внимание, что при управлении сложными организационно-техническими объектами на основе рассуждений, базирующихся на прецедентах, база должна иметь достаточно большое количество прецедентов, а количество их параметров может превышать тысячи, оптимальным в данном случае представляется вариант комплексирования методов для извлечения прецедентов из базы.

Для сокращения количества рассматриваемых прецедентов первоначально необходимо применить метод извлечения с учетом применимости прецедентов, исключив тем самым прецеденты, результаты решений по которым были отрицательными либо привели к негативным последствиям.

Далее можно воспользоваться методом извлечения прецедентов на основе знаний. Для этого необходимо заранее построить шкалы относительной значимости параметров для основных видов действий, характерных для конкретной предметной области. Шаблоны должны быть подготовлены с учетом знаний экспертов в данной предметной области. Применяя определенные шаблоны для текущей ситуации, следует в очередной раз сократить количество прецедентов, выбрав из оставшегося набора прецеденты, подходящие по степени значимости параметров для текущей ситуации.

На конечном этапе необходимо воспользоваться методом ближайших соседей. На практике применяются различные модификации этого метода. Обычно решение выбирается на основе нескольких ближайших точек (соседей), а не одной (метод k ближайших соседей). В алгоритме определения ближайшего соседа используется простое по координатное сопоставление текущей ситуации с прецедентом (каждый параметр для описания прецедентов рассматривается как одна из координат). Таким образом, определяется расстояние D между точкой, соответствующей текущей ситуации, и точкой, соответствующей прецеденту, а также вычисляется максимальное расстояние D_{\max} с использованием границ диапазонов параметров для описания прецедентов. Затем вычисляется значение степени сходства $SIM = 1 - D / D_{\max}$ или $SIM = (1 - D / D_{\max}) 100 \%$.

Рассмотрим алгоритм определения ближайшего соседа на следующем примере. Пусть прецеденты и текущая ситуация описываются только двумя параметрами: V — скорость объекта (диапазон изменения скорости 1000...5000 м/с); h — высота объекта над уровнем моря (диапазон изменения высоты 1000...5000 м).

В БП имеются два прецедента:

$$\Pi_1: V = 3000 \text{ м/с}, h = 3500 \text{ м}; \Pi_2: V = 4500 \text{ м/с}, h = 1500 \text{ м}.$$

Для текущей ситуации (цель) $V = 2000 \text{ м/с}, h = 3000 \text{ м}$.

Далее рассмотрим координатную плоскость (рис. 2), где по оси абсцисс отложены значения V , а по оси ординат — h . Цель имеет координаты (2000, 3000), а прецеденты — Π_1 (3000, 3500), Π_2 (4500, 1500).

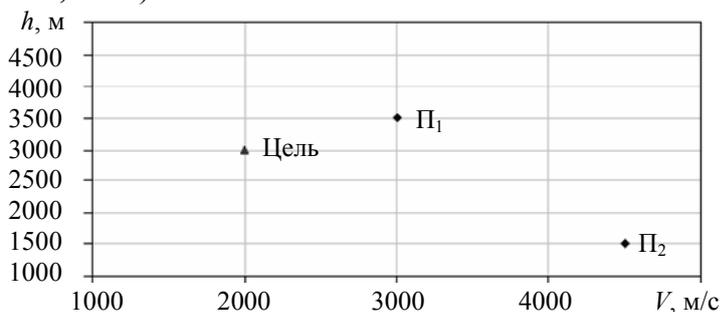


Рис. 2

Рассчитаем расстояние от цели до Π_1 (D_1) и Π_2 (D_2):

$$D_1 = \sqrt{(2000 - 3000)^2 + (3000 - 3500)^2} = 1118 \text{ м},$$

$$D_2 = \sqrt{(2000 - 4500)^2 + (3000 - 1500)^2} = 2915 \text{ м}.$$

Аналогичным образом вычисляется максимальное расстояние D_{\max} между точками с координатами (1000, 1000) и (5000, 5000), которое составит 5657 м.

Затем вычисляются значения степени сходства SIM текущей ситуации с двумя прецедентами из БП:

$$\text{— для } \Pi_1: \text{SIM}_1 = 1 - D_1 / D_{\max} = 1 - 1118 / 5657 = 0,8 \text{ (80 \%)};$$

$$\text{— для } \Pi_2: \text{SIM}_2 = 1 - D_2 / D_{\max} = 1 - 2915 / 5657 = 0,48 \text{ (48 \%)}.$$

Таким образом, прецедент Π_1 является более схожим с текущей ситуацией по сравнению с прецедентом Π_2 . При управлении сложными организационно-техническими объектами количество контролируемых и принимаемых к учету параметров иногда достигает нескольких сотен, а порой и тысяч. В этих случаях сравнение прецедентов с текущей ситуацией рассматривается не в двухкоординатной плоскости, а в n -мерном пространстве.

Заключение. Использование прецедентного метода при построении экспертных систем управления сложными организационно-техническими объектами обеспечивает, благодаря накопленным в базе прецедентов опытным данным, дополнительные возможности по повышению оперативности принятия управленческих решений и их адекватности, а также по прогнозированию. Для адекватности выдаваемых решений сложившейся ситуации необходимо наличие больших баз прецедентов в конкретной предметной области, что требует значительных временных затрат. С другой стороны, увеличение объема прецедентной базы приводит к увеличению ресурсов, требуемых для извлечения прецедентов из базы и нахождения наиболее схожей с текущей ситуацией. Комплексирование методов извлечения прецедентов позволяет существенно сократить трудоемкость и время поиска. В основу синтеза рассмотренных методов может быть положен метод ближайших соседей, а с использованием остальных методов осуществляется предварительная фильтрация прецедентов в базе.

Таким образом, применение прецедентного метода в экспертных системах позволяет:

— существенно сократить время на принятие решения;

— повысить адекватность принимаемых решений в отсутствие достаточного запаса времени на оценку обстановки;

— осуществить прогноз ситуации, а также состояния объекта после выдачи на него управляющего воздействия.

Исследования, выполненные по данной тематике, проводились при финансовой поддержке бюджетной темы № 0073–2018–0003.

СПИСОК ЛИТЕРАТУРЫ

1. Крылов А. В. Проактивное управление и его применение при управлении сложными организационно-техническими объектами // „Научная сессия ГУАП. Часть II. Технические науки“: Сб. докладов, 11—15 апр. 2016. СПб: СПбГУАП, 2016. С. 220—224.
2. Искусственный интеллект: Справочник. Кн. 2. Модели и методы / Под ред. Д. А. Поспелова. М.: Радио и связь, 1990.
3. Охтилев М. Ю., Мустафин Н. Г., Миллер В. Е., Соколов Б. В. Концепция проактивного управления сложными объектами: теоретические и технологические основы // Изв. вузов. Приборостроение. 2014. Т. 57, № 11. С. 7—15.
4. Залюбовский И. Компьютерный инструментарий инновационного менеджмента // Управление персоналом. 2000. № 1. С. 6—11.
5. Планирование на базе рассуждений по прецедентам [Электронный ресурс]: <<http://ai-center.botik.ru/planning/index.php?ptl=materials/071cbp.htm>>.
6. Aamodt A. Knowledge-intensive case-based reasoning and sustained learning // Proc. of the 9th European Conf. on Artificial Intelligence, Stockholm, 6—10 august, 1990. London: Pitman Publ., 1990. P. 1—6.
7. Зо Лин Кхаинг, Ар Кар Мьо, Варшавский П. Р., Алехин Р. В. Реализация прецедентного модуля для интеллектуальных систем // Программные системы и продукты. 2015. № 2. С. 26—31.
8. Варшавский П. Р., Алехин Р. В. Метод поиска решений в интеллектуальных системах поддержки принятия решений на основе прецедентов // Intern. Journal “Information Models and Analyses”. 2013. Vol. 2, N 4. P. 385—392.
9. Варшавский П. Р., Еремеев А. П. Моделирование рассуждений на основе прецедентов в интеллектуальных системах поддержки принятия решений // Искусственный интеллект и принятие решений. 2009. № 1. С. 45—57.
10. Варшавский П. Р., Еремеев А. П. Методы правдоподобных рассуждений на основе аналогий и прецедентов для интеллектуальных систем поддержки принятия решений // Новости искусственного интеллекта. 2006. № 3. С. 39—62.

Сведения об авторе

Алексей Валерианович Крылов — аспирант; СПИИРАН; лаборатория информационных технологий в системном анализе и моделировании; E-mail: KtAlex98@yandex.ru

Поступила в редакцию
27.08.18 г.

Ссылка для цитирования: Крылов А. В. Проблема извлечения знаний с использованием рассуждений на основе прецедентов // Изв. вузов. Приборостроение. 2018. Т. 61, № 11. С. 956—962.

**THE PROBLEM OF KNOWLEDGE RETRIEVING
WITH THE USE OF PRECEDENT-BASED REASONING****A. V. Krylov**

*St. Petersburg Institute for Informatics and Automation of the RAS,
199178, St. Petersburg, Russia
E-mail: KrAlex98@yandex.ru*

Methods of knowledge representation in the precedent knowledge base, as well as search and extraction of precedents similar to a current situation, are considered as applied to solving the problems of proactive management in decision-support systems. It is shown that for decision-support systems with large precedent bases, the most rational option is to combine the precedents extracting approaches based on the "k nearest neighbors" method. In this case, other methods are used for pre-filtering of precedents in the database.

Keywords: expert systems, decision support systems, precedent approach, methods for extracting precedents, complex organizational and technical objects

REFERENCES

1. Krylov A.V. *Nauchnaya sessiya GUAP. Chast' II. Tekhnicheskiye nauki* (Scientific session of the State University of Aviation Instrumentation. Part II. Technical Sciences), Collection of reports, 11–15 April, 2016, pp. 220–224. (in Russ.)
2. Pospelov D.A., ed., *Iskusstvennyy intellekt: Spravochnik. Kn. 2. Modeli i metody* (Artificial Intelligence: Reference book. Book 2. Models and Methods), Moscow, 1990. (in Russ.)
3. Okhtilev M.Yu., Mustafin N.G., Miller V.E., Sokolov B.V. *Journal of Instrument Engineering*, 2014, no. 11(57), pp. 7–15. (in Russ.)
4. Zalyubovskiy I. *Upravleniye personalom*, 2000, no. 1, pp. 6–11. (in Russ.)
5. <http://ai-center.botik.ru/planning/index.php?ptl=materials/071cbp.htm>. (in Russ.)
6. Aamodt A. *Proc. of the 9th European Conf. on Artificial Intelligence*, Stockholm, 6–10 August, 1990. London, Pitman Publ., 1990, pp. 1–6.
7. Zo Lin Kkhaing, Ar Kar M'o, Varshavskiy P.R., Alekhin R.V. *Programmnye Produkty i Sistemy* (Software & Systems), 2015, no. 2, pp. 26–31. (in Russ.)
8. Varshavskiy P.R., Alekhin R.V. *Intern. Journal "Information Models and Analyses"*. 2013, no. 4(2), pp. 385–392.
9. Varshavskiy P.R., Eremeyev A.P. *Scientific and Technical Information Processing*, 2009, no. 1, pp. 45–57.
10. Varshavskiy P.R., Eremeyev A.P. *Novosti iskusstvennogo intellekta*, 2006, no. 3, pp. 39–62. (in Russ.)

Data on author

Aleksey V. Krylov — Post-Graduate Student; St. Petersburg Institute for Informatics and Automation of the RAS, Laboratory of Information Technologies in Systems Analysis and Modeling; E-mail: KrAlex98@yandex.ru

For citation: Krylov A. V. The problem of knowledge retrieving with the use of precedent-based reasoning. *Journal of Instrument Engineering*. 2018. Vol. 61, N 11. P. 956–962 (in Russian).

DOI: 10.17586/0021-3454-2018-61-11-956-962