

Принятие решений  
Системы поддержки принятия решений

10 марта 2018 г.

# Оглавление

<b>1</b>	<b>Задачи</b>	<b>3</b>
1.1	Принятие решений с точки зрения системного анализа . . . . .	3
1.2	Некоторые формальные постановки задач . . . . .	4
1.2.1	Детерминированные задачи . . . . .	4
1.2.2	Недетерминированные задачи . . . . .	5
1.2.3	Информационные ситуации . . . . .	6
1.3	Пример: многоуровневая модель производства . . . . .	8
1.3.1	Уровень 0: планирование производства . . . . .	8
1.3.2	Уровень 1: производство . . . . .	8
1.3.3	Уровень 2: поставки, производство, сбыт . . . . .	10
1.3.4	Уровень 3: отраслевые игры . . . . .	11
<b>2</b>	<b>Некоторые методы принятия решений</b>	<b>13</b>
2.1	Метод анализа иерархий . . . . .	13
2.2	Теория ограничений Голдратта . . . . .	13
2.3	Формирование решений по управлению . . . . .	13
<b>3</b>	<b>Системы поддержки принятия решений</b>	<b>14</b>
3.1	Классификация СППР . . . . .	14
3.2	Структура СППР . . . . .	15
<b>4</b>	<b>Данные и знания</b>	<b>16</b>
4.1	Данные . . . . .	16
4.1.1	Модели данных . . . . .	16
4.1.2	Модель данных «сущность–связь» . . . . .	17
4.1.3	Реляционные базы данных . . . . .	19
4.1.4	Хранилища данных . . . . .	21
4.1.5	OLAP . . . . .	22
4.2	Знания . . . . .	23
4.2.1	Представление знаний . . . . .	23
4.2.2	Знания, данные, умения . . . . .	24
<b>5</b>	<b>Data Mining</b>	<b>29</b>
5.1	Качественные и количественные величины . . . . .	29
5.2	Регрессия . . . . .	29
5.3	Классификация . . . . .	29
5.4	Кластерный анализ . . . . .	30
5.5	Ассоциативные правила . . . . .	30
<b>A</b>	<b>Тестовые тесты</b>	<b>33</b>
A.1	Системы поддержки принятия решений . . . . .	33
A.1.1	№ 0 . . . . .	33
A.1.2	№ 1 . . . . .	34

А.2	Принятие решений . . . . .	34
А.2.1	№ 0 . . . . .	34
А.2.2	№ 1 . . . . .	35
А.2.3	№ 2 . . . . .	35

# Глава 1

## Задачи

Система поддержки принятия решений предназначена для поддержки многокритериальных (вообще говоря) решений в сложной информационной среде.

1. Какие задачи (типы задач, классы задач) будут рассматриваться;
2. Какие методы (способы, приёмы) принятия решений будут использоваться;
3. Какое программное обеспечение может быть использовано для облегчения работы лица, принимающего решения.

### 1.1 Принятие решений с точки зрения системного анализа

С точки зрения общей теории систем [6, стр. 14–19], система  $S$  может быть описана упорядоченным набором

$$S = \langle A, R, Z, E, \Delta T, N, L_N \rangle,$$

где  $A$  — составные части системы (элементы, компоненты, подсистемы);  $R$  — связи между составными частями;  $Z$  — цель существования системы (у искусственно созданных систем цель есть всегда);  $E$  — окружающая среда;  $\Delta T$  — интервал времени существования системы, на котором она рассматривается;  $N$  — наблюдатель, с точки зрения которого выполняется анализ;  $L_N$  — язык, используемый наблюдателем для описания системы.

Если интервал времени  $\Delta T$  достаточно большой, то следует учесть возможность изменения значения параметров системы, а также, возможно, множеств составных частей  $A$  и связей  $R$ . При малых значениях  $\Delta T$  множества  $A$  и  $R$  и, возможно, значения её параметров можно считать неизменными.

Выделив из набора  $A$  и  $R$  те составные части и связи, которые не изменяются на интервале  $\Delta T$ , получим описание структуры системы  $S$ .

Заметим, что цель  $Z$  системы единственна. «Цель можно определить как результат, на достижение которого направлены усилия. Необходимое условие — то, без чего достижение результата невозможно. Между этими понятиями существует логическая связь: следует соблюдать определённые условия, чтобы достичь цели» [8, стр. 33–35]. Таким образом для экономической, коммерческой системы  $S$  удовлетворение потребностей потребителей, получение конкурентных преимуществ, выход на новые рынки и т.п. можно рассматривать как необходимые условия достижения единственной цели — получения максимальной прибыли на заданном интервале времени.

Рассматривая взаимодействие системы  $S$  со средой  $E$ , можно выделить набор переменных  $x$  на входе системы и переменных  $y$  на выходе из неё. Значения  $x$  определяются средой и не зависят напрямую от поведения системы (множество независимых, экзогенных переменных).

Значения  $y$  определяются значениями  $x$  и функционированием системы (множество зависимых, эндогенных переменных). Формально связь  $x$  и  $y$  выражается следующим уравнением, описывающим функционирование системы:

$$y = f(x | \theta) + \varepsilon,$$

где  $\theta$  — параметры системы, а функция  $f$  описывает её поведение. Случайное слагаемое  $\varepsilon$  вводится для описания влияния неучтённых факторов.

От расположения наблюдателя  $N$  зависит набор способов воздействия на систему, управления ей.

- Наблюдатель, находящийся внутри системы, может непосредственно управлять её работой (например, изменяя структуру или параметры функции  $f$ ). Так директор предприятия влияет на глобальное поведение системы, начальник цеха — на поведение подсистемы и т.п.
- Наблюдатель, находящийся вне системы  $S$ , но внутри среды  $E$ , может влиять на поведение системы только косвенно. В этом случае множество переменных  $x$  разбивается на множество управляемых переменных  $x_1$ , значения которых определяются наблюдателем, и множеством неуправляемых переменных  $x_2$ . Наблюдатель выступает в роли управляющего элемента, который на основе заданного алгоритма устанавливает значения  $x_1$  в зависимости от значений  $y$ :

$$y_t = f(x_1(y_{t-1}), x_2 | \phi).$$

Например, министерство может задавать «правила игры» для предприятий данной отрасли с помощью приказов, предписаний, постановлений и т.п. Правительство (федеральное, региональное) влияет на поведение системы экономическими и юридическими методами.

- Наблюдатель, находящийся вне системы и вне среды никаким образом не может влиять на поведение системы.

## 1.2 Некоторые формальные постановки задач

### 1.2.1 Детерминированные задачи

#### Однокритериальные задачи с одним агентом

$$\begin{aligned} f(x, y) \rightarrow_x \text{extr} \\ g(x, y) \in G \end{aligned}$$

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, g: \mathbb{R}^n \rightarrow \mathbb{R}^m$$

#### Многокритериальные задачи с одним агентом

[12, 32]

$$\begin{aligned} f(x, y) \rightarrow_x \text{extr} \\ g(x, y) \in G \end{aligned}$$

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^{m_1}, g: \mathbb{R}^n \rightarrow \mathbb{R}^{m_2}$$

## Игровые, многоуровневые и смешанные задачи

[34]

Игра — агенты одновременно принимают решения, многоуровневая оптимизация — агенты по очереди принимают решения. В смешанных задачах решения на одном уровне принимаются одновременно, между уровнями — по очереди.

$$f_j(x_1, \dots, x_k, y) \rightarrow_{x_j} \text{extr}, \quad j = 1, \dots, k$$
$$g(x, y) \in G$$

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^{m_1}, \quad g: \mathbb{R}^n \rightarrow \mathbb{R}^{m_2}$$

## Задачи динамического программирования

[3]

$$f(x, y_0) = \sum_{j=1}^N f_j(x_j, y_{j-1}) \rightarrow \text{extr}, \quad y_i = \phi_i(y_{i-1}, x_{i-1})$$

## Поиск в пространстве состояний

[25]

- Найти решение.
- Найти оптимальное (субоптимальное) решение.

## 1.2.2 Недетерминированные задачи

### Интервальные величины

[1]

$$x \in [\underline{x}, \bar{x}]$$

### Стохастические величины

[18]  $\xi$

$$E \xi = \int_{-\infty}^{\infty} x \cdot p_{\xi}(x) dx \quad \text{или} \quad E \xi = \sum_{i=1}^n x_i \cdot P\{\xi = x_i\}$$

$$V \xi = E (\xi - E \xi)^2$$

### Нечёткие величины

[18]  $X \subset U$

$$\text{supp}_X = \{u \in U \mid \mu_X(u) > 0\}$$

$$E X = \frac{\int_{\text{supp}_X} x \cdot \mu_X(x) dx}{\int_{\text{supp}_X} \mu_X(x) dx} \quad \text{или} \quad E X = \frac{\sum_{i=1}^n x_i \cdot \mu_X(x_i)}{\sum_{i=1}^n \mu_X(x_i)}$$

$$X_{\alpha} = \{u \in U \mid \mu_X(u) \geq \alpha\}$$

$$X = \bigcup_{\alpha \in (0,1]} \alpha \cdot X_{\alpha}$$

### 1.2.3 Информационные ситуации

[30]

Состояния среды  $\Theta = \{\theta_1, \dots, \theta_m\}$ , альтернативы  $\Phi = \{\phi_1, \dots, \phi_n\}$ , оценочный функционал  $K: \Theta \times \Phi \rightarrow \mathbb{R}$ ,  $K = \{f_{jk}\}$

$$\max_{\phi_k \in \Phi} \{f_{jk}\}: F = F^+ = \{f_{jk}^+\}, \min_{\phi_k \in \Phi} \{f_{jk}\}: F = F^- = \{f_{jk}^-\}$$

Функция сожаления  $r$ :

$$l_j = \max_{\phi_k \in \Phi} f_{jk}^+, \quad r_j(\phi_k) = l_j - f_{jk}^+$$

$$L_j = \min_{\phi_k \in \Phi} f_{jk}^-, \quad r_j(\phi_k) = f_{jk}^- - L_j$$

#### Первая информационная ситуация

Известны значения вероятностей  $p_j = P\{\theta = \theta_j\}$ ,  $j = 1, \dots, m$

#### Критерий Байеса

$$B^+(p, \phi_{k_0}) = \max_{\phi_k \in \Phi} B^+(p, \phi_k) = \max_{\phi_k \in \Phi} \sum_{j=1}^m p_j f_{jk}^+$$

#### Критерий максимизации вероятности распределения оценочного функционала

$$\alpha_1 = \min_{\phi_k \in \Phi} \min_{\theta_k \in \Theta} f_{jk}^+, \quad \alpha_2 = \max_{\phi_k \in \Phi} \max_{\theta_k \in \Theta} f_{jk}^+$$

$$\alpha_1 < \alpha < \alpha_2$$

$$P\{f_{jk_0} \geq \alpha\} = \max_{\phi_k \in \Phi} P\{f_{jk} \geq \alpha\}$$

#### Критерий минимума дисперсии оценочного функционала

$$\sigma_k^2 = \sum_{j=1}^m p_j \cdot (f_{jk} - B^+(p, \phi_k))^2$$

$$\sigma_{k_0}^2 = \min_{\phi_k \in \Phi} \sigma_k^2$$

#### Модальный критерий

$$p_{j_1} = \max_{\theta_j \in \Theta} P\{\theta = \theta_j\}$$

$$f_{j_1 k_0}^+ = \max_{\phi_k \in \Phi} f_{j_1 k}^+$$

#### Модифицированный критерий

$$0 \leq \lambda \leq 1$$

$$\varkappa(p, \phi_k) = (1 - \lambda)B^+(p, \phi_k)^2 - \lambda\sigma_k^2$$

$$\varkappa(p, \phi_{k_0}) = \max_{\phi_k \in \Phi} \varkappa(p, \phi_k)$$

## Вторая информационная ситуация

Известны параметризованные значения вероятностей  $p_j(\omega) = P\{\theta = \theta_j\}$ ,  $j = q, \dots, m$ .  $\omega \in \Omega$

$$\int_{\Omega} \rho(\omega) d\omega = 1 \quad \text{или} \quad \sum_{\theta_i \in \Theta} \rho_i = 1, \quad \rho_i = P\{\omega = \omega_i\}$$

$$B_{\Omega}^+(\bar{p}, \phi_k) = E B^+(p(\omega), \phi_k)$$

$$E B^+(p(\omega), \phi_k) = \sum_{j=1}^m f_{jk} \int_{\Omega} p_j(\omega) \rho(\omega) d\omega \quad \text{или} \quad E B^+(p_j(\omega), \phi_k) = \sum_{j=1}^m f_{jk} \sum_{\theta_i \in \Theta} p_j(\omega_i) \rho_i$$

## Третья информационная ситуация

Известно, что вероятности  $p_j = P\{\theta = \theta_j\}$ ,  $j = q, \dots, m$  упорядочены: например,

$$p_1 \geq p_2 \geq \dots \geq p_m, \quad \sum_{j=1}^m p_j = 1$$

## Точечные оценки Фишборна

$$\tilde{p}_j = \frac{2(m-j+1)}{m(m+1)} \quad \text{при} \quad p_1 \geq p_2 \geq \dots \geq p_m$$

$$\tilde{p}_j = \frac{2^{m-j}}{2^m - 1} \quad \text{при} \quad p_j \geq p_{j+1} + \dots + p_m$$

## Четвёртая информационная ситуация

Известно, что  $p_j = P\{\theta = \theta_j\}$ ,  $j = 1, \dots, m$ . Нет активного противодействия.

## Критерий Бернулли-Лапласа

$$\tilde{p}_j = \frac{1}{m}$$

## Критерий Хоменюка

$$\tilde{p}_j = \frac{\sum_{i=1}^n (\max_{\phi_k \in \Phi} f_{jk}^+ - f_{ji}^+)}{\sum_{l=1}^m \sum_{i=1}^n (\max_{\phi_k \in \Phi} f_{lk}^+ - f_{li}^+)}, \quad \tilde{p}_j = \frac{\sum_{i=1}^n f_{ji}^-}{\sum_{l=1}^m \sum_{i=1}^n f_{li}^-}$$

## Пятая информационная ситуация

Имеется активное противодействие.

## Критерий Вальда

$$\tilde{f}_{k_0} = \max_{\phi_k \in \Phi} \min_{\theta_j \in \Theta} f_{jk}^+$$

## Критерий Севиджа

$$\tilde{f}_{k_0} = \min_{\phi_k \in \Phi} \max_{\theta_j \in \Theta} f_{jk}^-$$

## Шестая информационная ситуация

Промежуточная между первой и пятой информационными ситуациями.



## Критерий Гурвица

$$0 \leq \lambda \leq 1$$

$$\lambda \min_{\theta_j \in \Theta} f_{jk_0}^+ + (1 - \lambda) \max_{\theta_j \in \Theta} f_{jk_0}^+ = \max_{\phi_j \in \Phi} \left( \lambda \min_{\theta_j \in \Theta} f_{jk}^+ + (1 - \lambda) \max_{\theta_j \in \Theta} f_{jk}^+ \right)$$

## 1.3 Пример: многоуровневая модель производства

### 1.3.1 Уровень 0: планирование производства

В простейшей модели планирования производства [14, стр. 17–19] предполагается ненасыщенность рынка, то есть можно продать всё, что произведено. Так как целью  $Z$  (коммерческой) производственной системы является получение прибыли, то в качестве критерия естественно использовать максимизацию прибыли предприятия.

На уровне 0 систему можно рассматривать в виде чёрного ящика: структура системы, окружающая её среда, интервал времени — всё это не имеет значения. План производства  $x$  определяется только имеющимися ресурсами  $b$  и значениями дохода от единицы продукции  $c$ :

$$\begin{aligned} \sum_{i=1}^n c_i x_i &\rightarrow \max, \\ \sum_{i=1}^n a_{ij} x_i &\leq b_j, \quad j = 1, \dots, m, \\ x_i &\geq 0, \quad i = 1, \dots, n. \end{aligned}$$

Здесь  $x_i$  — объём выпуска  $i$ -го вида продукции ( $i = 1, \dots, n$ ),  $c_i$  — прибыль от продажи единицы продукции  $i$ -го вида (рыночная цена минус себестоимость),  $b_j$  — запасы  $j$ -го вида сырья ( $j = 1, \dots, m$ ),  $a_{ij}$  — производственные коэффициенты (расход сырья  $j$ -го вида при производстве единицы продукции  $i$ -го вида).

В случае насыщенного рынка план производства определяется заключёнными контрактами.

### 1.3.2 Уровень 1: производство

План производства, полученный на предыдущем уровне, требуется выполнить. Очевидно, что, вообще говоря, невозможно произвести всю запланированную продукцию  $x$  моментально, поэтому необходимо определить последовательность действий для достижения цели — выполнения плана в срок.

На данном уровне при построении модели уже требуется учитывать как структуру системы  $S$ , так и горизонт планирования  $\Delta T$ .

#### Технологическая операция

Пусть имеется  $N_o$  технологических операций, причём  $i$ -я,  $i = 1, \dots, N_o$ , операция описывает преобразование

$$m_i \xrightarrow{o_i} p_i$$

исходного набора сырья  $m_i = \langle m_{i_1}, \dots, m_{i_{n_i}} \rangle$  в набор конечной продукции  $p_i = \langle p_{i_1}, \dots, p_{i_{p_i}} \rangle$ .

Формально, технологическая операция  $o_i$  описывается кортежем  $\langle i, t_i, m_i, p_i \rangle$ , где  $t_i$  — продолжительность технологической операции;  $m_i = \{m_{i,1}, \dots, m_{i,n_{m_i}}\}$  — исходный набор сырья ( $m_{i,j} = \langle k_{i,j}, v_{i,j}, t_{i,j} \rangle$ , где  $k_{i,j}$  — тип сырья;  $v_{i,j}$  — расход сырья на производство единицы продукции (технологический коэффициент);  $t_{i,j}$  — момент подачи сырья в переработку относительно начала операции);  $p_i = \{p_{i,1}, \dots, p_{i,n'_{m_i}}\}$  — набор конечной продукции

$(p_{i,j} = \langle k_{i,j}, v_{i,j}, t_{i,j} \rangle)$ , где  $k_{i,j}$  — тип продукции;  $v_{i,j}$  — удельный объём производства (в расчёте, что  $\max_{i,j} v_{i,j} = 1$ );  $t_{i,j}$  — время получения продукции относительно начала операции).

При описании технологической операции требуется указать

- все исходные компоненты подаются одновременно или требуется указывать время подачи (относительно начала операции);
- есть ли обратные связи ( $\exists k, l: m_{i_k} = p_{i_l}$ )
- длительность операции — число (строго определённая; продолжительность) или неопределённая величина (например, случайная величина с определённой дисперсией);
- требуется ли учитывать дополнительные факторы (например температурный режим), от которых зависит длительность (качество, что-то ещё) операции.

## Технологический процесс

Технологический процесс — упорядоченный набор операций, предназначенный для получения из исходного сырья готовой продукции.

Пусть имеется  $N_T$  технологических процессов, где  $i$ -й процесс — это набор технологических операций:  $\langle o_{i_1}, \dots, o_{i_{n_i}} \rangle$ .

При описании технологического процесса требуется указать

- имеются ли обратные связи внутри процесса и связи между процессами (выходной продукт какой-либо операции данного процесса является исходным сырьём предыдущей операции этого же процесса или какой-то операции другого процесса);
- следует ли учитывать возможность брака;
- имеются ли одинаковые подмножества/поднаборы технологических операций в разных технологических процессах;
- могут ли полуфабрикаты храниться на складе или же они должны сразу поступать на вход следующей операции.

## Процессор

Процессор — устройство, выполняющее заданную операцию. В химической промышленности процессором может быть, например, химический реактор.

Пусть имеется  $N_p$  процессоров, причём  $i$ -й процессор описывается кортежем  $\langle i, c, o \rangle$ , где  $i$  — номер процессора;  $c$  — описание производительности процессора;  $o$  — множество технологических операций, доступных для процессора.

При описании процессора требуется указать

- зависит ли длительность операции от процессора;
- имеется ли технологическая пауза между двумя следующими друг за другом технологическими операциями одного типа;
- имеется ли технологическая пауза между двумя следующими друг за другом технологическими операциями разного типа;
- возможно ли неполное использование объёма процессора (частичная загрузка);
- какие существуют связи между процессорами, а также между процессорами и контейнерами:
  - каждый с каждым;
  - ненаправленный граф;
  - ориентированный граф.

## Планирование производства

При планировании производства в модели уровня 1 требуется распределить работы по  $N_p$  процессорам таким образом, чтобы

- выполнить план производства  $x$ ;
- выполнить условия заключённых контрактов (если они есть);
- затратить на работу наименьшее время.

При решении данной задачи можно использовать методы сетевого планирования [28, стр. 299–318].

Один подход к решению некоторых вариантов данной задачи (случай одностадийных техпроцессов) описывается в следующей главе.

### 1.3.3 Уровень 2: поставки, производство, сбыт

В модели уровня 1 предполагалось, что сырьё для производства имеется в любой момент в достаточном количестве, а также имеется достаточно места для хранения готовой продукции перед отправкой её потребителю/заказчику.

В модели уровня 2 происходит объединение модели производства уровня 1 с моделями управления запасами (склад, поставщики) и работы с потребителями.

В данном случае в систему  $S$  входит не только собственно производственное предприятие. В неё также включаются поставщики и потребители (возможно в виде бесструктурных элементов — чёрных ящиков).

#### Управление запасами

Пусть требуется хранить  $N_m$  видов сырья (полуфабрикатов, готовой продукции).  $i$ -й вид сырья описывается кортежем  $\langle i, b, t \rangle$ , где  $b$  — количество сырья  $i$ -го вида в момент времени  $t$ .

Пусть на складе имеется  $N_c$  контейнеров для хранения.  $i$ -й, контейнер описывается кортежем  $\langle i, c, v, j, k, t \rangle$ , где  $c$  — ёмкость контейнера;  $v$  — заполненный объём контейнера;  $j \in \{1, \dots, N_m\}$  — вид хранящегося сырья (если  $v > 0$ );  $k = \{j_1, \dots, j_{n_i}\}$  — виды сырья, которые *можно* хранить в контейнере.

Пусть для восполнения запасов имеется  $N_s$  поставщиков.  $i$ -й поставщик описывается кортежем  $\langle i, m_i \rangle$ , где  $m_i = \langle m_{i,1}, \dots, m_{i,n_i} \rangle$  — набор поставляемого сырья. Каждый элемент  $m_{i,j}$  имеет вид  $\langle k_{i,j}, v_{i,j}, \tau_{i,j} \rangle$ , где  $k_{i,j} \in \{1, \dots, M_n\}$  — вид сырья;  $v_{i,j}$  — мощность поставщика (максимальный объём поставки в единицу времени, максимальный объём разовой поставки);  $\tau_{i,j}$  — время от момента заказа до момента попадания сырья на склад.

В зависимости от вида величин  $\tau_{i,j}$ , а также от того, насколько велика вероятность изменения планов в процессе их выполнения (например, из-за возможности заключения новых, срочных контрактов) для описания динамики запасов можно использовать те или иные детерминированные или вероятностные модели управления запасами [28, стр. 471–504, стр. 607–624].

#### Сбыт

Пакет заказов  $R$  в данный момент времени можно представить в виде кортежа  $\langle R_1, R_2, R_3, R_4 \rangle$ , где  $R_1$  — заключённые контракты;  $R_2$  — незаключённые, но предполагаемые с большой степенью уверенности, контракты;  $R_3$  — возможные контракты;  $R_4$  — неожиданные контракты.

$i$ -й заказ описывается кортежем  $\langle i, p_i \rangle$ , где  $p_i = \langle j_i, v_i, t_i \rangle$  — набор продукции:  $j_i \in \{1, \dots, N_m\}$  — вид продукции;  $v_i$  — объём заказа;  $t_i$  — время поставки.

## Планирование производства

Объединяя модель уровня 1, управление запасами и сбыт, получаем модель уровня 2.

В модели уровня 2 планирование собственно производства в определённый момент времени  $t$  зависит от наличия в этот момент необходимых запасов сырья, а также от наличия места для хранения произведённой продукции. Так как в модель включены поставщики и потребители, имеется возможность более гибко управлять запасами сырья и готовой продукции, чтобы избежать и дефицита и излишков.

Заметим, что поставщики и потребители как составные части рассматриваемой системы  $S$  приносят неустранимую неопределённость, которую требуется учитывать при составлении планов. Различные варианты сведения неопределённых задач к детерминированным аналогам рассматриваются, например, в [18].

### 1.3.4 Уровень 3: отраслевые игры

Объединяя предприятия отрасли в одну систему, переходим к модели уровня 3. Составными частями такой системы являются предприятия–подсистемы, поставщики и потребители.

#### Одноуровневая игра

Рассматривая предприятия–подсистемы как активные элементы одного уровня, получаем задачу в игровой постановке:

$$\begin{aligned} f_i(x) &\rightarrow \max_{x_i}, \quad i = 1, \dots, n \\ g(x) &\geq 0. \end{aligned}$$

Здесь  $n$  — общее число предприятий,  $f_i$  — целевая функция модели уровня 2  $i$ -го предприятия, учитывающая деятельность остальных предприятий;  $x = (x_1, \dots, x_n)$  — вектор стратегий предприятий;  $g$  — общее описание ограничений, накладываемых на возможные значения  $x$ . Так как на результаты деятельности  $i$ -го предприятия влияет выбор других предприятий, аргументов функции  $f_i$  будет весь вектор  $x$ . Однако само  $i$ -е предприятие может выбирать только значение подвектора  $x_i$ .

В зависимости от способа действий активных агентов решением задачи может быть Парето-оптимальное решение, равновесие Неша, различные варианты кооперативного поведения игроков и т.п. См., например, [19].

#### Иерархические игры

Добавление в систему с одноуровневыми активными агентами управляющего элемента даёт систему с двухуровневой иерархической структурой:

$$\begin{aligned} F(x, y) &\rightarrow \max_y, \\ G(y) &\geq 0, \\ f_i(x, y) &\rightarrow \max_{x_i}, \quad i = 1, \dots, n \\ g(x, y) &\geq 0. \end{aligned}$$

Здесь  $F$  и  $G$  соответственно целевая функция и описание ограничений *ведущего* элемента системы.

1. Ведущий элемент задаёт значение  $y^*$  вектора  $y$ .
2. Ведомые элементы выбирают стратегии  $x^*$  поведения с помощью решения игровой задачи с заданным параметром  $y^*$ .

3. Ведущий элемент, получив значение  $x^*$ , вычисляет значение целевой функции в точке  $(x^*, y^*)$ .

Цель ведущего элемента (например, отраслевого министерства) — установить такие «правила игры»  $y$ , чтобы заставить ведомые элементы, действующие в своих интересах (максимизирующие свои целевые функции  $f_i$ ), в то же время действовать так, чтобы достигался максимум функции  $F$ .

Варианты решения задач двухуровневой оптимизации рассматриваются, например, в [34].

# Глава 2

## Некоторые методы принятия решений

### 2.1 Метод анализа иерархий

[26, 27]

- Цель
- Подцели
- Критерии
- Подкритерии
- Альтернативы

Метод попарных сравнений

### 2.2 Теория ограничений Голдратта

[8, 31]

- Дерево текущей реальности
- Дерево будущей реальности
- Диаграмма разрешения конфликтов
- Дерево перехода
- План преобразований

### 2.3 Формирование решений по управлению

[23]

- Цели, ситуации и планы
- Стратегии планирования
- Функциональные модели для планирования
- Сценарии и логики действий
- Экстраполяция решений
- Особенности рассуждений у человека

# Глава 3

## Системы поддержки принятия решений

- Выбор наилучшего (наилучших) решений из множества возможных.  
Оптимизация.
- Упорядочение всевозможных решений по предпочтительности.  
Ранжирование.

### Методы СППР

- информационный поиск;
- интеллектуальный анализ данных;
- поиск знаний в базах данных;
- рассуждение на основе прецедентов;
- имитационное моделирование;
- эволюционные вычисления и генетические алгоритмы;
- нейронные сети;
- ситуационный анализ;
- когнитивное моделирование;
- ...

### 3.1 Классификация СППР

#### Взаимодействие с пользователем

**Пассивные** помогают в процессе принятия решений, но не могут выдвинуть конкретного предложения

**Активные** непосредственно участвуют в разработке правильного решения

**Кооперативные** предполагают взаимодействие СППР с пользователем. Выдвинутое системой предложение пользователь может доработать, усовершенствовать, а затем отправить обратно в систему для проверки. После этого предложение вновь представляется пользователю, и так до тех пор, пока он не одобрит решение.

## Способы поддержки

- Модельно-ориентированные СППР, используют в работе доступ к статистическим, финансовым или иным моделям.
- СППР, основанные на коммуникациях, поддерживают работу двух и более пользователей, занимающихся общей задачей.
- СППР, ориентированные на данные, имеют доступ к временным рядам организации. Они используют в работе не только внутренние, но и внешние данные.
- СППР, ориентированные на документы, манипулируют неструктурированной информацией, заключённой в различных электронных форматах.
- СППР, ориентированные на знания, предоставляют специализированные решения проблем, основанные на фактах.

## Архитектура СППР

- Функциональные СППР
- СППР, использующие независимые витрины данных
- СППР на основе двухуровневого хранилища данных
- СППР на основе трёхуровневого хранилища данных

## 3.2 Структура СППР

- Информационные хранилища данных
- Средства и методы извлечения, обработки и загрузки данных
- Многомерная база данных и средства анализа OLAP
- Средства Data Mining



# Глава 4

## Данные и знания

### 4.1 Данные

[7, 13]

#### 4.1.1 Модели данных

##### Иерархическая модель

Иерархическая древовидная структура представления. Дерево, образованное попарными связями. Один корень. Все узлы, кроме корня, связываются только с одним узлом на более высоком по отношению к ним самим уровне.

Пример: организационная структура

- Организация
  - Отделы
    - \* Начальник
    - \* Сотрудники
    - \* Оборудование
      - Компьютеры
      - Факсы
  - Филиалы

##### Сетевая модель

Каждый порождённый элемент может иметь более одного исходного. Элемент может быть связан с любым другим без ограничений.

Сетевая база данных состоит из набора записей, соответствующих каждому экземпляру объекта предметной области, и набора связей между ними.

- Сотрудник  $i$
- Проект  $j$
- $(i, j)$

##### Реляционная модель

Вся информация представляется в виде прямоугольных таблиц.

Основные понятия реляционной алгебры: таблица, отношение, строка, столбец, первичный ключ.

## Модель на основе инвертированных списков

Похожа на реляционную модель, но хранимые таблицы и пути доступа к ним видны пользователям. При этом

- строки таблиц упорядочены системой в некоторой физической последовательности;
- для каждой таблицы можно определить произвольное число ключей поиска, для которых строятся индексы. Эти индексы автоматически поддерживаются системой, но явно видны пользователям.

## ER-модель

Модель данных, позволяющая описывать концептуальные схемы предметной области. Используется при высокоуровневом проектировании баз данных. В качестве стандартной графической нотации применяются диаграммы «сущность–связь».

### 4.1.2 Модель данных «сущность–связь»

Entity–relationship model

- Сущность — абстрактный объект определённого вида
- Атрибут — свойство сущности
- Связь — соединение между двумя и большим числом множеств сущностей

Диаграмма сущностей и связей

- прямоугольник — множество сущностей
- овал — атрибуты
- ромб — связи

Пример:

- Сущности
  - Movies (title, year, length, filmType)
  - Stars (name, address)
  - Studios (name, address)
- Связи
  - Stare-in (Movies, Stars)
  - Owns (Movies, Studios(1))
- Бинарные связи
  - один ко многим — каждый член множества  $E$  посредством связи  $R$  может быть соединён не более чем с одним членом  $F$ :  $E \rightarrow F$
  - один к одному:  $E \leftrightarrow F$ 
    - «Не более чем один»
    - Studios  $\leftrightarrow$  Presidents
  - многие ко многим:  $E - F$

- Многосторонние ( $n$ -арные связи) связи  
Contracts (Stars, Movies, Studios)
- Роли: одно и то же множество сущностей многократно упоминается в контексте единственной связи  
Sequel-of (Movies(Sequel)→Movies(Original))  
Contracts (Studio<sub>1</sub>, Studio<sub>2</sub>, Stars, Movie)
- Связи с атрибутами  
Contracts (Salary)

Преобразование многосторонних связей в бинарные: четырёхсторонняя связь Contracts заменяется соединяющим множеством сущностей Contracts.

Базовые классы и подклассы:

- is-a (Movies, Cartoons)
- is-a (Movies, Murder-Mysteries)

**Упражнение.** Требуется спроектировать структуру, адекватно описывающую базу данных поступающих в вуз, которая должна включать в себя информацию об абитуриентах, факультетах, направлениях подготовки, экзаменах и экзаменационных оценках.

**Упражнение.** Требуется спроектировать структуру, адекватно описывающую производственную базу данных, которая должна включать в себя информацию о заявках, процессорах и технологических цепочках.

## Принципы проектирования

- Достоверность: множества сущностей и атрибутов должны соответствовать реальным требованиям
- Отсутствие избыточности: объекты структуры не должны повторять друг друга
- Простота: в проект включаются только те структурные элементы, без которых совершенно нельзя обойтись
- Выбор подходящих связей
- Использование элементов адекватных типов

Множество сущностей  $E$  целесообразно заменить атрибутом или набором атрибутов нескольких других сущностей, если

- Все связи, в которые вовлечено множество  $E$ , обладают стрелками, указывающими на  $E$
- Атрибуты множества  $E$  в совокупности обязаны однозначно определять любую сущность множества

## Моделирование ограничений

- Ключ — атрибут или множество атрибутов, уникальным образом определяющие некоторую сущность в составе множества ограничений
- Ограничение уникальности: определённое значение в некотором контексте должно быть уникальным
- Ограничение ссылочной целостности: некоторое значение, на которое ссылается другой объект, должно существовать в базе данных
- Ограничение домена: значение атрибута должно выбираться из некоторого конечного множества значений или принадлежать определённому диапазону изменения
- Ограничение общего вида: произвольное ограничение, которое должно быть зафиксировано в базе данных

## Слабые множества сущностей

Ключ для некоторого множества сущностей приходится формировать на основе атрибутов, которые полностью или частично принадлежат другому множеству сущностей.

Crews: number, Unit-of(Studios:name)

Если множество  $E$  на диаграмме обозначено двойным прямоугольником, оно является слабым. Атрибуты множества  $E$ , отмеченные подчёркиванием, в совокупности с ключевыми атрибутами тех множеств сущностей, с которыми  $E$  соединено посредством связей типа «многие к одному», обозначенных двойными ромбами, образуют ключ множества  $E$

### 4.1.3 Реляционные базы данных

Данные представлены в виде набора двумерных таблиц (отношений, relations).

- Атрибуты: названия столбцов
- Схемы: наименования отношения и атрибутов этого отношения
- Кортежи: строки отношения

Кортеж содержит по одному компоненту для каждого атрибута отношения

- Домены: базовые типы

Каждый компонент кортежа должен быть атомарным, т.е. относиться к некоторому базовому типу. В качестве компонентов кортежа не разрешается использовать записи, множества, списки, массивы или иные составные объекты, которые допускают естественное разбиение на более мелкие элементы

- Отношения: множества кортежей
- Экземпляры отношения: конкретные множества кортежей отношения

Преимущества

- Имеется развитая теория реляционной модели данных
- Другие модели данных можно свести к реляционной модели
- Реляционной моделью поддерживаются специальные средства ускоренного доступа к информации

- Можно манипулировать данными не зная конкретной физической организации базы данных во внешней памяти
- Имеется стандартный высокоуровневый язык запросов

Функциональные отношения

$$A_1 A_2 \dots A_n \rightarrow B$$

Если два кортежа отношения  $R$  совпадают в атрибутах  $A_1, A_2, \dots, A_n$ , то они должны совпадать и в другом атрибуте,  $B$ .

Операции реляционной алгебры:

- объединение отношений — включает все строки отношений-операндов
- пересечение отношений — отношение, содержащее строки, которые входят одновременно в оба отношения-операнда
- разность отношений — строки, которые входят в первое отношение-операнд и не входят во второе
- произведение отношений — каждая строка первого отношения сцепляется с каждой строкой второго отношения-операнда
- деление отношений
- ограничение отношения — выбор строк отношения-операнда по некоторому условию
- проекция отношения — выбор из отношения-операнда определённых столбцов
- соединение отношений
  - соединение по условию — строка, конкатенации строк отношений-операндов, включается в результат, если она удовлетворяет заданному критерию
  - естественное соединение — сравнение производится на равенство значений общих для отношений-операндов атрибутов

Нормализация отношений

- Первая нормальная форма: значения атрибутов атомарны (в каждом столбце находится только одно значение), все неключевые атрибуты функционально зависят от ключа
- Вторая нормальная форма: выполняются ограничения 1NF и каждый неключевой атрибут функционально полно зависит от первичного ключа
- Третья нормальная форма: выполняются ограничения 2NF и все неключевые атрибуты отношения взаимно независимы и полностью зависят от первичного ключа (отсутствуют транзитивные зависимости неключевых атрибутов от ключа)
- ...

Средства ускорения доступа к данным

- Индексирование
- Хэширование

Некоторые операторы SQL:

SELECT	Выбрать данные из базы данных
INSERT	Добавить данные
UPDATE	Обновить данные
DELETE	Удалить данные
GRANT	Предоставить привилегии пользователю
REVOKE	Отменить привилегии пользователя
COMMIT	Зафиксировать текущую транзакцию
ROLLBACK	Прервать текущую транзакцию

#### 4.1.4 Хранилища данных

Предметно-ориентированный, интегрированный, неизменяемый и поддерживающий хронологию набор данных, предназначенный для обеспечения принятия управленческих решений.

- Ориентация на предметную область: хранилище должно разрабатываться с учётом специфики предметной области, а не приложений, оперирующих данными
- Интегрированность: информация загружается в хранилище из приложений, созданных разными разработчиками, поэтому необходимо объединить данные этих приложений, приведя их к единому синтаксическому и семантическому виду
- Неизменяемость данных: после загрузки в хранилище данные остаются неизменными, внесения каких-либо изменений, кроме добавления записей, не допускается
- Поддержка хронологии: в структуры хранилища данных вводятся ключевые атрибуты «дата» и/или «время»

Свойство	OLTP	СППР
Назначение данных	Оперативный поиск, несложные виды обработки	Аналитическая обработка, прогнозирование, моделирование
Уровень агрегации данных	Детализированные	Агрегированные
Период хранения данных	От нескольких месяцев до одного года	От нескольких лет до десятков лет
Частота обновления	Высокая частота, обновление маленькими порциями	Малая частота, обновление большими порциями

В хранилищах данных данные подготавливаются для последующего анализа

- Интеграция и согласование данных из различных источников
- Разделение наборов данных, используемых системами обработки транзакций и системами поддержки принятия решений

Цели

- Своевременное обеспечение аналитиков всей информацией, необходимой для выработки решения
- Создание единой модели данных организации
- Создание интегрированного источника данных, предоставляющего удобный доступ к разнородной информации и гарантирующего получение одинаковых ответов на одинаковые запросы из различных аналитических подсистем

Метаданные — высокоуровневые средства отражения информационной модели

- Описание структур данных хранилища
- Описание структур данных, импортируемых из разных источников
- Сведения о периодичности импортирования, методах загрузки и обобщения данных
- Средства доступа и правила представления информации
- Оценки приблизительных затрат времени на получение ответа на запрос

Хранятся в репозитории метаданных

#### 4.1.5 OLAP

OnLine Analysis Processing

Система поддержки принятия решений, основанная на концепции хранилища данных и обеспечивающая малое время выполнения аналитических запросов

#### MOLAP

Multidimensional OLAP

Основные понятия:

- измерение: множество, образующее одну из граней гиперкуба
- значения: качественные или количественные данные, которые находятся в ячейках гиперкуба

Основные операции манипулирования измерениями

- сечение: формируется подмножество гиперкуба, в котором значение одного или более измерений фиксировано
- вращение: изменение порядка представления измерений  
Обычно применяется к двумерным таблицам, для представления их в более удобной для восприятия форме
- детализация: переход от обобщённых данных к детализированным  
весь год → месяц
- свёртка: одно из значений измерения заменяется значением более высокого уровня иерархии  
месяц → весь год

#### ROLAP

Relational OLAP

Гиперкуб эмулируется СУБД на логическом уровне

Радиальная схема организации хранилища данных («звезда»)

- Таблицы фактов: данные, наиболее интенсивно используемые для анализа (соответствуют ячейке гиперкуба)
- Справочные таблицы (таблицы измерений): содержат возможные значения одного из измерений гиперкуба. Каждое измерение описывается отдельной справочной таблицей.

Фактологическая таблица индексируется по сложному ключу, скомпонованному из индивидуальных ключей справочных таблиц.

Для увеличения производительности в фактологической таблице могут храниться не только детализированные, но и предварительно вычисленные агрегированные данные.

Если имеется большое число измерений, то можно использовать схему «снежинка»: атрибуты справочных таблиц могут быть детализированы в дополнительных справочных таблицах.

## Киоски данных

Специализированные тематические хранилища, обслуживающие одно из направлений деятельности организации.

Данные, которые реально нужны конкретным аналитическим приложениям, выделяются в отдельный набор, который может быть реализован в многомерной базе данных.

## 4.2 Знания

[9, 11, 25]

- Априорные и апостериорные
- Процедурные, декларативные и неявные

Пирамида знаний

- Мудрость: способность использовать знания наилучшим образом
- Метазнания: знания о знаниях
- Знания: способность использовать информацию
- Информация: потенциальный источник знаний
- Данные: потенциальный источник информации
- Шум: отсутствие видимых признаков информации

### 4.2.1 Представление знаний

Сетевые модели

Семантические сети

$$H = \langle I, C_1, C_2, \dots, C_n, \Gamma \rangle$$

$I$  — множество информационных единиц,  $C_1, \dots, C_n$  — множество типов связей между информационными единицами,  $\Gamma$  — связи между информационными единицами.

MemberOf(Mary, FemalePerson): Mary ∈ FemalePerson

$$\forall x \in \text{Persons} \Rightarrow (\forall y \text{ HasMother}(x, y) \Rightarrow y \in \text{FemalePersons})$$

SubsetOf, MemberOf



## Фреймы

Протофрейм — жёстко фиксированная структура информационных единиц.

(Имя фрейма  
Имя слота 1 (значение слота 1)  
Имя слота 2 (значение слота 2)  
...  
Имя слота  $K$  (значение слота  $K$ )

При конкретизации фрейма ему и слотам присваивают конкретные имена и происходит заполнение слотов. Из протофреймов получаются фреймы-экземпляры.

## Логические модели

$$M = \langle T, P, A, B \rangle$$

$T$  — множество базовых элементов,  $P$  — множество синтаксических правил,  $A$  — множество аксиом,  $B$  — множество правил вывода.

$$\text{Bachelor} = \text{And}(\text{Unmarried}, \text{Adult}, \text{Male})$$
$$\text{Bachelor}(x) \Leftrightarrow \text{Unmarried}(x) \wedge \text{Adult}(x) \wedge \text{Male}(x)$$

## Продукционные системы

Продукции — правила вывода. Знания описываются в виде семантической сети. Вывод на знаниях.

$$(i); Q; P; A \Rightarrow B; N.$$

$i$  — имя продукции,  $Q$  — сфера применения продукции,  $A \Rightarrow B$  — ядро продукции («ЕСЛИ  $A$ , ТО  $B$ » или «ЕСЛИ  $A$ , ТО  $B_1$ , ИНАЧЕ  $B_2$ »),  $P$  — условие применимости ядра продукции,  $N$  — постусловия продукции.

## Процедурные знания

### Объектно-ориентированное моделирование

[5, 24]

### 4.2.2 Знания, данные, умения

[20]

Данные хранятся, знания преобразуются, умения применяются

## Уровень насекомого

Единичная циклическая жизненная программа, выполнение которой разнообразится непосредственными реакциями на внешние раздражители. И сама программа, и реакции заложены заранее и совершенствованию в течение жизни не подлежат.

Преимущества

- Насекомое почти невозможно сбить с толку
- Фиксированные рефлексы могут на самом деле быть достаточно сложными программами

Недостатки

- Полное отсутствие обучаемости

## Стереотипное реагирование

Уровень условных рефлексов: распознавание ситуации и применение в ней некоторого фиксированного действия с тем, чтобы получить желаемый результат.

Знание заложено в системе распознавания.

Данные: элементарные факты вида

$$P(c_1, \dots, c_n) \quad \text{или} \quad \neg P(c_1, \dots, c_n)$$

или их булевы комбинации.

Знания:

$$\forall x_1, \dots, x_n (A_1(x) \wedge \dots \wedge A_k(x) \Rightarrow Q(x))$$

$A_i$  — предикаты или их отрицания предикатов

Достоинства

- Быстрота ответа на знакомую реакцию, соединённая с возможностью перевести незнакомую ситуацию в разряд знакомых
- Стимул к обучению, когда ситуация распознаётся как не подходящая ни под одно из стандартных действий

Недостатки

- Возможность растерянности и паники при возникновении непредусмотренной ситуации
- Возможность ошибки при условии, что сразу несколько стереотипных действий кажутся применимыми

## Тупость

Тупой человек не может распознать ситуацию, но пытается применить то, что он видел в качестве успешных действий других.

Выход: затвердить несколько фиксированных правил (очень мало, чтобы не сбиваться), и выполнять данную систему правил как программу с инстинктивными реакциями.

## Комбинационное (комбинаторное) планирование

В комбинации есть целая последовательность (возможно, разветвлённая) действий, результат каждого из которых предсказуем, на настоящей целью осуществляющего комбинацию является лишь результат последнего действия. Начальные действия подготавливают почву для заключительных.

$$\forall x (A_1(x) \wedge \dots \wedge A_k(x) \Rightarrow B_1(x) \vee \dots \vee B_m(x))$$

Опасность: соблазн поиска лобовых решений там, где надо было бы переформулировать задачу в соответствии с системой ценностей и искать новую цель. Возможность пирровой победы.

Достоинства

- Возможность полной смены обстановки в результате последовательности хорошо спланированных действий
- Человек, комбинирующий свои приёмы, вынужден пересматривать их на предмет согласованности, и поэтому даже база стандартных приёмов у него обычно лучше структурирована, чем у того, кто работает на уровне условных рефлексов

- Комбинатор обычно оптимистичен, поскольку мир кажется ему единым целым, и наслаждение от поиска новых решений добавляет положительного настроения в его жизнь

#### Недостатки

- Необходимость заранее продумать и спланировать действия, как правило, противоречит скорости реакции
- Ненадёжность создаваемых планов, даже если каждое действие вполне надёжно и результат его предсказуем.

Чем длиннее и разветвлённее план, тем он уязвимее

- Люди, которым свойственно комбинаторное мышление, плохо приспособляются к ординарным ситуациям

Люди, владеющие комбинаторным мышлением вместе с логическими либо математическими орудиями, сразу видят несогласованности как новых концепций с тем, что считается догмой, так и внутри самих новых концепций.

#### Глупость

Человек не желает пользоваться известными ему рецептами, пытается рассчитывать вперёд, но сам себя запутывает в своих расчётах, не принимая во внимание возможные (а то и наиболее вероятные) последствия своих действий.

#### Стратегическое планирование и преобразование действий

Часто достаточно преобразовать рутинную формулу, чтобы она повернулась другой стороной.

Логический уровень преобразований — применение пропозициональных и предикатных эквивалентностей.

Применять композиции и преобразования вместе крайне опасно, поскольку ненадёжности обоих способов великолепно дополняют друг друга.

#### Достоинства

- Возможность представить имеющиеся знания в нужной форме или перенести их на случай, подобный рассмотренным, но не рассмотренный ранее.  
Здесь впервые знания становятся настоящими знаниями
- Преобразования позволяют лучше структурировать пространство знаний, и заставляют иметь единую модель мира, что обеспечивает цельность восприятия.
- Преобразования позволяют трезво оценить ситуацию, определить, в чём заключается стратегический выигрыш, и идти к нему, невзирая на частичные поражения.

#### Недостатки

- Отсутствие дальнего расчёта и игнорирование побочных эффектов преобразований
- Отсутствие быстрой реакции на мимолётные возможности
- Слишком часто преобразования ограничены узкой областью и базируются на излишне жёсткой модели мира, что ведёт к узости кругозора человека, работающего на данном уровне

## Релятивизм

Основной чертой релятивизма является наглое передёргивание фактов, положений, законов. Поскольку по виду эквивалентные предложения часто не имеют между собой ничего общего, а, соединяя комбинаторику и преобразования, можно из формально противоречивой системы знаний вывести всё, что угодно, человек зачастую превращается в законченного скептика и циника, тем более, что ему при этом способствует и практика, и теория нынешнего общества.

Возможный вывод в релятивизме: поскольку лишь точные науки выдерживают проверку комбинацией композиций и преобразований, то лишь они имеют право на существование, а гуманитарное знание знанием просто не является.

## Владение методом

Метод — это общий способ преобразования планов.

Характерная черта такого уровня — взаимосогласование модели мира и действий.

Применение метода требует подстановки в него целых планов и условий. Уровень метода соответствует тому, что часто называют системным взглядом.

Недостатки

- Иллюзия завершённости знания

Абсолютизация метода и успехи, связанные с его применением, вызывают некритическое восприятие и у самого человека, и (ещё чаще) у окружающих его учеников и почитателей

- Общий недостаток высших уровней — на них трудно удержаться

## Умничание, мессианство

Умничание: человек, на самом деле не овладев методом, пытается его применять и запутывается в конкретизациях высокоуровневых понятий.

Мессианство: человек, достаточно овладевший методом, чтобы его применять, но недостаточно им овладевший, чтобы понять, когда применять его не нужно.

## Многоуровневое мышление

Альтернативные модели мира и действий, каждая из которых внутренне согласована, и которые дают многомерный взгляд на проблему.

Недостатки

- По сравнению с человеком, владеющим методом, человек с многоуровневым мышлением становится менее активным. Он теряет ощущение целостности, приобретённое после овладения методом. Он начинает видеть недостатки и ловушки в слишком многих местах. Он не может быть единомышленником даже самому себе. В любой схватке он пытается, даже встав на одну из сторон, морально оказаться над битвой.
- Такой человек всё время подмечает недостатки в положительно охарактеризованном решении и достоинства в большинстве неудач. Это может сбить с толку прямолинейно мыслящих коллег.

## Мудрствование, интуитивизм

### Дао

Состояние понимания целостности мира.

Общая идея поведения: отшельничество и недеяние или полностью незаинтересованное активное действие.

## **Лжепророки**

Принимают свои озарения, которые даются им ценой максимального напряжения интеллектуальных и духовных сил, за абсолютную истину, и тем самым убивают элементы Истины, которые в них содержатся.

# Глава 5

## Data Mining

[2, 33, 35]

- Регрессия
- Классификация
- Кластеризация
- Поиск ассоциативных правил
- Разное

### 5.1 Качественные и количественные величины

[22]

- Качественные величины
  - Номинальные
  - Порядковые
- Количественные величины
  - Интервальная шкала
  - Шкала разностей
  - Шкала отношений
  - Абсолютная шкала

### 5.2 Регрессия

$(x_i, y_i)$

$$\hat{y} = \phi(x, \theta)$$

$y$  — количественная величина

### 5.3 Классификация

$(x_i, y_i)$

$$\hat{y} = \phi(x, \theta)$$

$y$  — качественная величина

5.4 Кластерный анализ

5.5 Ассоциативные правила

# Литература

- [1] Алефельд Г., Херцбергер Ю. *Введение в интервальные вычисления*. Пер. с англ. — М.: Мир, 1987. — 360 с.
- [2] Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. *Методы и модели анализа данных: OLAP и Data Mining*. — СПб.: БХВ-Петербург, 2004. — 336 с.
- [3] Беллман Р. *Динамическое программирование*. Пер. с англ. — М.: Издательство иностранной литературы, 1960. — 400 с.
- [4] Бурков В. Н. *Основы математической теории активных систем*. — М.: Наука, 1977. — 255 с.
- [5] Буч Г. и др. *Объектно-ориентированный анализ и проектирование с примерами приложений*. Пер. с англ. — 3-е изд. — М.: Издательский дом «Вильямс», 2008. — 720 с.
- [6] Волкова В., Денисов А. А. *Теория систем*. — М.: Высшая школа, 2006. — 511 с.
- [7] Гарсиа-Молина Г., Ульман Дж., Уидом Дж. *Системы баз данных. Полный курс*. Пер. с англ. — М.: Издательский дом «Вильямс», 2003. — 1088 с.
- [8] Детмер У. *Теория ограничений Голдратта: Системный подход к непрерывному совершенствованию*. Пер. с англ. — 2-е изд. — М.: Альпина Бизнес Букс, 2008. — 444 с.
- [9] Джарратано Р., Райли Г. *Экспертные системы: принципы разработки и программирование*. Пер. с англ. — 4-е изд. — М.: Издательский дом «Вильямс», 2007. — 1152 с.
- [10] Заде Л. *Понятие лингвистической переменной и его применение к принятию решений*. Пер. с англ. — М.: Мир, 1976. — 167 с.
- [11] *Искусственный интеллект*. — В 3-х кн. Кн. 2. *Модели и методы*. / Под. ред. Д. А. Поспелова — М.: Радио и связь, 1990. — 304 с.
- [12] Кини Р. Л., Райфа Х. *Принятие решений при многих критериях: предпочтения и замещения*. Пер. с англ. — М.: Радио и связь, 1981. — 560 с.
- [13] Корнеев В. В., Гареев А. Ф., Васютин С. В., Райх В. В. *Базы данных. Интеллектуальная обработка информации*. — М.: Нолидж, 2000. — 352 с.
- [14] Кремер Н. Ш., Путко Б. А., Тришин И. М., Фридман М. Н. *Исследование операций в экономике: Учеб. пособие для вузов*. — М.: ЮНИТИ, 2003. — 407 с.
- [15] Ларичев О. И. *Наука и искусство принятия решений*. — М.: Наука, 1979. — 200 с.
- [16] Ларичев О. И. *Объективные модели и субъективные решения*. — М.: Наука, 1987. — 144 с.
- [17] Ларичев О. И. *Теория и методы принятия решений, а также Хроника событий в Волшебных Странах*. — М.: Логос, 2000. — 296 с.



- [18] Лю Б. *Теория и практика неопределённого программирования*. Пер. с англ. — М.: БИНОМ. Лаборатория знаний, 2005. — 416 с.
- [19] Мулен Э. *Теория игр с примерами из математической экономики*. Пер. с франц. — М.: Мир, 1985. — 200 с.
- [20] Непейвода Н. Н., Скопин И. Н. *Основания программирования* — 2002.
- [21] Новиков Д. А., Петраков С. Н. *Курс теории активных систем*. — М.: СИНТЕГ, 1999. — 104 с.
- [22] Орлов А. И. *Эконометрика*. — М.: Экзамен, 2002. — 409 с.
- [23] Поспелов Д. А. *Ситуационное управление: теория и практика*. — М.: Наука, 1986. — 288 с.
- [24] Рамбо Дж., Блаха М. *UML 2.0. Объектно-ориентированное моделирование и разработка*. Пер. с англ. — 2-е изд. — СПб.: Питер, 2007. — 544 с.
- [25] Рассел С., Норвиг П. *Искусственный интеллект: современный подход*. Пер. с англ. — 2-е изд. — М.: Издательский дом «Вильямс», 2006. — 1408 с.
- [26] Саати Т. Л. *Принятие решений при зависимостях и обратных связях: Аналитические сети*. Пер. с англ. — М.: Издательство ЛКИ, 2008. — 360 с.
- [27] Саати Т., Кернс К. *Аналитическое планирование. Организация систем*. Пер. с англ. — М.: Радио и связь, 1991. — 224 с.
- [28] Таха Х. А. *Исследование операций*. Пер. с англ. — 7-е изд. — М.: Издательский дом «Вильямс», 2005. — 912 с.
- [29] Трахтенгерц Э. А. *Компьютерная поддержка принятия решений*. — Москва, 1998. — 246 с.
- [30] Трухачев Р. И. *Модели принятия решений в условиях неопределённости*. — М.: Наука, 1981. — 258 с.
- [31] Шрагенхайм Э. *Управленческие дилеммы: Теория ограничений в действии*. Пер. с англ. — М.: Альпина Бизнес Букс, 2007. — 288 с.
- [32] Штойер Р. *Многокритериальная оптимизация. Теория, вычисления и приложения*. Пер. с англ. — М.: Радио и связь, 1992. — 504 с.
- [33] Cios K. J., Pedrycz W., Swiniarsky R. W., Kurgan L. A. *Data Mining. A Knowledge Discovery Approach*. — Springer, 2007. — 600 p.
- [34] Dempe S. *Foundations of Bilevel Programming*. — Kluwer Academic Publishers, 2002. — 309 p.
- [35] Han J., Kamber M. *Data Mining: Concepts and Techniques*. — Elsevier Inc., Morgan Kaufmann Publishers, 2006. — 772 p.

# Приложение А

## Тестовые тесты

### А.1 Системы поддержки принятия решений

#### А.1.1 № 0

1. В чём различие между качественными и количественными величинами? Номинальными и порядковыми?

Какими величинами являются экзаменационные оценки? Почему? Как вычислить средний балл?

2. Пусть в портфеле ценных бумаг (модель Марковица) имеются три актива со случайными доходностями  $\xi_1$ ,  $\xi_2$  и  $\xi_3$ , для которых известны  $\bar{x}_i$ ,  $s_{x_i}^2$  и  $\text{cov}(x_i, x_j) < 0$  при  $i \neq j$ . На графике с координатами *доходность*–*риск* нарисуйте множество допустимых портфелей и эффективное множество.

Как изменятся графики при добавлении в портфель безрискового актива  $\eta$ ?

3. Для непрерывно распределённой в диапазоне  $[a, b]$  случайной величины  $\xi$  известна функции плотности вероятности  $p_\xi(x)$ . Как вычислить первый начальный момент, второй центральный момент, вероятности событий  $\xi \leq q$  и  $\xi \geq q$  (при  $q \leq a$ ,  $a < q < b$  и  $q \geq b$ )?

4. Функция  $f: \mathbb{R} \rightarrow \mathbb{R}$  как минимум дважды непрерывно дифференцируема. Как выглядят необходимое и достаточное условия существования экстремума этой функции? Существования минимума этой функции?

Сформулируйте необходимое и достаточное условия существования минимума функции многих переменных  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ .

5. Нарисуйте схему системы  $S$ , имеющей входы  $x_1$  (управляющая переменная) и  $x_2$  (неуправляющая переменная) и выход  $y$ . Входы и выход связаны функцией  $f$  с параметром  $\beta$ .

Как изменится схема при добавлении внешнего к  $S$  регулятора  $\Phi$ , образующего отрицательную обратную связь?

6. В игровой задаче «Дилемма заключённого» матрица выигрышей имеет вид

	$y_1$	$y_2$
$x_1$	(-1, -1)	(0, -10)
$x_2$	(-10, 0)	(-5, -5)

$x$  — стратегии первого игрока,  $y$  — второго.

Как выглядит Парето-оптимальное решение? Решение, оптимальное по Нэшу?

7. Есть мнение, что в России сильна коррупция. Опишите (и обоснуйте) эффект от проведения антикоррупционных демонстраций и митингов.  
В Китае взяточников расстреливают. Может ли это быть более эффективным? Почему? Предложите способ оценки эффекта от воздействия на систему.
8. Элементами множества  $X$  являются двухэлементные векторы  $x = (x_1, x_2)$ . Множество  $X$  разбито на три подмножества:  $X = X_1 \cup X_2 \cup X_3$  ( $X_i \cap X_j = \emptyset$  при  $i \neq j$ ).  
Как выглядит классификация методом бинарного выбора, если множества  $X_1$ ,  $X_2$  и  $X_3$  — линейно разделимы? Квадратично разделимы?  
Изменится ли процедура классификации, если известно, что классы  $X_i$  представляют *порядковые* качественные величины?
9. В песне поётся: «*В хоккее играют настоящие мужчины. Трус не играет в хоккей.*»  
Должен ли каждый настоящий мужчина играть в хоккей? Можно ли сказать, что мужчина, не играющий в хоккей является поддельным?  
Можно ли сказать, что женщина, играющая в хоккей, — мужчина?  
Для ответа используйте диаграмму Эйлера–Венна.
10. Мужчина  $m_2$  женился на вдове  $f_1$ , имеющей дочь  $f_2$ . Его отец  $m_1$  женился на  $f_2$ . Можно ли сказать, что  $m_i$  — отец  $m_j$  ( $i \neq j$ ) и дед  $m_i$ ?

### А.1.2 № 1

1. Опишите варианты постановки задачи оптимизации портфеля ценных бумаг.
  - Формальная постановка задачи.
  - Известны значения курса акции за прошлый год. Как оценить её доходность?
  - Виды неопределённости.
  - Способы сведения неопределённой задачи к детерминированной.
2. Опишите постановку задачи оптимизации работы промышленного предприятия, сельскохозяйственного предприятия.
3. Как можно сформулировать задачу, рассматриваемую в вашей выпускной работе?

## А.2 Принятие решений

### А.2.1 № 0

1. В чём различие между качественными и количественными величинами?
2. Какие виды неопределённости могут присутствовать в задачах принятия решений? Что является причиной их появления?
3. Нарисуйте график функции полезности  $U: \mathbb{R} \rightarrow \mathbb{R}$  для товаров первой необходимости, товаров роскоши.
4. Нарисуйте график производственной функции  $Y: \mathbb{R} \rightarrow \mathbb{R}$ .

5. Пусть в портфеле ценных бумаг (модель Марковица) имеются три актива со случайными доходностями  $\xi_1, \xi_2$  и  $\xi_3$ , для которых известны  $\bar{x}_i, s_{x_i}^2$  и  $\text{cov}(x_i, x_j) < 0$  при  $i \neq j$ . На графике с координатами *доходность–риск* нарисуйте множество допустимых портфелей и эффективное множество.  
Как изменятся графики при добавлении в портфель безрискового актива  $\eta$ ?
6. Функция  $f: \mathbb{R} \rightarrow \mathbb{R}$  как минимум дважды непрерывно дифференцируема. Как выглядят необходимое и достаточное условия существования экстремума этой функции? Существования минимума этой функции?  
Сформулируйте необходимые и достаточные условия существования минимума функции  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ .
7. Для функции  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  известна процедура, позволяющая вычислить значение функции для любых значений аргументов. Предложите метод оценки значения  $x^* = \text{argmin}_x f(x)$ .
8. Нарисуйте схему системы  $S$ , имеющей входы  $x_1$  (управляющая переменная) и  $x_2$  (неуправляющая переменная) и выход  $y$ . Входы и выход связаны функцией  $f$  с параметром  $\beta$ .  
Как изменится схема при добавлении внешнего к  $S$  регулятора  $\Phi$ , образующего отрицательную обратную связь?
9. Система  $S$  имеет вход  $x$  и выход  $y$ . Функция  $U$  используется для оценки качества поведения системы.  
Как обеспечить желаемое поведение системы  $S$ ? Подавить её нежелательное поведение?
10. Есть мнение, что в России сильна коррупция. Опишите (и обоснуйте) эффект от проведения антикоррупционных демонстраций и митингов.  
В Китае взяточников расстреливают. Может ли это быть более эффективным? Почему?

### А.2.2 № 1

1. Опишите варианты постановки задачи оптимизации портфеля ценных бумаг.
  - Формальная постановка задачи.
  - Известны значения курса акции за прошлый год. Как оценить её доходность?
  - Виды неопределённости.
  - Способы сведения неопределённой задачи к детерминированной.
2. Опишите постановку задачи оптимизации работы промышленного предприятия, сельскохозяйственного предприятия.
3. Как можно сформулировать задачу, рассматриваемую в вашей выпускной работе?

### А.2.3 № 2

1. [28, стр. 591] Армия полковника Блотто сражается с вражеской армией за контроль над двумя стратегически важными позициями. Полковник имеет в своём распоряжении два полка, а его противник — три. Каждый из противников может посылать на любую позицию только целое число полков или совсем не посылать. Позиция будет захвачена армией, которая атакует бóльшим количеством полков, иначе результат сражения является ничейным.  
Какая армия выиграет сражение?  
Что изменится, если убрать условие целочисленности?

2. [26, стр. 40] Семья со средними доходами хочет купить дом. Определены восемь критериев для выбора лучшего дома, распределённые на три категории: экономические, географические и физические: размер дома, транспортное сообщение, ближайшие окрестности дома, возраст дома, пространство двора, современные средства обслуживания, общее состояние дома, финансовые условия.

Сформулируйте задачу выбора наилучшей альтернативы методом анализа иерархий.

3. [26, стр. 84] Сформулируйте задачу (варианты задачи) о строительстве нового стадиона XYZ-клубом.

- Выгоды
  - Экономический критерий: использование для нужд города; привлечение средств из фондов; налогообложение строителя
  - Социальный критерий: статус первоклассного города; солидарность граждан; отдых и развлечения
- Возможности: сохранение городской XYZ-команды; дополнительные рабочие места; развитие бизнеса; освоение неиспользуемой территории города; проведение спортивных игр
- Издержки: увеличение налогов; модификация инфраструктуры; уменьшение финансирования других программ; стоимость земли; потеря доходов от старого стадиона
- Риски: снижение престижа города; команда может покинуть город (в будущем), отсутствие гражданской солидарности; долги; появление новых социальных проблем

4. В процессе решения задачи из пункта 3 была получена следующая таблица:

	Выгоды	Возможности	Издержки	Риски
Строить	0.696	0.815	0.769	0.576
Не строить	0.304	0.185	0.231	0.424

Проанализируйте возможные варианты решения.

5. [18, стр. 93] В задаче размещения-распределения объектов требуется определить координаты размещения новых объектов таким образом, чтобы стоимость транспортировки от объектов к потребителям была минимальной.

Объекты:  $i = 1, \dots, n$ . Координаты размещения  $i$ -го объекта:  $(x_i, y_i)$ . Производительности  $i$ -го объекта:  $s_i$ .

Потребители:  $j = 1, \dots, m$ . Координаты размещения  $j$ -го потребителя:  $(a_j, b_j)$ . Неопределённый спрос  $j$ -го потребителя:  $\xi_j$ .

Объём поставок от объекта  $i$  потребителю  $j$  после того, как величины спроса  $\xi_j$  приняли конкретные значения (реализации):  $z_{i,j}$ .

Опишите возможные варианты формальной постановки задачи.

6. [8, стр. 207] Проанализируйте диаграмму разрешения конфликта:

- В рабочем состоянии всегда находится весь парк оборудования.
  - Быстро восстанавливать неисправное оборудование ← Все ресурсы сосредоточить на ремонте неисправного оборудования
  - Предотвращать поломки исправного оборудования ← Все ресурсы сосредоточить на плановом техобслуживании оборудования