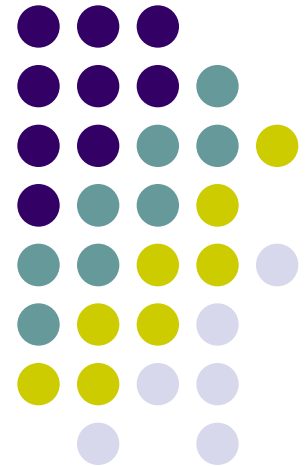
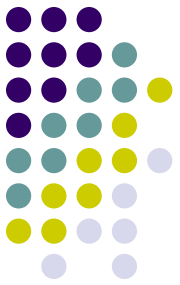


**ПРЕДСТАВЛЕНИЕ
ЗНАНИЙ:**

НЕЧЕТКИЕ ЗНАНИЯ

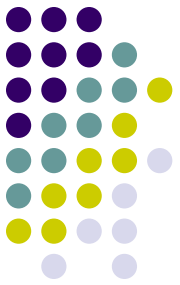


СОДЕРЖАНИЕ



1. Формализация нечётких знаний
2. Нечёткие множества и нечёткая логика
 - лингвистические и нечеткие переменные
3. Нечёткие знания и неточный вывод в ПС
 - механизм неточного вывода
 - меры достоверности
4. ЭС с нечеткими знаниями
 - MYCIN: коэффициенты уверенности
 - PROSPECTOR : байесовское дерево
5. Представление знаний: шкалы
6. Заключение

НЕЧЕТКИЕ ЗНАНИЯ



Истоки – *не-факторы*: неопределённость/
нечёткость/ неполнота знаний,
имеют разнообразный характер:

- Недоопределённость: не все объекты и связи ПО можно априори полностью определить
- Нечёткость качественных оценок (*красиво, хорошо, часто, молодой/старый*)
- Неточность получаемых количественных оценок опытных данных и эмпирических фактов (при замерах оказывает влияние погрешность прибора)
- Многозначность интерпретаций ЕЯ фраз
- Неоднозначность, ненадёжность выводов, в частности: недетерминированность выводов в ПС
- Некорректность (противоречивость) БЗ

НЕЧЕТКИЕ ЗНАНИЯ: ФОРМАЛИЗАЦИЯ



Разработаны специальные формализмы и логики, отличающиеся от классической

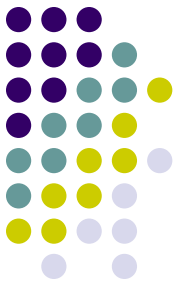
- Нечеткие множества (*fuzzy sets*) – теория Л.Заде
Лингвистические переменные
- Нечёткая логика (*fuzzy logic*)
- Вероятностная логика (основана на теореме Байеса)
- Модальная логика

ЛИНГВИСТИЧЕСКИЕ ПЕРЕМЕННЫЕ



- *Лингвистическая переменная*: ее значения – слова и фразы ЕЯ с неточным значением:
 - *Возраст – молодой / зрелый / пожилой*
 - *Температура – высокая / нормальная / низкая*
- *Нечеткая переменная (молодой, высокий, горячий)* может принимать значения из некоторого интервала, границы которого могут меняться в зависимости от контекста:
 - *холодный день*: границы меняются в зависимости от сезона: зима/весна/лето/осень
 - *старый атлет* ($\approx$ 36-40 лет) намного моложе *молодого пенсионера* ($\approx$ 55-60 лет)
 - *маленький слон* больше *крупной мыши*
- Нечеткие переменные описываются
нечёткими множествами:
Интервальный способ представления неточности

НЕЧЕТКИЕ МНОЖЕСТВА



- ❖ Вводится универсум U , на котором будут определять нечёткие множества
- ❖ Каждому элементу x универсума ставится в соответствие функция

$\mu_M(x)$ – *характеристическая функция* множества M :

$$\mu_M(x)=1, \quad \text{если } x \in M$$

$$\mu_M(x)=0, \quad \text{если } x \notin M$$

$$\mu_M(x)=y, \quad y \in (0,1), \quad \text{в ином случае}$$

y – *истинностное значение* принадлежности x к M

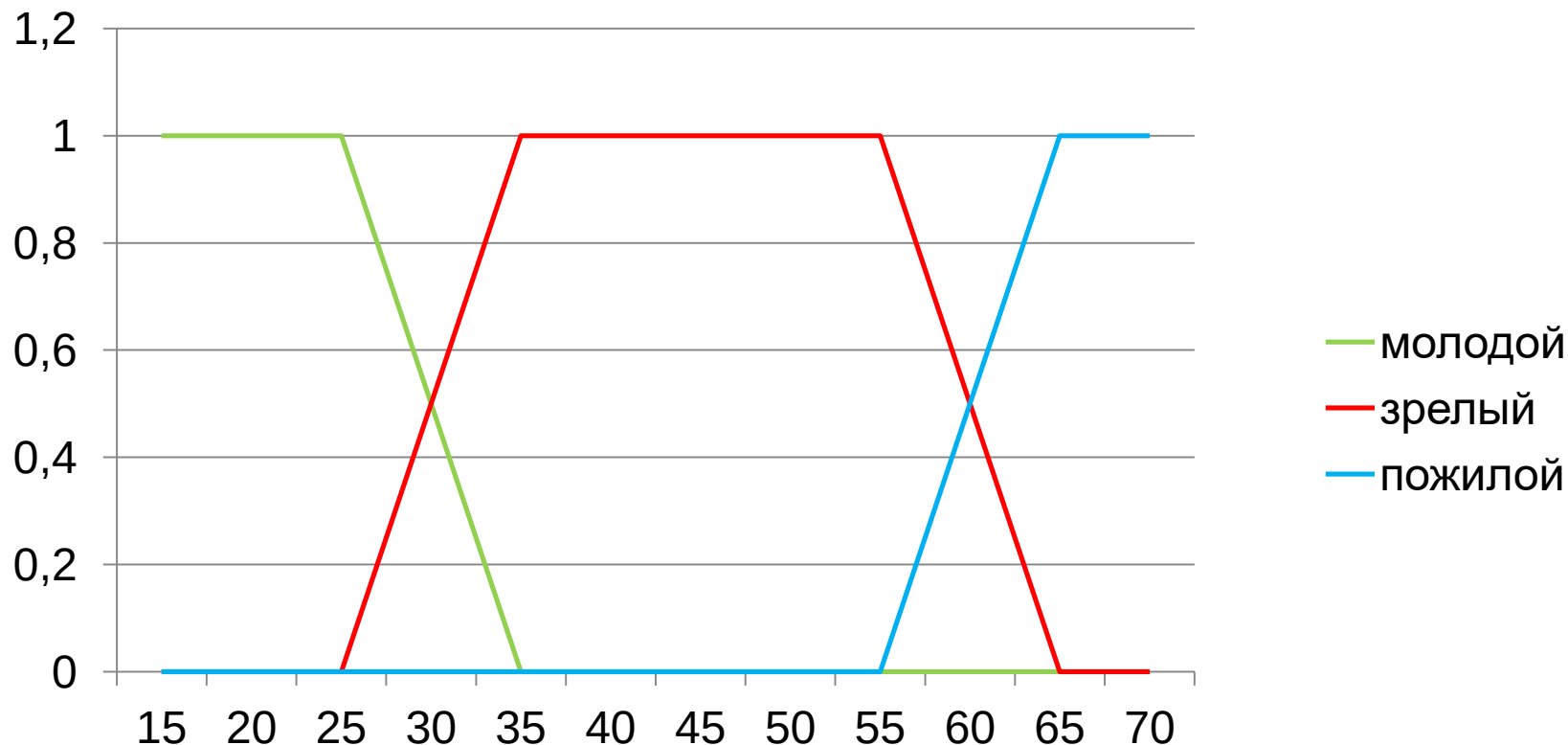
- ❖ Такое *истинностное значение* – произвольное субъективное значение, не имеющее в общем случае никакого статистического/вероятностного смысла
- ❖ Множество M описывается в виде набора пар:

$$x / \mu_M(x), \quad \text{где } x \in U$$

ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ НЕЧЕТКИХ МНОЖЕСТВ



Пример: Лингв.переменная – возраст человека:
молодой / зрелый / пожилой (нечеткие переменные)



НЕЧЕТКИЕ МНОЖЕСТВА: ОПЕРАЦИИ



Операции над нечёткими множествами:

- Пересечение (конъюнкция)
- Объединение (дизъюнкция)
- Дополнение (отрицание)

Фактически,

для любых двух нечётких множеств $M1$ и $M2$ производятся соответствующие операции с их характеристическими функциями:

$$\mu_{M1}(x) \& \mu_{M2}(x) = \min [\mu_{M1}(x), \mu_{M2}(x)]$$

$$\mu_{M1}(x) \vee \mu_{M2}(x) = \max [\mu_{M1}(x), \mu_{M2}(x)]$$

$$\neg \mu_M(x) = 1 - \mu_M(x)$$

НЕЧЕТКИЕ МНОЖЕСТВА: ПРИМЕР



Пусть универсум: множество элементов

$$\{ a, b, c, d, e, f \}$$

$$M1 = \{ a/1, \quad b/0.3, \quad c/0.8, \quad d/0, \quad e/0.5, \quad f/0.2 \}$$

$$M2 = \{ a/0.4, \quad b/1, \quad c/0.7, \quad d/0.9, \quad e/0, \quad f/0.6 \}$$

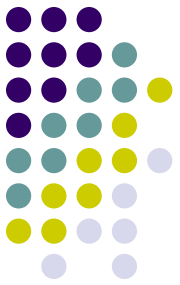
Тогда:

$$M1 \cap M2 = \{ a/0.4, \quad b/0.3, \quad c/0.7, \quad d/0, \quad e/0, \quad f/0.2 \}$$

$$M1 \cup M2 = \{ a/1, \quad b/1, \quad c/0.8, \quad d/0.9, \quad e/0.5, \quad f/0.6 \}$$

$$\neg M1 = \{ a/0, \quad b/0.7, \quad c/0.2, \quad d/1, \quad e/0.5, \quad f/0.8 \}$$

НЕЧЕТКАЯ ЛОГИКА



- Универсум: множество высказываний A_i
- Высказывание нечеткое, если его истинность не может быть точно установлена
- Вводится функция $t(A_i)$, определяющая истинность/ложность высказывания A_i : $t(A_i) \in [0,1]$
- Значение $t(A_i)$ может быть интерпретировано как *степень достоверности* A_i , в общем случае оно не имеет статистического/вероятностного смысла
- Скалярный способ представления неточности
- Алгебра нечеткой логики:

$$t(A_1 \& A_2) = \min \{t(A_1), t(A_2)\}$$

$$t(A_1 \vee A_2) = \max \{t(A_1), t(A_2)\}$$

$$t(\neg A) = 1 - t(A)$$

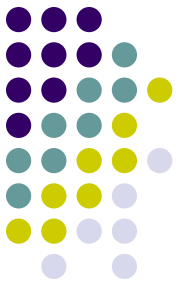
ПРИМЕНЕНИЕ НЕЧЕТКИХ МНОЖЕСТВ И НЕЧЕТКОЙ ЛОГИКИ



Широкое применение нечетких множеств
и нечеткой логики с 80-х гг.:

- Экспертные системы с неточными знаниями
- Сфера финансов и экономики
- Транспортная, авиа и аэрокосмическая промышленность
- Бытовая техника: управляющие микроконтроллеры на основе нечеткой логики, например, в современных стиральных машинах и пылесосах
- Ситуационные центры руководителей западных стран
- Программные продукты (*RuleMaker*, *FuziCalc*, ...)

ПРОДУКЦИОННЫЕ СИСТЕМЫ: НЕЧЁТКИЕ ЗНАНИЯ



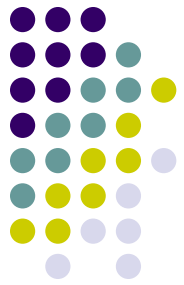
Модель оперирования с неточными знаниями:

- ❖ Способ представления меры точности/достоверности высказываний – *скалярный* либо *интервальный*
- ❖ Механизм вывода на неточных знаниях

Два типа механизмов вывода:

- Первый тип – механизмы, носящие «присоединённый характер»: пересчет мер достоверности присоединен к правилам вывода для точных высказываний
- Второй тип – механизмы, для которых характерно наличие особых схем вывода, специально ориентированных на используемый язык представления неточности

МЕХАНИЗМ НЕТОЧНОГО ВЫВОДА В ПС



Простейший механизм неточного вывода
присоединённого характера для ПС

с правилами продукции вида $r: p \rightarrow q$
получается путём введения функций для пересчёта:

- Меры достоверности $t(p)$ левой части правила по мерам достоверности составляющих его частей:

$$t(p) = f_p(t(e_1), t(e_2), \dots, t(e_n))$$

- Меры достоверности правой части (заключения правила) по мерам достоверности левой части и самого правила

$$t(q) = f_r(t(r), t(p))$$

где $t(e_k)$ – мера достоверности утверждения/факта e_k в РП
 $t(r)$ – мера достоверности правила r (достоверности
заключения при истинности посылок)

ДОСТОВЕРНОСТЬ ГИПОТЕЗЫ



В ПС некоторое утверждение/факт может быть выведен несколькими возможными способами, и в случае нечеткого вывода

- Необходима объединённая мера $t_{\Sigma}(h)$ достоверности выводимого факта-гипотезы h по мерам достоверности $t_1(h), t_2(h), \dots, t_k(h)$ нескольких ее доказательств:

$$t_{\Sigma}(h) = f_h(t_1(h), t_2(h), \dots, t_k(h))$$

- Функция f_h должна быть:
 - монотонно возрастающей
 - симметричной (порядок $t_k(h)$ не важен)
 - больше любой меры $t_k(h)$

Например: $f_h(t_1, t_2) = t_1 + t_2 - t_1 * t_2$

ВЫЧИСЛЕНИЕ МЕР ДОСТОВЕРНОСТИ



- Для вычисления меры достоверности $t(p)$ левой части правила могут быть использованы операции нечеткой логики (их суперпозиции):

$$t(e_1 \& e_2) = \min \{t(e_1), t(e_2)\}$$

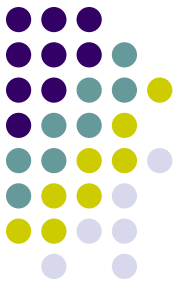
$$t(e_1 \vee e_2) = \max \{t(e_1), t(e_2)\}$$

$$t(\neg e) = 1 - t(e)$$

где $t(e_k)$ – мера достоверности утверждения e_k

- Один из способов вычисления достоверности (истинностного значения) заключения q правила продукции $r: p \rightarrow q$: $t(q) = \min \{ t(p), t(r) \}$
- Пример: $r: e_1 \& e_2 \vee \neg e_3 \rightarrow q$
 $t(e_1) = 0.67, \quad t(e_2) = 0.8, \quad t(e_3) = 0.4, \quad t(r) = 0.69$
 $t(q) = \min \{ \max [\min (0.67, 0.8), 1 - 0.4], 0.69 \} = 0.67$

ЭС С НЕЧЕТКИМИ ЗНАНИЯМИ



- Первые две известные ЭС:
MYCIN и *PROSPECTOR*, в обеих:
 - Скалярная форма представления неточности
 - ПС с присоединенным механизмом неточного вывода
- Цель обеих ЭС: минимизация временных и финансовых затрат в ходе принятия решений
 - MYCIN – диагностика и подбор лекарств для инфекционных заболеваний крови
 - PROSPECTOR – рекомендации по геологоразведочным мероприятиям в ходе поиска месторождений полезных ископаемых

ЭКСПЕРТНАЯ СИСТЕМА MYCIN



- Назначение: диагностика инфекционных заболеваний крови и поиск адекватных лекарственных средств (антибиотиков)
- Экспертные знания носят эвристический характер:
«скорее всего...», «в большинстве случаев...»
- Исходным фактам и правилам приписываются *коэффициенты уверенности (certainty factors) =* меры достоверности
- Для правила продукции этот коэффициент (*КУ*) означает степень надёжности заключения при *ИСТИННОСТИ* посылок



ЭС MYCIN: РАССЧЕТЫ

- Коэффициент уверенности ($KУ$) предпосылки правила вычисляется по формулам нечёткой логики
- $KУ$ заключения правила – произведение $KУ$ самого правила и $KУ$ предпосылки этого правила
- Для суммирования нескольких доказательств гипотезы h были введены:
 - *мера доверия mb* , отражающая наличие свидетельств «за»
 - *мера недоверия md* опровергающих свидетельств
 - вычисление обеих мер – с использованием вероятностей и *теоремы Байеса*

$$mb(\neg h \mid e) = md(h \mid e)$$

ТЕОРЕМА БАЙЕСА

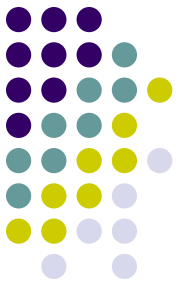


- *Условная вероятность* $P(A | B)$ – вероятность события A с учетом того, что произошло событие B
- $P(A \& B) = P(A | B) * P(B)$, тогда $P(A | B) = P(A \& B) / P(B)$
 $P(B)$ – априорная или безусловная вероятность
- Задача – вычисление *обратной вероятности* $P(B | A)$
вероятность предыдущего события B при условии, что произошло последующее событие A
- Формула Байеса
$$P(B | A) = \frac{P(A | B) * P(B)}{P(A)}$$

апостериорная или *обратная вероятность* события B

- Апостериорные вероятности позволяют пересматривать значения априорных вероятностей для получения более точных результатов

ЭС МУСИН: ФОРМУЛЫ



- Формулы расчета mb и md :

$$mb(h | e) = \begin{cases} 1, & \text{если } p(h)=1 \\ \frac{\max[p(h | e), p(h)] - p(h)}{1 - p(h)} \end{cases}$$

$$md(h | e) = \begin{cases} 1, & \text{если } p(h)=0 \\ \frac{p(h) - \min[p(h | e), p(h)]}{p(h)} \end{cases}$$

- Объединённая мера доверия гипотезы h по мерам доверия двух различных доказательств:

$$mb(h | e_1, e_2) = mb(h | e_1) + mb(h | e_2) * [1 - mb(h | e_1)]$$

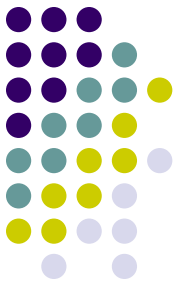
- Смысл формулы: эффект второго свидетельства e_2 на гипотезу h при заданном свидетельстве e_1 – в смещении mb в сторону полной определённости на расстояние, зависящее от второго свидетельства

MYCIN: КУ ГИПОТЕЗЫ



- Для md аналогичная формула:
$$md(h \mid e_1, e_2) = md(h \mid e_1) + md(h \mid e_2) * [1 - md(h \mid e_1)]$$
- Два важных свойства этих формул:
 - симметричность относительно e_1 и e_2
 - монотонное возрастание mb и md по мере накопления подкрепляющих/опровергающих свидетельств
- Коэффициент уверенности гипотезы, $KУ$ (*certainty factor*):
$$cf(h \mid e) = mb(h \mid e) - md(h \mid e)$$
$$-1 \leq cf(h \mid e) \leq 1$$
- Однако в итоге этот $KУ$ при выводе в ЭС не применялся, вывод был упрощен

MYCIN: ОСОБЕННОСТИ ВЫВОДА



- Логический вывод на основе правила продукции считается верным, если KU заключения превышает заранее заданное **пороговое значение**
- Пороговое значение $KU \geq 0.2$ – подобрано опытным путем

Важно: Коэффициент уверенности не отождествляется с условной вероятностью:

- ❖ Эксперты отказывались принимать положительное значение вероятности на основании опровергающего доказательства (хотя по теории вероятностей из опровержения гипотезы с вероятностью 0.7 следует подтверждение её с вероятностью 0.3)
- ❖ Субъективная вера людей вступает в противоречие с аксиомами теории вероятностей

ЭКСПЕРТНАЯ СИСТЕМА *PROSPECTOR*



- Назначение: определение перспективных площадок для разведки полезных ископаемых, получение рекомендаций о проведении или отказе от проведения геологоразведочных мероприятий (сейсмические исследования, бурение скважины и т.д.)
- Правила продукции задают априорную **сеть логического вывода**, событиям (действиям) задаются априорные вероятности
- Сеть логического вывода состоит из нескольких уровней, промежуточные гипотезы используются для вывода гипотез верхнего уровня
- Использован **метод байесовского принятия решений**
- Рассчитываются апостериорные вероятности событий, используемые для вычисления возможных прибылей или убытков компании

ЭС *PROSPECTOR*: РАССЧЕТЫ



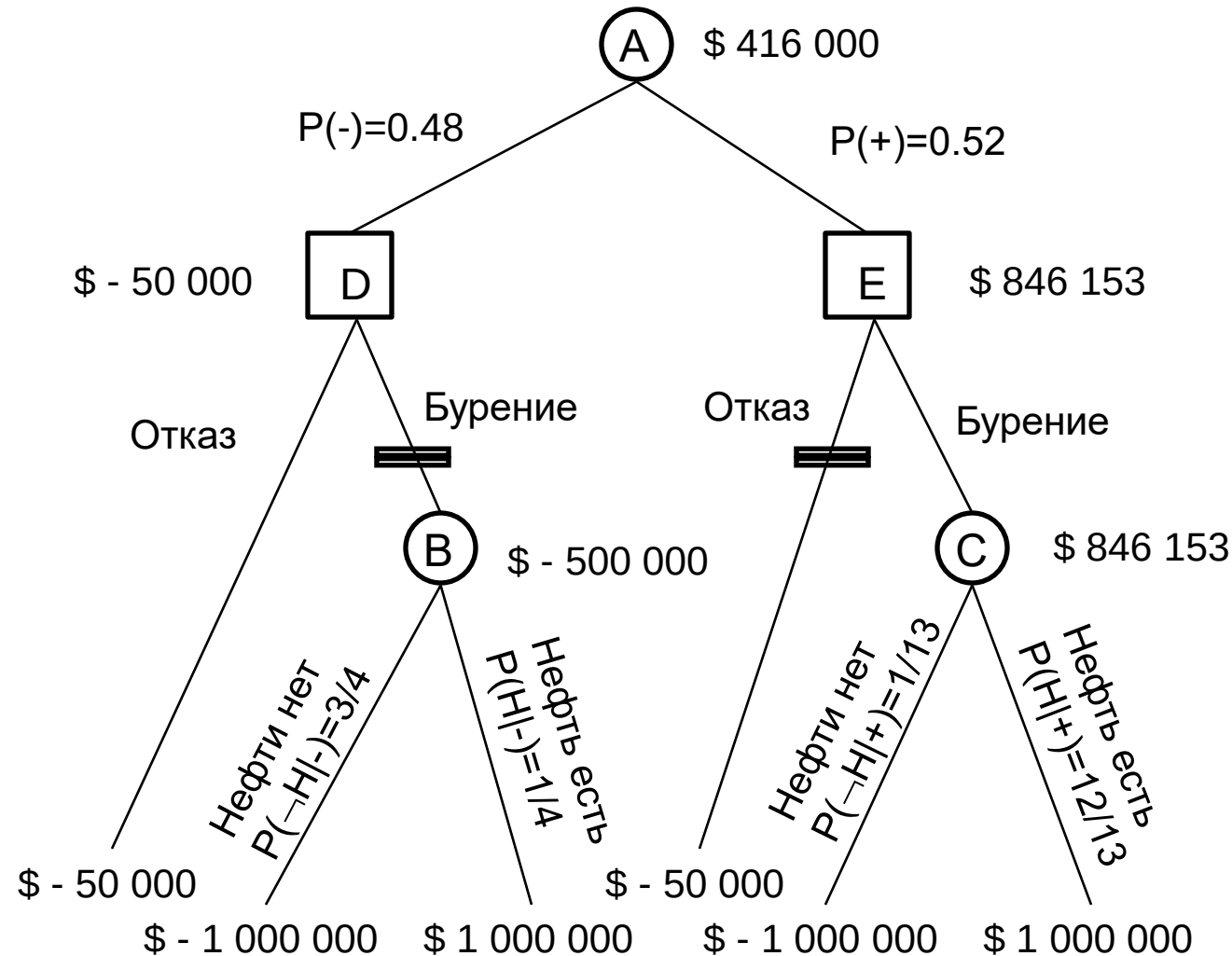
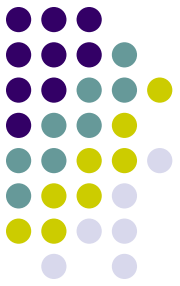
- Построение дерева (сети) вывода и вычисление апостериорных вероятностей событий
(= коэффициентов достоверности)
с учетом априорных вероятностей как положительных, так и отрицательных результатов
(наличие/отсутствие месторождения):

$$P(S) = P(S | E) * P(E) + P(S | \neg E) * P(\neg E)$$

- Использование минимаксной процедуры и отсечения неперспективных ветвей дерева
для вычисления возможных прибыли или убытка компании от принятия решений о проведении или отказе от проведения геологоразведочных мероприятий

ЭС PROSPECTOR: ПРИМЕР

Полное байесовское дерево решений
для задачи разведки нефти



Событие

Результат проверки
сейсмич. исследований
положительный (+) или
отрицательный (-)

Действие

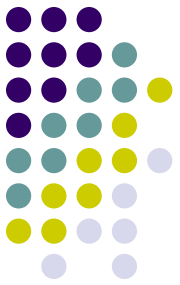
Отказ от дальнейших
исследование или
разведывательное
бурение

Событие

Нефть имеется или
отсутствует

Выигрыш

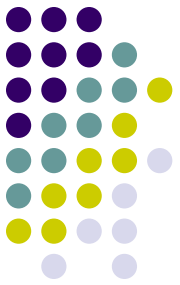
ПРЕДСТАВЛЕНИЕ ЗНАНИЙ: ШКАЛЫ



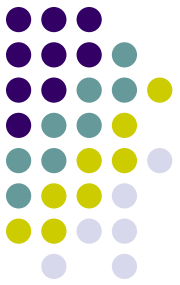
Для фиксации соотношений отдельных информационных единиц человек использует различные шкалы:

- *Метрические*, с помощью которых можно устанавливать количественные соотношения и порядок информационных единиц;
прежде всего: шкалы времени и пространства
- *Порядковые*, в которых фиксируется лишь порядок информационных единиц:
например, шкала оценок: *неуд / уд / хор / отл*
- *Оппозиционные (биполярные)* основаны на бинарной оппозиции: *сильный / слабый, хороший / плохой, острый/тупой...*

МЕТРИЧЕСКИЕ ШКАЛЫ



- ❖ Формируются в процессе развития человека
- ❖ Используются также процедуры перехода между этими шкалами (например: год – месяц)
- ❖ Пример: упорядочение по значению возраста
- ❖ Подвиды метрических шкал:
 - абсолютные: фиксированная точка отсчета (например, дата)
 - относительные: точка отсчёта может изменяться или точно не известна, например:
два часа езды от Москвы
3 км от Москвы
через 2 часа после этого разговора



ПОРЯДКОВЫЕ ШКАЛЫ

- ❖ Фиксируется лишь порядок информационных единиц, например: *солдат / сержант / офицер*
- ❖ Особый вид – *размытые порядковые шкалы*:
никогда / редко / нечасто / нередко / часто / всегда
 - Границы каждой информационной единицы размыты, зависят от индивидуального восприятия
 - Свойства размытых шкал аналогичны лингвистическим и нечетким переменным:
недалеко / близко, хорошо /отлично / превосходно
 - Для формализации размытой шкалы можно использовать технику нечётких множеств и нечеткую логику

ОПОЗИЦИОННЫЕ ШКАЛЫ



- ❖ Первые попытки построения: Ч. Осгуд (когнитивный психолог), в 50-х гг.
- ❖ Образуются при помощи пар слов-антонимов: *высокий/низкий, красивый/безобразный, острый/тупой*
- ❖ Маркированы лишь концы шкалы – абсолютные значения соответствующих качеств и свойств
- ❖ На шкалах и их совокупности можно отображать разные оценочные сведения об окружающем мире
- ❖ Эти оценочные шкалы можно проецировать одну на другую (в том числе абстрактную на физическую и наоборот), что объясняет метафоричность мышления: *недалёкий человек, острый ум, тяжёлое горе*
- ❖ Осгуд – три базовые шкалы (определяются методом факторного анализа): **оценка** (хорошо/плохо), **сила** (сильный/слабый), **активность** (активный/пассивный)

ЗАКЛЮЧЕНИЕ



Человеческие знания по своей природе

- принципиально не полны
- принципиально противоречивы
- ситуативны, эвристичны

Модель представления знаний должна отражать эти особенности, поэтому необходимо представление неточных знаний и возможность оперирования с ними



СПАСИБО ЗА ВНИМАНИЕ!