



ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

УПРАВЛЕНИЕ ДИСТАНЦИОННОГО ОБУЧЕНИЯ И ПОВЫШЕНИЯ
КВАЛИФИКАЦИИ

Кафедра «Управление качеством»

«Принятие решений на основе прецедентов»

Методические указания

к проведению практических занятий
по дисциплине

«Основы моделирования управленческих задач»

Авторы:
Димитров В.П.
Катаев В.С.

Ростов-на-Дону, 2013



Аннотация

Методические указания предназначены для студентов, обучающихся по направлению 221400 «Управление качеством». Приводится методика принятия решений на основе прецедентов. Приводятся индивидуальные задания и методика решения задач.

Авторы

Заведующий кафедрой «Управление качеством» ДГТУ д.т.н., профессор Димитров Валерий Петрович

Старший преподаватель кафедры «Управление качеством» Катаев Виктор Сергеевич





Оглавление

ВВЕДЕНИЕ.....	4
1.ПРИНЯТИЕ РЕШЕНИЙ НА ОСНОВЕ ПРЕЦЕДЕНТОВ	4
2. МОДЕЛЬНЫЙ ПРИМЕР	9
ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ.....	13
РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА.....	15



ВВЕДЕНИЕ

Целью работы является приобретение навыков использования накопленного опыта для решения подобных задач принятия решения

Задачи. Научиться извлекать прецеденты из базы прецедентов (БП), определять степень близости между текущими условиями задачи принятия решения и решениями из БП.

1. ПРИНЯТИЕ РЕШЕНИЙ НА ОСНОВЕ ПРЕЦЕДЕНТОВ

Вывод на основе прецедентов (CBR – Case Based Reasoning) является подходом, позволяющим решить новую, неизвестную задачу, используя или адаптируя решение уже известной задачи, т.е. используя уже накопленный опыт решения подобных задач.

Большинство практических задач, ориентированных на открытые и динамические предметные области, являются плохо формализованными, причем неопределенность может иметь невероятностный характер. При поиске решения таких задач необходимо применение методов правдоподобного вывода, позволяющих найти некоторое приемлемое (которое может и не быть оптимальным) решение. Один из подходов базируется на том факте, что человеку (эксперту, ЛПР) свойственно на первом этапе поиска решения новой (неизвестной) задачи пытаться использовать решения, которые принимались ранее в подобных случаях, и при необходимости адаптировать их к возникшей проблеме (текущей проблемной ситуации). Данный подход с использованием накопленного опыта лег в основу методов моделирования рассуждений на основе прецедентов.

Как правило, процесс вывода на основе прецедентов включает четыре основных этапа, образующих так называемый цикл рассуждения на основе прецедентов или CBR-цикл. Основными этапами CBR-цикла являются:

Этап 1. Извлечение наиболее адекватного (подобного) прецедента (или прецедентов) для сложившейся ситуации из библиотеки прецедентов (БП).

На первом этапе CBR-цикла – извлечении прецедентов – выполняется определение степени сходства текущей ситуации с прецедентами из БП и последующее их извлечение с целью



Принятие решений на основе прецедентов

разрешения возникшей новой проблемной ситуации. Для успешной реализации рассуждений на основе прецедентов необходимо обеспечить корректное извлечение прецедентов из БП.

Прецедент в общем случае может включать следующие компоненты:

- описание задачи (проблемной ситуации);
- решение задачи (диагностирование проблемной ситуации и рекомендации ЛПР);
- результат (или прогноз) применения решения.

Результат может включать список выполненных действий, дополнительные комментарии и ссылки на другие прецеденты. Прецедент может иметь как положительный, так и отрицательный исход применения решения, также в некоторых случаях может приводиться обоснование выбора предложенного решения и возможные альтернативы.

Основные способы представления прецедентов можно разделить на следующие группы:

- параметрические;
- объектно-ориентированные;
- специальные (в виде деревьев, графов, логических формул и т.д.).

В большинстве случаев для представления прецедентов достаточно простого параметрического представления, т.е. представления прецедента в виде набора параметров с конкретными значениями и решением (диагнозом и рекомендациями ЛПР):

$$\text{CASE}=(x_1, \dots, x_n, R),$$

где x_1, \dots, x_n – параметры ситуации, описывающей данный прецедент; $x_i \in X_1, \dots, x_n \in X_n$, n – количество параметров прецедента, X_1, \dots, X_n – области допустимых значений соответствующих параметров, R – диагноз и рекомендации ЛПР. Дополнительно может присутствовать описание результата применения найденного решения и дополнительные комментарии.

Существует целый ряд методов извлечения прецедентов и их модификаций.

Метод ближайшего соседа (NN – Nearest Neighbor) – наиболее используемый метод сравнения и извлечения прецедентов. Он позволяет довольно легко вычислить степень сходства текущей проблемной ситуации и прецедентов из БП. Для определения степени сходства на множестве параметров, используемых для описания прецедентов и текущей ситуации, вводится некоторая метрика. Далее в соответствии с выбранной метрикой определяется расстояние от целевой точки, соответствующей теку-

щей проблемной ситуации, до точек, представляющих прецеденты из БП, и выбирается ближайшая точка к целевой.

Основными преимуществами метода ближайшего соседа 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 неэффективность при работе с неполными и плохо определенными (так называемыми «зашумленными») исходными данными.

Пусть задан прецедент (C) и текущая проблемная ситуация (T) в n -мерном пространстве признаков (параметров, свойств), тогда степень сходства или близости $S(C, T)$ прецедента C и ситуации T можно определить, используя одну из следующих основных метрик, определяющих расстояние между двумя точками:

- Евклидова метрика (евклидово расстояние)

$$d_{CT} = \sqrt{\sum_{i=1}^n (x_i^C - x_i^T)^2}$$

- Квадрат евклидова расстояния:

$$d_{CT} = \sum_{i=1}^n (x_i^C - x_i^T)^2$$

- Манхэттенская метрика:

$$d_{CT} = \sum_{i=1}^n |x_i^C - x_i^T|$$

- Расстояние Чебышева:

$$d_{CT} = \max |x_i^C - x_i^T|$$

Для определения значения степени сходства $S(C, T)$ необходимо найти максимальное расстояние d_{max} в выбранной метрике, используя границы диапазонов параметров для описания прецедентов $x_i^{нач}$, $x_i^{кон}$, $i = 1, \dots, n$. Затем можно вычислить значение степени сходства $S(C, T) = 1 - d_{CT}/d_{max}$ или в процентах $S(C, T) = (1 - d_{CT}/d_{max}) * 100\%$.

Если степень сходства $S(C, T)$ превышает пороговое значение степени сходства H , то прецедент можно использовать для решения текущей проблемной ситуации.

Метод извлечения прецедентов на основе деревьев решений базируется на нахождении требуемых прецедентов путем разрешения вершин дерева решений. Каждая вершина дерева указывает, по какой ее ветви следует осуществлять дальней-



ший поиск решения. Выбор ветви производится на основе информации о текущей проблемной ситуации. Необходимо добраться до концевой вершины, которая соответствует одному или нескольким прецедентам.

Если концевая вершина связана с некоторым подмножеством прецедентов, то тогда для выбора наиболее подходящего из них может использоваться метод ближайшего соседа. Такой подход рекомендуется применять для больших БП, т.к. основная часть работы по извлечению прецедентов выполняется заранее на этапе построения дерева решений, что значительно сокращает время поиска решения.

Метод извлечения прецедентов на основе знаний в отличие от предыдущих методов позволяет учесть знания экспертов (ЛПР) по конкретной предметной области (коэффициенты важности параметров, выявленные зависимости и т.д.) при извлечении прецедентов. Метод реализует подход, основанный на индексации прецедентов специальным образом (семантической индексации). При определении прецедентов учитываются важности параметров прецедентов, заданные экспертом или ЛПР, и другая информация, позволяющая учесть знания о конкретной предметной области. За счет этого значительно сокращается время поиска решения, что является существенным достоинством данного метода. Метод может успешно применяться совместно с другими методами извлечения прецедентов, особенно когда БП имеет большие размеры и предметная область является открытой и динамической.

Метод извлечения прецедентов с учетом их применимости. В большинстве систем, использующих механизмы рассуждений на основе прецедентов, предполагается, что наиболее схожие с текущей проблемной ситуацией прецеденты являются и наиболее применимыми в этой ситуации. Однако это не всегда так. В основе методов извлечения на основе применимости прецедентов лежит тот факт, что извлечение прецедентов базируется не только на их сходстве с текущей проблемной ситуацией, но и на том, насколько хорошую для желаемого результата модель они собой представляют. Таким образом, на выбор извлекаемых прецедентов влияет возможность их успешного применения (адаптации) в конкретной ситуации, т.е. наличие сведений о их применимости в сложившейся ситуации. В некоторых системах эта проблема решается путем сохранения прецедентов вместе с комментариями по их применению. Использование данного метода позволяет сделать поиск решения более эффективным, зара-

Принятие решений на основе прецедентов

нее отбрасывая часть заведомо неперспективных прецедентов.

Этап 2. Повторное использование извлеченного прецедента для попытки решения текущей проблемы.

Этап 3. Пересмотр и адаптация в случае необходимости полученного решения в соответствии с текущей проблемой;

Этап 4. Сохранение (запоминание) вновь принятого решения как части нового прецедента.

Для реализации CRB-цикла методом ближайшего соседа предлагается следующий алгоритм:

Входные данные: текущая ситуация **T** (т.е. должны быть заданы числовые значения параметров, описывающие ситуацию), **CL** – непустое множество прецедентов, хранящееся в БП, w_1, \dots, w_n – веса (коэффициенты важности) параметров, m – количество рассматриваемых прецедентов из БП и пороговое значение степени сходства H .

Выходные данные: Множество прецедентов SC (Set of Cases), которые имеют степень сходства (близости) больше или равную порогового значения H .

Промежуточные данные: Вспомогательные переменные i, j (параметры цикла).

Шаг 1. $SC = \emptyset, j = 1$; переход к следующему шагу.

Шаг 2. Если $j \leq m$, то выбрать прецедент C_j из БП ($C_j \in CL$) и переход к шагу 3; иначе считать, что все прецеденты из БП рассмотрены и переход к шагу 6.

Шаг 3. Вычислить расстояние d_{CT} в евклидовой метрике между выбранным прецедентом C_j и текущей ситуацией T с учетом коэффициентов важности параметров:

$$d_{CT} = \sqrt{\sum_{i=1}^n w_i (x_i^{C_j} - x_i^T)^2}$$

Если значение параметра $x_i^{C_j}$ в описании прецедента C_j отсутствует, то вычислить расстояния d_{CT} , учитывая, что $x_i^{C_j} = x_i^T$, а если отсутствует значение параметра x_i^T в описании текущей ситуации T , то вычислить расстояние d_{CT} , полагая:

$$x_i^T = \begin{cases} x_i^{\text{нач}}, & \text{если } (x_i^C - x_i^{\text{нач}}) > (x_i^{\text{кон}} - x_i^C) \\ x_i^{\text{кон}}, & \text{если } (x_i^C - x_i^{\text{нач}}) \leq (x_i^{\text{кон}} - x_i^C) \end{cases}$$

Переход к шагу 4.

Шаг 4. Вычислить степень сходства $S(C_j, T) = 1 - d_{CT} / d_{\max}$ (или $S(C_j, T) = 1 - d_{CT} / d_{\max} * 100\%$, если пороговое значение H задано в процентах), учитывая при вычислении d_{\max} веса параметров;



Принятие решений на основе прецедентов

переход к шагу 5.

Шаг 5. Если $S(C_j, T) \geq H$, то извлечь прецедент C_j из БП и добавить в результирующее множество SC ; присвоить $j=j+1$ и переход к шагу 2.

Шаг 6. Если $SC \neq \emptyset$, то прецеденты для текущей ситуации успешно извлечены, выдать их список пользователю (ЛПР) и переход шагу 7. Иначе, если $SC = \emptyset$ и прецеденты для текущей проблемной ситуации не найдены, выдать сообщение пользователю (ЛПР) о необходимости уменьшения порогового значения H и переход к шагу 7.

Шаг 7. Конец алгоритма.

Отметим, что найденные прецеденты могут быть упорядочены по убыванию значений их степеней сходства с текущей ситуацией и выданы пользователю (ЛПР), который может с учетом собственных предпочтений выбрать наиболее подходящие прецеденты и на их основе получить решение (диагноз и рекомендации) для текущей проблемной ситуации. Заметим также, что для вычисления расстояния на шаге 3 могут быть использованы и другие из рассмотренных метрик, выбор которых делается с учетом специфики проблемной области.

2. МОДЕЛЬНЫЙ ПРИМЕР

Рассмотрим алгоритм извлечения прецедентов методом ближайшего соседа с использованием евклидовой метрики.

Постановка задачи.

Предметная область: Расчёт тормозного усилия для автомобиля по заданной длине тормозного пути.

На длину тормозного пути оказывают влияние следующие факторы: Масса автомобиля; коэффициент сцепления шин с дорогой, тормозное усилие.

Коэффициент сцепления шин с дорогой зависит от влажности и покрытия дороги (таблица 1).

Таблица 1 – Коэффициенты сцепления шин с дорогой

Покрывтие дороги	Коэффициент сцепления	
	Сухая поверхность	Мокрая поверхность
Асфальт	0,7 – 0,8	0,3 – 0,4
Бетон	0,6 – 0,7	0,3 – 0,4
Булыжная	0,5 – 0,6	0,3 – 0,35
Грунтовая	0,5 – 0,6	0,3 – 0,4
Глина	0,5 – 0,6	0,2 – 0,4
Укатанный (плотный) снег	0,2 – 0,3	
Лёд	0,08 – 0,1	0,08 и менее

Необходимо определить необходимое тормозное усилие для автомобиля заданной массы при заданном состоянии дороги, обеспечивающее установленный тормозной путь. Использовать метод принятия решений на основе прецедентов.

Пусть есть некоторая база прецедентов (таблица 2), заданных параметрически:

$$CASE=(x_1, x_2, x_3, R)$$

где x_1 – масса автомобиля, кг; x_2 – коэффициент сцепления шин с дорогой; x_3 – тормозной путь, м; R – тормозное усилие, кН.

Таблица 2 – База прецедентов

№ прецедента	x_1	x_2	x_3	R
1	2100	0,7	20	1000
2	2100	0,4	60	900
3	1900	0,6	40	950
4	2000	0,5	45	900
5	2100	0,8	30	800
6	1950	0,65	50	850
7	2000	0,6	20	1100
8	2050	0,75	40	700
9	1900	0,55	40	930
10	2100	0,7	30	850

Границы диапазона варьирования параметров представлены в таблице 3.

Принятие решений на основе прецедентов

Таблица 3 – Границы диапазона варьирования параметров

Параметр	$X_{лев}$	$X_{прав}$	$w(X_{прав} - X_{лев})$
x_1	1700	2300	180
x_2	0,2	0,8	0,24
x_3	15	80	19,5

Текущая ситуация **T** представляет следующий набор параметров: $T=(1970; 0,66; 35)$, параметры имеют следующие веса: $w_{x1} = 0,3; w_{x2} = 0,4; w_{x3} = 0,3$, количество рассматриваемых прецедентов из БП = 10; пороговое значение степени сходства $H = 0,8$. Промежуточные данные: Вспомогательные переменные i, j

Используем описанный ранее алгоритм.

Шаг 1. $SC = \emptyset, j=1$; переходим к следующему шагу

Шаг 2. $j=1 < m$, выбираем прецедент C_1 из БП, переходим к шагу 3.

Шаг 3. Вычислить расстояние d_{C1T} в евклидовой метрике между выбранным прецедентом C_1 и текущей ситуацией T с учетом коэффициентов важности параметров:

$$d_{ст1} = \sqrt{0,3(2100 - 1970)^2 + 0,4(0,7 - 0,66)^2 + 0,3(35 - 20)^2} = 71,68$$

Переход к шагу 4.

Шаг 4. Вычислить степень сходства $S(C_1, T) = 1 - d_{C1T}/d_{max}$, учитывая при вычислении d_{max} веса параметров: $d_{max} = 180$, тогда:

$$S(C_1, T) = 1 - 71,68/180 = 0,602$$

переход к шагу 5.

Шаг 5. $S(C_1, T) < H$, не извлекаем прецедент; присвоим $j=j+1$ и переход к шагу 2.

Шаг 2. $j=2 < m$, выбираем прецедент C_2 из БП, переходим к шагу 3.

Шаг 3. Вычислить расстояние $d_{C2T} = 72,51$. Переход к шагу 4.

Шаг 4. $S(C_2, T) = 0,6$. Переход к шагу 5.

Шаг 5. $S(C_2, T) < H$, не извлекаем прецедент; присвоим $j=j+1$ и переход к шагу 2.

Шаг 2. $j=3 < m$, выбираем прецедент C_3 из БП, переходим к шагу 3.

Шаг 3. Вычислить расстояние $d_{C3T} = 38,44$. Переход к шагу 4.

Шаг 4. $S(C_3, T) = 0,79$. Переход к шагу 5.

Шаг 5. $S(C_3, T) < H$, не извлекаем прецедент; присвоим $j=j+1$ и переход к шагу 2.

Шаг 2. $j=4 < m$, выбираем прецедент C_4 из БП, переходим к шагу 3.

Принятие решений на основе прецедентов

- Шаг 3. Вычислить расстояние $d_{C_4T} = 17,32$. Переход к шагу 4.
- Шаг 4. $S(C_4, T) = 0,904$. Переход к шагу 5.
- Шаг 5. $S(C_4, T) > H$, извлекаем прецедент и добавляем в результирующее множество SC ; присвоим $j=j+1$ и переход к шагу 2.
- Шаг 2. $j=5 < m$, выбираем прецедент C_5 из БП, переходим к шагу 3.
- Шаг 3. Вычислить расстояние $d_{C_5T} = 71,26$. Переход к шагу 4.
- Шаг 4. $S(C_5, T) = 0,604$. Переход к шагу 5.
- Шаг 5. $S(C_5, T) < H$, не извлекаем прецедент; присвоим $j=j+1$ и переход к шагу 2.
- Шаг 2. $j=6 < m$, выбираем прецедент C_6 из БП, переходим к шагу 3.
- Шаг 3. Вычислить расстояние $d_{C_6T} = 13,7$. Переход к шагу 4.
- Шаг 4. $S(C_6, T) = 0,924$. Переход к шагу 5.
- Шаг 5. $S(C_6, T) > H$, извлекаем прецедент и добавляем в результирующее множество SC ; присвоим $j=j+1$ и переход к шагу 2.
- Шаг 2. $j=7 < m$, выбираем прецедент C_7 из БП, переходим к шагу 3.
- Шаг 3. Вычислить расстояние $d_{C_7T} = 18,37$. Переход к шагу 4.
- Шаг 4. $S(C_7, T) = 0,9$. Переход к шагу 5.
- Шаг 5. $S(C_7, T) > H$, извлекаем прецедент и добавляем в результирующее множество SC ; присвоим $j=j+1$ и переход к шагу 2.
- Шаг 2. $j=8 < m$, выбираем прецедент C_8 из БП, переходим к шагу 3.
- Шаг 3. Вычислить расстояние $d_{C_8T} = 43,9$. Переход к шагу 4.
- Шаг 4. $S(C_8, T) = 0,756$. Переход к шагу 5.
- Шаг 5. $S(C_8, T) < H$, не извлекаем прецедент; присвоим $j=j+1$ и переход к шагу 2.
- Шаг 2. $j=9 < m$, выбираем прецедент C_9 из БП, переходим к шагу 3.
- Шаг 3. Вычислить расстояние $d_{C_9T} = 38,44$. Переход к шагу 4.
- Шаг 4. $S(C_9, T) = 0,79$. Переход к шагу 5.
- Шаг 5. $S(C_9, T) < H$, не извлекаем прецедент; присвоим $j=j+1$ и переход к шагу 2.
- Шаг 2. $j=10 = m$, выбираем прецедент C_{10} из БП, переходим к шагу 3.
- Шаг 3. Вычислить расстояние $d_{C_{10}T} = 71,26$. Переход к шагу 4.

Принятие решений на основе прецедентов

Шаг 4. $S(C_{10}, T) = 0,604$. Переход к шагу 5.

Шаг 5. $S(C_{10}, T) < H$, не извлекаем прецедент; присвоим $j = j + 1$ и переход к шагу 2.

Шаг 2. $j = 11 > m$, все прецеденты из БП рассмотрены, переход к шагу 6

Шаг 6. $SC \neq \emptyset$ - прецеденты для текущей ситуации успешно извлечены, выдать их список пользователю (ЛПР) (таблица 4)

Таблица 4 – Прецеденты с удовлетворяющей степенью близости

№ прецедента	x_1	x_2	x_3	R	$S(C, T)$
4	2000	0,5	45	900	0,904
6	1950	0,65	50	850	0,924
7	2000	0,6	20	1100	0,9

переход шагу 7.

Шаг 7. Конец алгоритма.

Наибольшую степень близости имеет прецедент 6, следовательно разумно выбрать его в качестве решения текущей задачи и обеспечить тормозное усилие около 850 кН.

ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

В соответствии с вариантом индивидуального задания определить необходимое тормозное усилие. № варианта соответствует последним двум цифрам зачётной книжки (таблицы 6,7).

База прецедентов общая для всех студентов.

Таблица 5 – База прецедентов

№ прецедента	x_1	x_2	x_3	R
1	2500	0,7	20	1200
2	2100	0,4	60	900
3	1400	0,6	40	650
4	2200	0,8	40	870
5	2100	0,8	30	800
6	1950	0,65	50	850
7	2150	0,6	30	1000
8	2050	0,75	40	700
9	1700	0,65	40	730
10	2100	0,7	30	850

Принятие решений на основе прецедентов

Таблица 6 – Индивидуальные задания (Параметры)

Параметры текущей ситуации			
Последняя цифра зачётной книжки	x_1	x_2	x_3
0	1500	0,7	40
1	2300	0,45	50
2	1800	0,2	80
3	1550	0,47	42
4	2150	0,68	35
5	1650	0,85	30
6	2250	0,8	40
7	2150	0,6	55
8	1650	0,8	25
9	2050	0,6	35

Таблица 7 – Индивидуальные задания (Пороговое значение)

Последняя цифра зачётной книжки										
	0	1	2	3	4	5	6	7	8	9
π	0,95	0,8	0,85	0,9	0,7	0,85	0,75	0,92	0,8	0,75



РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Моделирование рассуждений на основе прецедентов в интеллектуальных системах поддержки принятия решений / Варшавский П.Р., Еремеев А.П. // М.: Искусственный интеллект и принятие решений – 2009, №1 – С.45-57.