Рассуждения в условиях неопределенности

Проблемы

- Составить план действий вида «Успеть на занятия». Например, «*Собраться и выйти из дома в 8:30, добраться до остановки, ожидать. Если в течение 15 минут появился автобус или маршрутка – воспользоваться. Иначе ждать еще 15 маршрутку, иначе поймать такси*». Заменяем на более слабое утверждение: план позволит достичь цели при условиях отсутствия пробок. Однако логически установить успешность плана нельзя.

Проблемы логики первого порядка

- *Экономия усилий*. Для формирования множества антецедентов и консеквентов правила без исключений требуется много усилий, а применение такого правила – проблемно;
- *Отсутствие теоретических знаний*. Медицинская наука не располагает исчерпывающей теорией.
- *Отсутствие практических знаний*. Все необходимые обследования выполнить сложно.

Особенности

- *Свидетельство* – результат восприятия;
- Используются *коэффициенты уверенности* из [0..1] – не путать с коэффициентами истинности в нечеткой логике!
- Необходимо иметь возможность получить информацию о *предпочтениях* между возможными *результатами* – используется *теория полезности*.

Теория решений = теория вероятности + теория полезности

Основы

- Отсылка к некоторой части «мира», состояние которой изначально неизвестно. Например, *Cavity* может быть представлено случайной переменной.
- Разделяем булевы, дискретные и непрерывные случайные величины.
- Рассматриваются *атомарные события* – взаимно исключающие, и взаимно исчерпывающие.
- Если нет опытных данных – рассматриваем априорную вероятность, иначе – апостериорную. Есть распределение случайной величины априорное.
- Полное совместное распределение вероятностей – все варианты с указанием вероятностей комбинаций.

Условная вероятность

- Вероятность наступления события:

$$P(Cavity|Toothache)=0.8$$

- Другая запись

$$P(a|b)=P(a \wedge b)/P(b)$$

$$P(a \wedge b)=P(a|b)P(b)$$

Существуют различные подходы – фреквентистский, объективистский, субъективистский и т.д.

Логический вывод

- Полное совместное распределение:

Полное совместное распределение для мира *Toothache*, *Cavity*, *Catch*

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

- Правило маргинализации:

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z}} \mathbf{P}(\mathbf{Y}, \mathbf{z})$$

- Правило обусловливания:

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z}} \mathbf{P}(\mathbf{Y} | \mathbf{z}) P(\mathbf{z})$$

- Вывод:

$$\mathbf{P}(X | \mathbf{e}) = \alpha\, \mathbf{P}(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} \mathbf{P}(X, \mathbf{e}, \mathbf{y})$$

Правило Байеса

- Позволяет «менять местами» посылку и следствие, то есть использовать причинную (структурную информацию) вместо статистической.

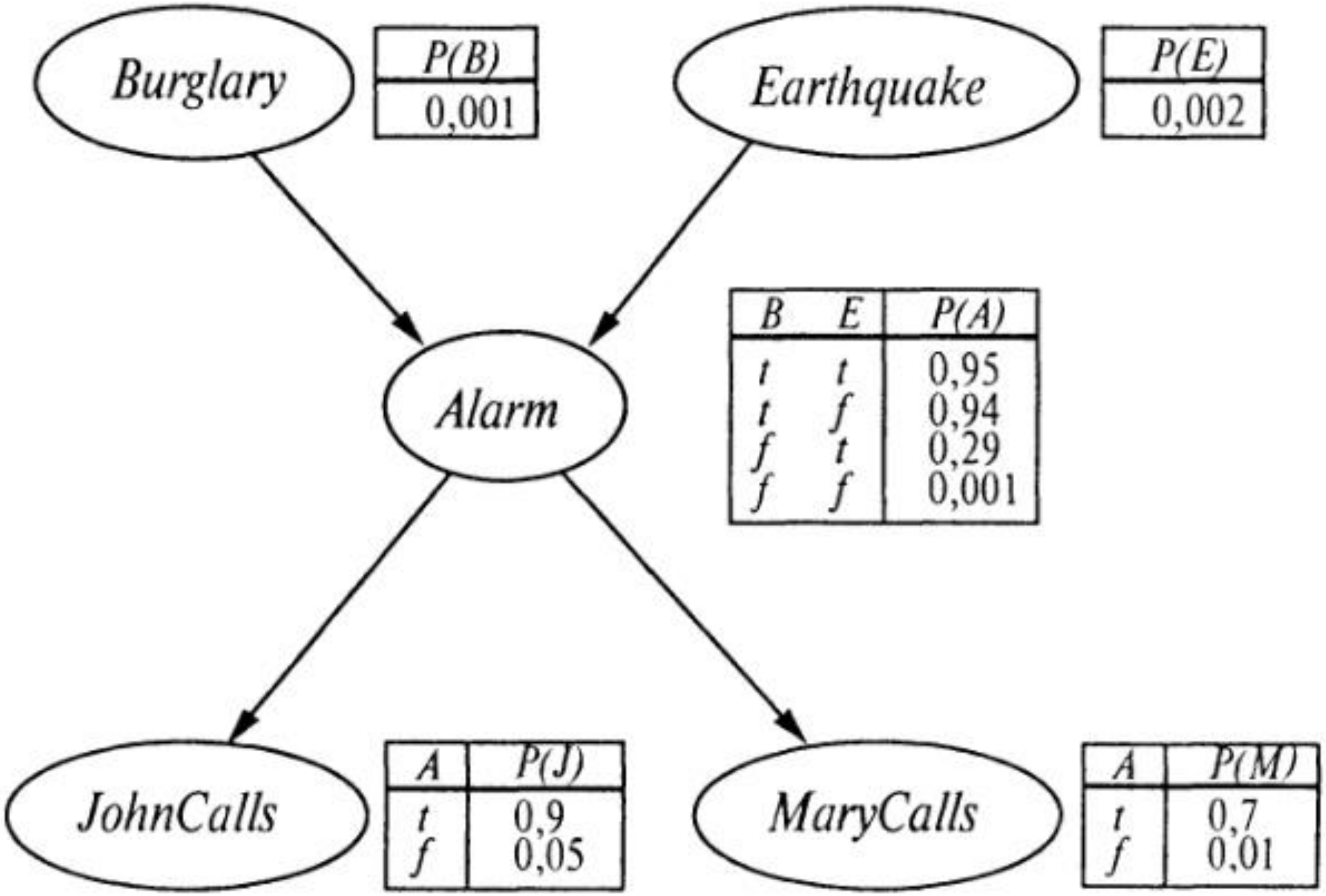
$$\mathbf{P}(Y \mid X) = \frac{\mathbf{P}(X \mid Y)\,\mathbf{P}(Y)}{\mathbf{P}(X)}$$

$$\mathbf{P}(Y \mid X, \mathbf{e}) = \frac{\mathbf{P}(X \mid Y, \mathbf{e})\,\mathbf{P}(Y \mid \mathbf{e})}{\mathbf{P}(X \mid \mathbf{e})}$$

Байсовские сети доверия

- Вершина – случайная переменная;
- Ребра – ориентированные, от X к Y означает, что X – родительская, Y – дочерняя;
- Вершина характеризуется распределением условных вероятностей, количественно оценивающих влияние родительских на эту вершину;
- Граф ациклический (Directed Acyclic Graph).
- Позволяют неявно задавать полные совместные распределения.

Пример сети



P(B)
0,001

P(E)
0,002

B	E	P(A)
t	t	0,95
t	f	0,94
f	t	0,29
f	f	0,001

A	P(J)
t	0,9
f	0,05

A	P(M)
t	0,7
f	0,01

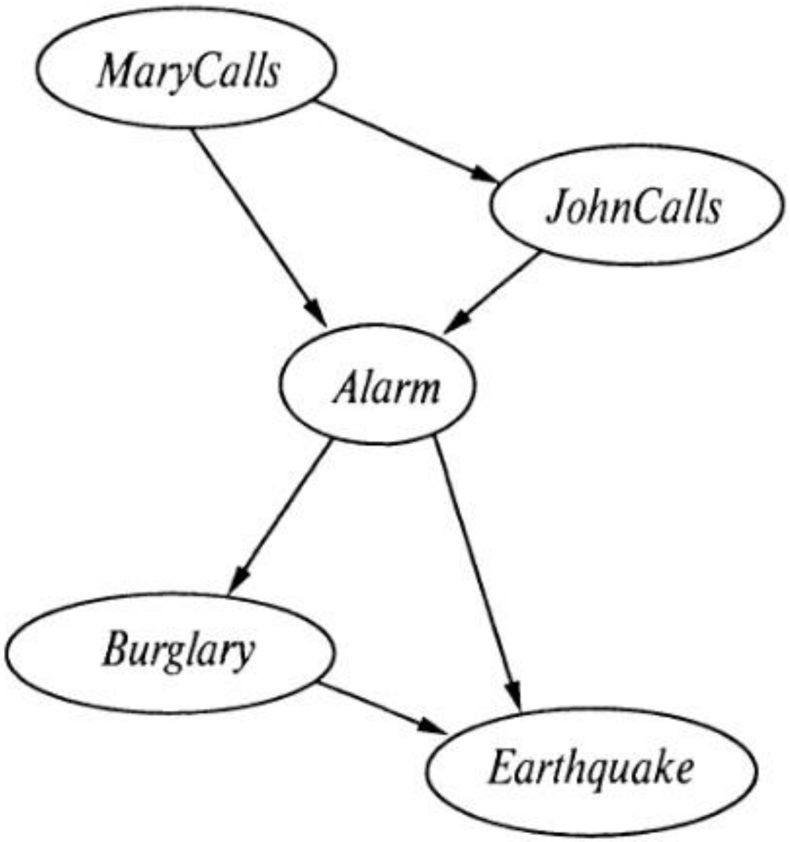
Основное уравнение

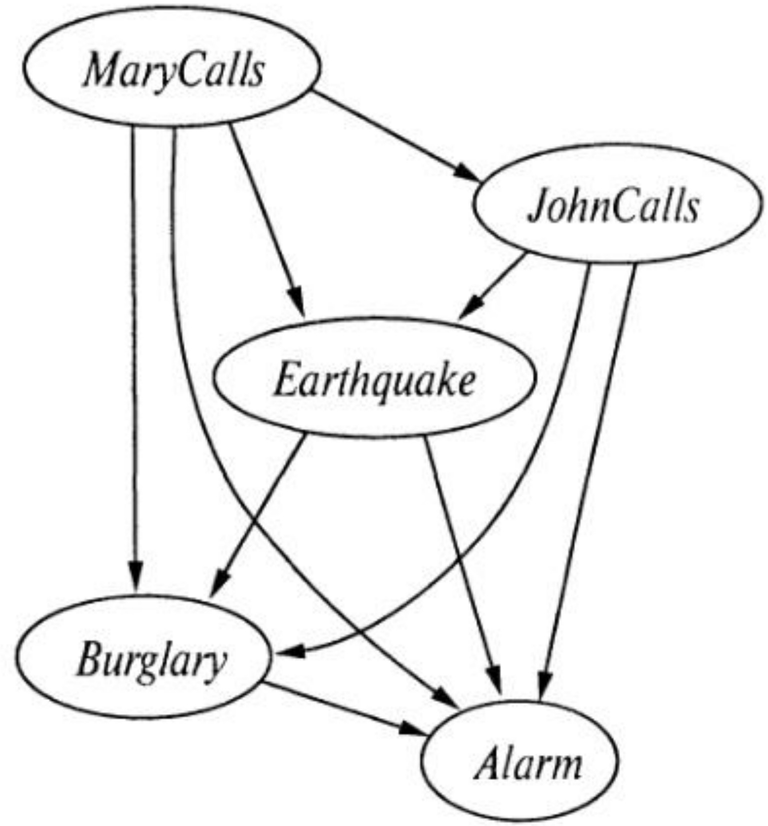
- Позволяет осуществлять вывод на сети, позволяет строить сети:

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid parents(X_i))$$

$$\begin{aligned} P(j \wedge m \wedge a \wedge \neg b \wedge \neg e) \\ &= P(j \mid a) P(m \mid a) P(a \mid \neg b \wedge \neg e) P(\neg b) P(\neg e) \\ &= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998 = 0.00062 \end{aligned}$$

Примеры некорректного построения сети



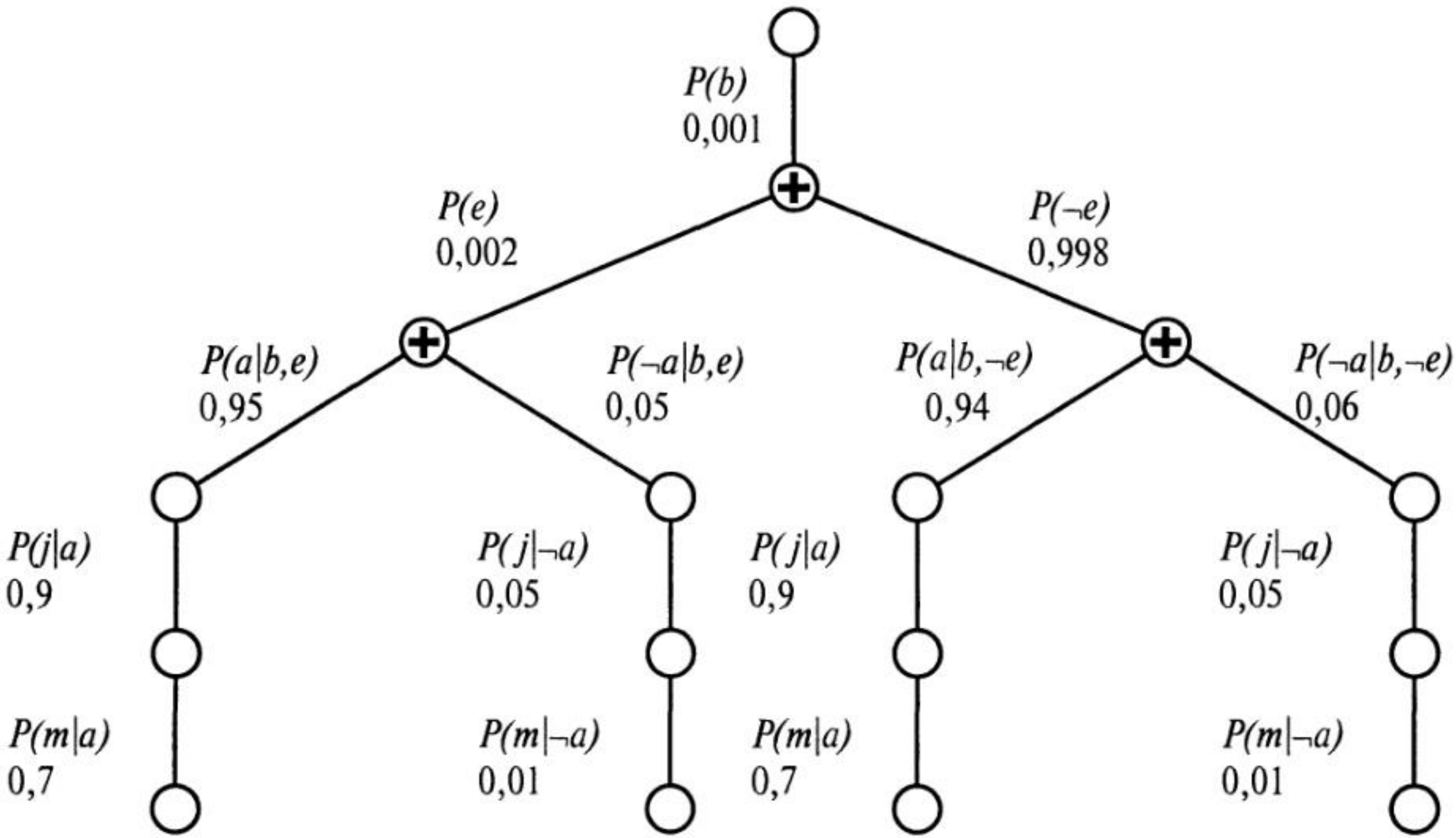


Точный вероятностный вывод

- Очень высокая сложность – соответствует полному перебору – хуже NP;
- Наблюдаем некоторые переменные (есть значения) – это эксперимент, E;
- Ненаблюдаемые переменные – скрытые, множество Y;
- Последовательно строим совместные присваивания и считаем:

$$\mathbf{P}(X|\mathbf{e}) = \alpha\,\mathbf{P}(X,\mathbf{e}) = \alpha \sum_{\mathbf{y}} \mathbf{P}(X,\mathbf{e},\mathbf{y})$$

Пример



P(b)
0,001
P(e)
0,002
P(¬e)
0,998
P(a|b,e)
0,95
P(¬a|b,e)
0,05
P(a|b,¬e)
0,94
P(¬a|b,¬e)
0,06
P(j|a)
0,9
P(j|¬a)
0,05
P(j|a)
0,9
P(j|¬a)
0,05
P(m|a)
0,7
P(m|¬a)
0,01
P(m|a)
0,7
P(m|¬a)
0,01

Абдуктивный вывод

- Свойство <u>монотонности</u> логических систем?
- Механизмы вывода – дедукция, индукция, абдукция (необоснованный метод вывода);
- Абдукция позволяет сформировать гипотезы для проверки и анализа. Аннулируемые модели, требующие механизмов поддержки истинности (логические системы поддержки истинности);

```
bird(X) unless penguin(X) → canFly(X);
bird(X) unless ab bird(X) → canFly(X);
```

Default logic

- Логика умолчаний – Реймонд Рейтер, 1980;
- `a(Z) and :b(Z) → c(Z)` – *«Если истинно a(Z) и не противоречит (согласуется) b(Z), то предполагаем истинность c(Z)»*
- Механизм вывода позволяет строить множество моделей, однако выбор между ними остаётся проблемным – имеем дело с множеством гипотез;

```
∀X good_student(X) ^ M study_hard(X) → graduates(X)
∀X party_person(X) ^ M ¬study_hard(X) → ¬graduates(X)
party_person(Peter)
study_hard(Peter)
```

«Замкнутость мира»

Замкнутость мира – свойство, при котором если формула А не может быть доказана, то считается верным not(A). Для использования минимальной модели требуется выполнение трёх предположений (аксиом):

- Уникальность имён – каждое имя обозначает уникальную сущность;
- Замкнутость мира – все отношения являются следствием имеющихся дизъюнктов;
- Замкнутость предметной области – все известные факты учтены, других нет.

Минимальные модели не требуют систем поддержки истинности.

Фактор уверенности

Стэнфордская теория фактора уверенности – используется certainty factor (Shortliffe and Buchanan, 1975). Использовалась в MYCIN;

$$MB(\mathrm{H},\mathrm{E}) = \begin{cases} 1, & \text{если } \mathrm{P(H)=1} \\ \dfrac{\max[\mathrm{P(H\,|\,E)},\mathrm{P(H)}] - \mathrm{P(H)}}{\max[1,0] - P(H)}, & \text{иначе} \end{cases}$$

$$MD(\mathrm{H},\mathrm{E}) = \begin{cases} 0, & \text{если } \mathrm{P(H)=0} \\ \dfrac{\min[\mathrm{P(H\,|\,E)},\mathrm{P(H)}] - \mathrm{P(H)}}{\min[1,0] - P(H)}, & \text{иначе} \end{cases}$$

Measure of belief: *MB*(H,E) – measure of belief в H при E

Measure of disbelief: *MD*(H,E) – в недостоверности H при E

Certainty factor: *CF*(H|E) = *MB*(H|E)-*MD*(H|E)

Вычисления

Правила вычислений:

and, or, not – элементарно

В 1977 в системе MYCIN изменили формулу вычисления CF:

$$CF = \frac{MB - MD}{1 - \min[MB, MD]}$$

Проверить для MB=0.999, MD=0.799

Комбинирующая функция:

$$CF_4 = \begin{cases} CF_1 + CF_2 - CF_1 CF_2 & CF_1, CF_2 \geq 0 \\ CF_1 + CF_2 + CF_1 CF_2 & CF_1, CF_2 < 0 \\ \frac{CF_1 + CF_2}{1 - \min(|CF_1|, |CF_2|)} & \text{otherwise} \end{cases}$$

Проблемы

1. Коэффициенты уверенности всё ещё выбираются на основе некоторой произвольной оценки – человеческий фактор;
2. В некоторых случаях можем получать не совсем корректные результаты:

$$P(H_1) = 0.8$$

$$P(H_1 \mid E) = 0.9$$

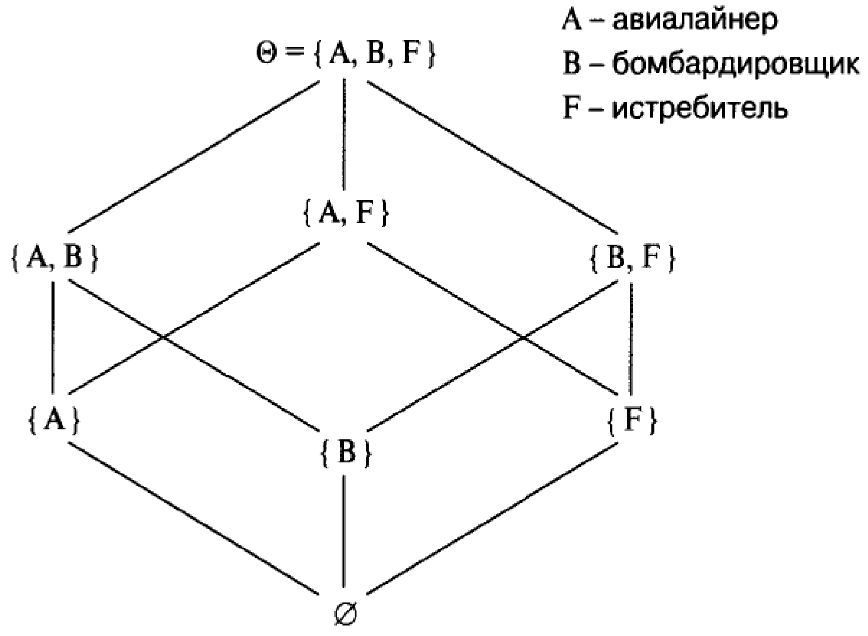
$$P(H_2) = 0.2$$

$$P(H_2 \mid E) = 0.8$$

3. «Веса» свидетельств не учитываются;
4. Остались проблемы теории вероятностей.

Теория Демпстера-Шафера

Рассматривается «среда», или универсум, и все возможные комбинации его элементов. Каждое подмножество – возможный ответ на некоторый вопрос. Если все элементы среды могут быть «ответами», и только один – правильный, то это *frame of discernment – рамки различения*.



Масса

Используется аналог физической массы, описывающей степень доверия к свидетельству – *мера доверия к свидетельству*. Позволяет нивелировать проблемы *принципа безразличия*, гласящего, что при отсутствии знаний все возможные исходы принимаем равновероятными. Если нет информации о свидетельствах, то масса (степень доверия) остаётся атрибутом среды – степень отсутствия доверия.

P(hostile) = 0.7, P(non-hostile) – ?

Сумма масс показательного множества нормирована 1, но подмножество может иметь массу большую, нежели надмножество.

Пьер-Симон Лаплас исследовал с помощью этого принципа вероятность восхода Солнца (задача Прайса).

Куб со стороной 2-4

Комбинирование масс

	$m_2(\{B\}) = 0.9$	$m_2(\Theta) = 0.1$
$m_1(\{B,F\}) = 0.7$	$\{B\}0.63$	$\{B,F\}0.07$
$m_1(\Theta) = 0.3$	$\{B\}0.27$	$\Theta 0.03$

- В простейшем случае получаем, что минимальное значение степени доверия к *B* равно 0.9, максимально правдоподобное – 1.0
- В общем случае – правило комбинирования Демпстера (прямая, или ортогональная сумма). Новые массы называются *консенсусом*:

$$m_1 \oplus m_2(Z) = \sum_{X \cap Y = Z} m_1(X) m_2(Y)$$

- Потерянные массы называются «конфликтом», и сумма масс подлежит нормированию на эту величину.

Доверие и правдоподобие

Доверие к гипотезе (Belief) = {сумма масс свидетельств, однозначно поддерживающих гипотезу}. От 0 до 1, 0 – отсутствие свидетельств в пользу гипотезы, 1 – все доступные свидетельства поддерживают гипотезу.

Правдоподобие (Plausibility) = 1 − {сумма масс всех свидетельств, противоречащих гипотезе}.

$$bel(A) = \sum_{B \subseteq A} m(B)$$

$$pls(A) = \sum_{B \cap A \neq \varnothing} m(B)$$

$$pls(A) = 1 - bel(\overline{A})$$

Методы на основе правил

- *Локальность*. Если из А следует В, то в логике не учитываем влияние других правил. В вероятностных системах это не так;
- *Отделение*. В логике если В доказано, то можно его использовать в доказательствах без учета источника. В вероятностных системах это не так;
- *Истинностная функциональность*. В логике значение выражения можно вычислить на основе значений переменных. При учете неопределенностей это не так (если нет сильных предположений о независимости).

Эти свойства не распространяются на вероятностные системы.

Проблемы

$P(A)$	$P(B)$	$P(A \vee B)$
	$P(H_1) = 0.5$	$P(H_1 \vee H_1) = 0.50$
$P(H_1) = 0.5$	$P(T_1) = 0.5$	$P(H_1 \vee T_1) = 1.00$
	$P(H_2) = 0.5$	$P(H_1 \vee H_2) = 0.75$

$$Rain \mapsto WetGrass \text{ и } WetGrass \mapsto Rain$$

$$Sprinkler \mapsto WetGrass \text{ и } WetGrass \mapsto Rain$$