Севастопольский государственный университет
Институт информационных технологий

Методы и системы искусственного интеллекта

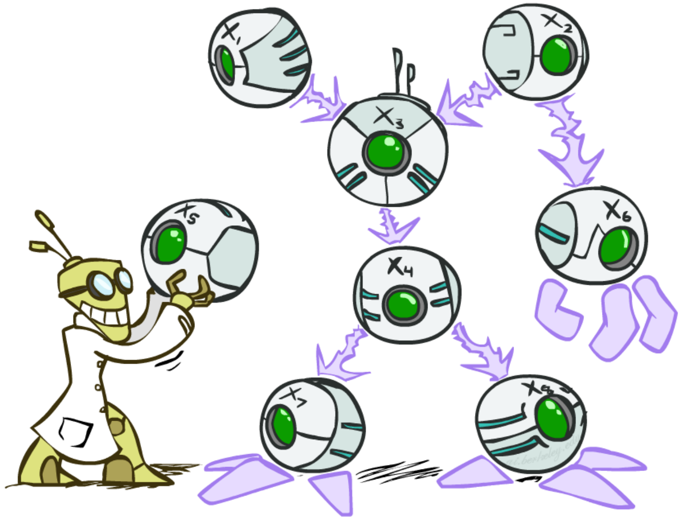
Бондарев Владимир Николаевич

Вероятностный вывод в сетях Байеса

Представление сетей Байеса

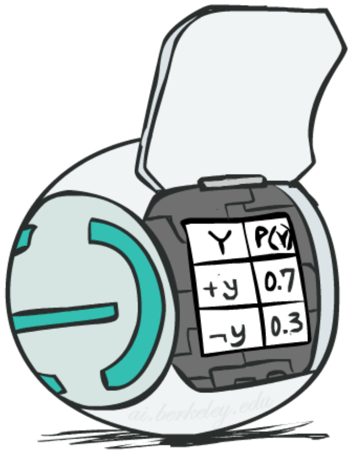
- Направленный ациклический граф, в котором вершина – случайная переменная
- С каждой вершиной связана таблица условных вероятностей (CPT- conditional probability table) :
  - распределение X для каждой комбинации значений родительских вершин

$$P(X|a_1 \ldots a_n)$$



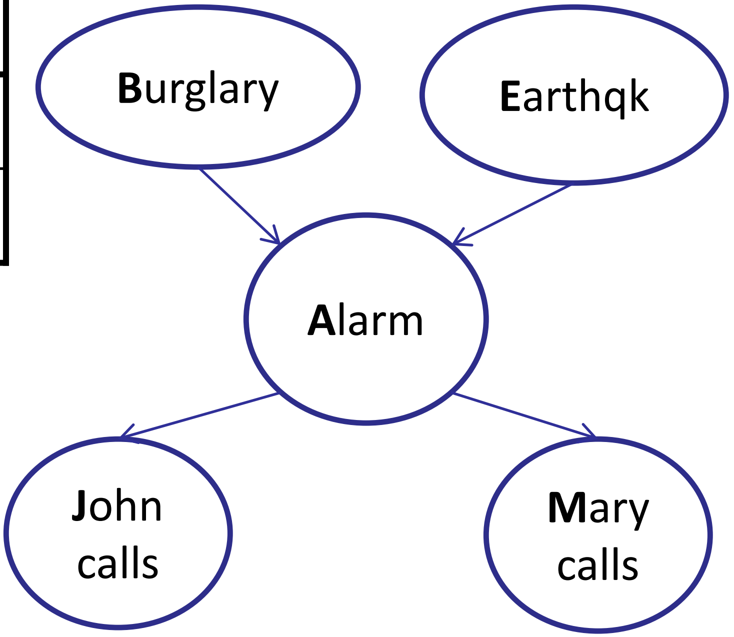
- Сети Байеса неявно кодируют совместное распределение:
  - в виде произведения локальных условных распределений;
  - чтобы определить, какая вероятность соответствует полному присваиванию в BN, перемножьте все соответствующие условные вероятности:

$$P(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} P(x_i|parents(X_i))$$

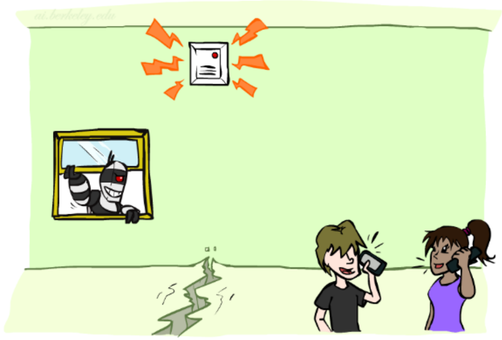


Пример: Сеть Тревоги

B	P(B)
+b	0.001
-b	0.999



E	P(E)
+e	0.002
-e	0.998



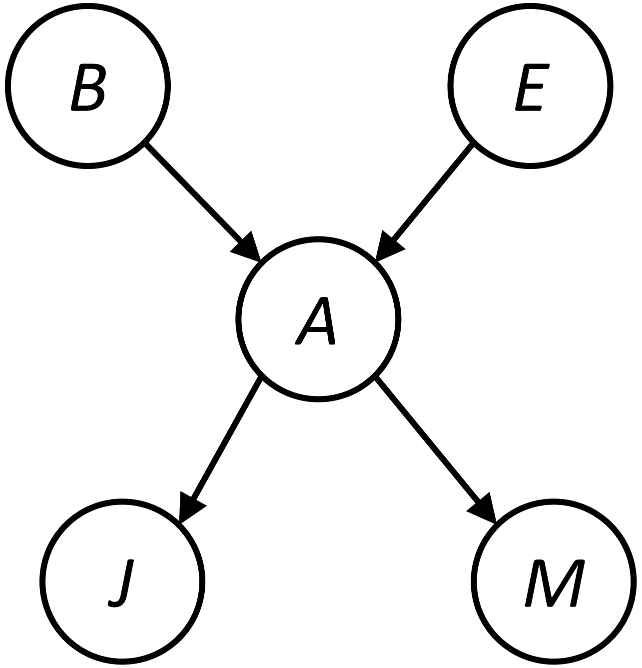
A	J	P(J|A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95

A	M	P(M|A)
+a	+m	0.7
+a	-m	0.3
-a	+m	0.01
-a	-m	0.99

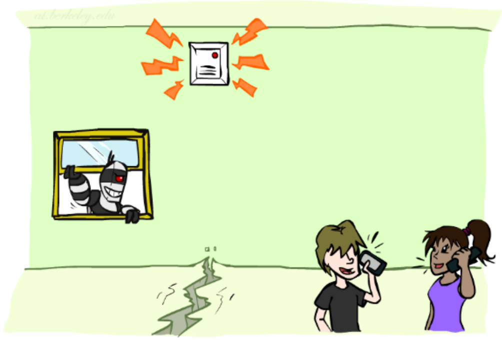
B	E	A	P(A|B,E)
+b	+e	+a	0.95
+b	+e	-a	0.05
+b	-e	+a	0.94
+b	-e	-a	0.06
-b	+e	+a	0.29
-b	+e	-a	0.71
-b	-e	+a	0.001
-b	-e	-a	0.999

Пример: Сеть Тревоги

B	P(B)
+b	0.001
-b	0.999



E	P(E)
+e	0.002
-e	0.998



A	J	P(J|A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95

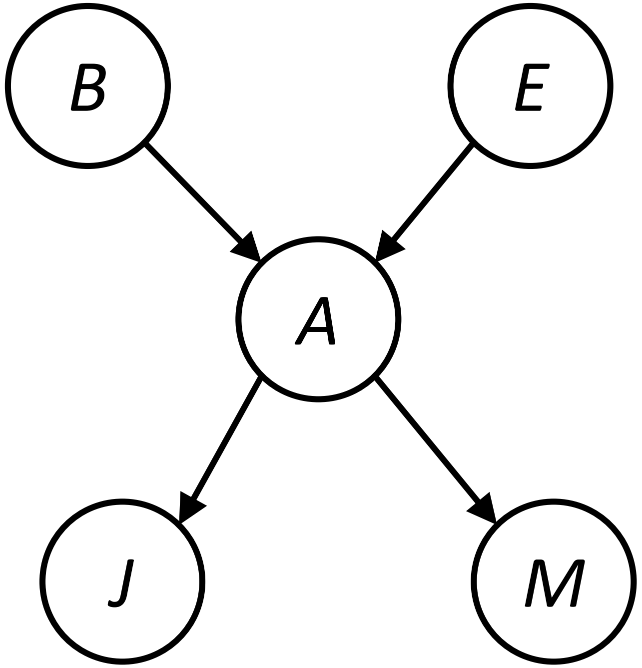
A	M	P(M|A)
+a	+m	0.7
+a	-m	0.3
-a	+m	0.01
-a	-m	0.99

B	E	A	P(A|B,E)
+b	+e	+a	0.95
+b	+e	-a	0.05
+b	-e	+a	0.94
+b	-e	-a	0.06
-b	+e	+a	0.29
-b	+e	-a	0.71
-b	-e	+a	0.001
-b	-e	-a	0.999

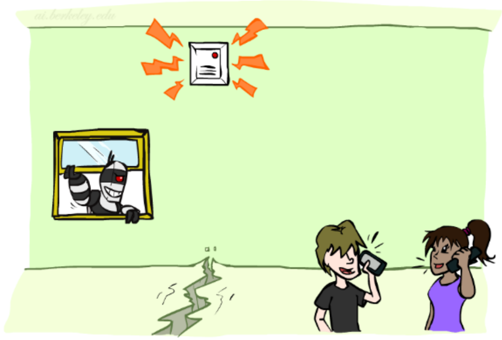
$$P(+b, -e, +a, -j, +m) =$$

Пример: Сеть Тревоги

B	P(B)
+b	0.001
-b	0.999



E	P(E)
+e	0.002
-e	0.998



A	J	P(J|A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95

A	M	P(M|A)
+a	+m	0.7
+a	-m	0.3
-a	+m	0.01
-a	-m	0.99

B	E	A	P(A|B,E)
+b	+e	+a	0.95
+b	+e	-a	0.05
+b	-e	+a	0.94
+b	-e	-a	0.06
-b	+e	+a	0.29
-b	+e	-a	0.71
-b	-e	+a	0.001
-b	-e	-a	0.999

$$P(+b,-e,+a,-j,+m) =$$
$$P(+b)P(-e)P(+a|+b,-e)P(-j|+a)P(+m|+a) =$$
$$0.001 \times 0.998 \times 0.94 \times 0.1 \times 0.7$$

Сети Байеса

- Представление
- Условная независимость
- Вероятностный вывод:
  - вывод простым перебором (точный, экспоненциальная сложность);
  - исключение переменных (точный, в наихудшем случае экспоненциальная сложность);
  - вывод относится к NP-полным задачам;
  - выборочный метод (аппроксимация);
- Обучение сетей Байеса на данных

Задачи вывода

- **Вывод** – вычисление вероятностей утверждений относительно переменных запроса на основе совместного распределения вероятностей

- Задачи вывода:
  - Вычисление апостериорного распределения запроса *Q*

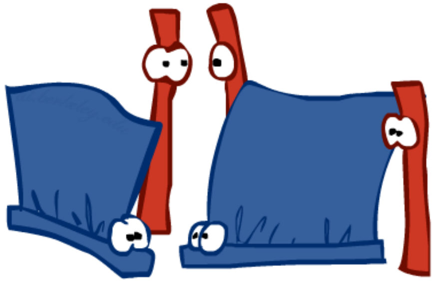
    $$P(Q|E_1 = e_1, \dots E_k = e_k)$$

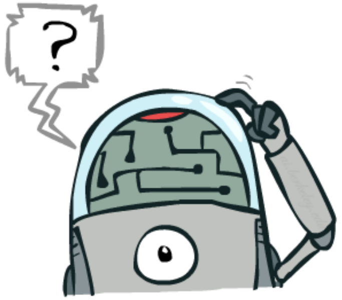
  - Наиболее вероятное объяснение:

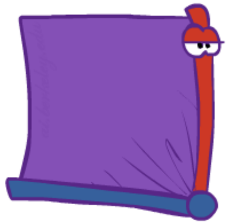
    $$\text{argmax}_q \; P(Q = q|E_1 = e_1 \dots)$$

  - Вычисление апостериорных распределений конъюнктивных запросов:

$$\mathbf{P}(X_i, X_j|\mathbf{E}=\mathbf{e}) = \mathbf{P}(X_i|\mathbf{E}=\mathbf{e})\mathbf{P}(X_j|X_i, \mathbf{E}=\mathbf{e})$$







Вывод путем перебора (enumeration) значений

- Дано:
  - Свидетельства: $E_1 \dots E_k = e_1 \dots e_k$
  - Переменная запроса*: Q
  - Скрытые переменные: $H_1 \dots H_r$

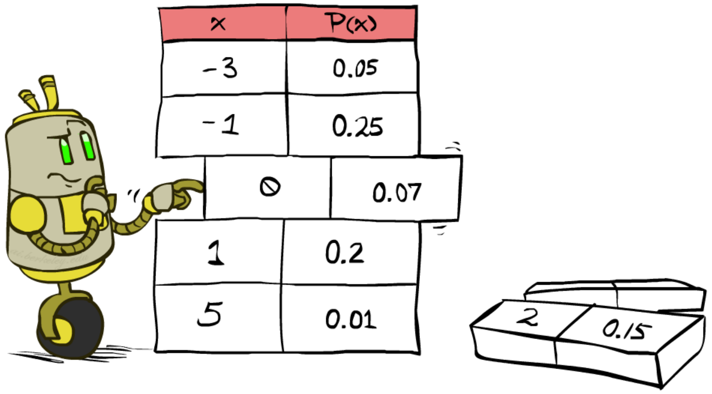
  $\left.\begin{matrix} E_1 \dots E_k = e_1 \dots e_k \\ Q \\ H_1 \dots H_r \end{matrix}\right\} X_1, X_2, \dots X_n$ *Все переменные*

- Требуется:

$$P(Q|e_1 \dots e_k)$$

** Также работает хорошо со мн-вом переменных запроса*

- Шаг 1: Выбрать входы CPT с учетом имеющихся свидетельств



- Шаг 2: Суммировать по H, чтобы получить совместную вероятность запроса и свидетельств

$$P(Q, e_1 \dots e_k) = \sum_{h_1 \dots h_r} \underbrace{P(Q, h_1 \dots h_r, e_1 \dots e_k)}_{X_1, X_2, \dots X_n}$$

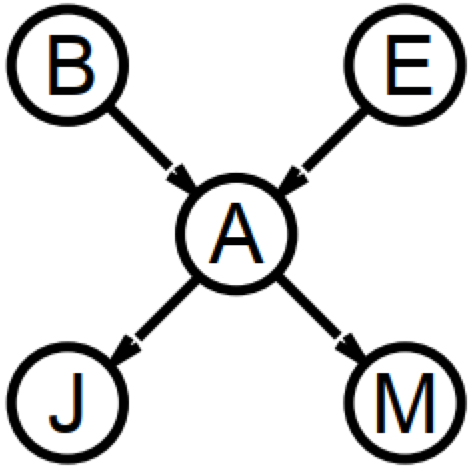
- Шаг3: Нормализация

$$Z = \sum_q P(Q, e_1 \cdots e_k)$$

$$P(Q|e_1 \cdots e_k) = \frac{1}{Z} P(Q, e_1 \cdots e_k)$$

Вывод путем перебора в сетях Байеса

Любую условную вероятность можно вычислить, суммируя элементы из полного совместного распределения:



$$P(B \mid +j, +m) \propto_B P(B, +j, +m) = \sum_{e,a} P(B, e, a, +j, +m)$$

Представим совместное распределение через произведение условных :

$$\sum_{e,a} P(B, e, a, +j, +m) = \sum_{e,a} P(B)P(e)P(a|B,e)P(+j|a)P(+m|a)$$

$$= P(B)P(+e)P(+a|B,+e)P(+j|+a)P(+m|+a) + P(B)P(+e)P(-a|B,+e)P(+j|-a)P(+m|-a)$$

$$+ P(B)P(-e)P(+a|B,-e)P(+j|+a)P(+m|+a) + P(B)P(-e)P(-a|B,-e)P(+j|-a)P(+m|-a)$$

Для вычисления этого выражения необходимо сложить четыре терма, каждый из которых вычисляется путем умножения пяти чисел. В наихудшем случае, когда приходится находить сумму почти по всем переменным, сложность этого алгоритма для сети с n булевыми переменными равна $O(n2^n)$.

Рекурсивный перебор в глубину: пространств. сложность – $O(n)$, временная – $O(d^n)$

Дерево оценивания

Для упрощения получим соответствующее выражение только для случая Burglary = true:

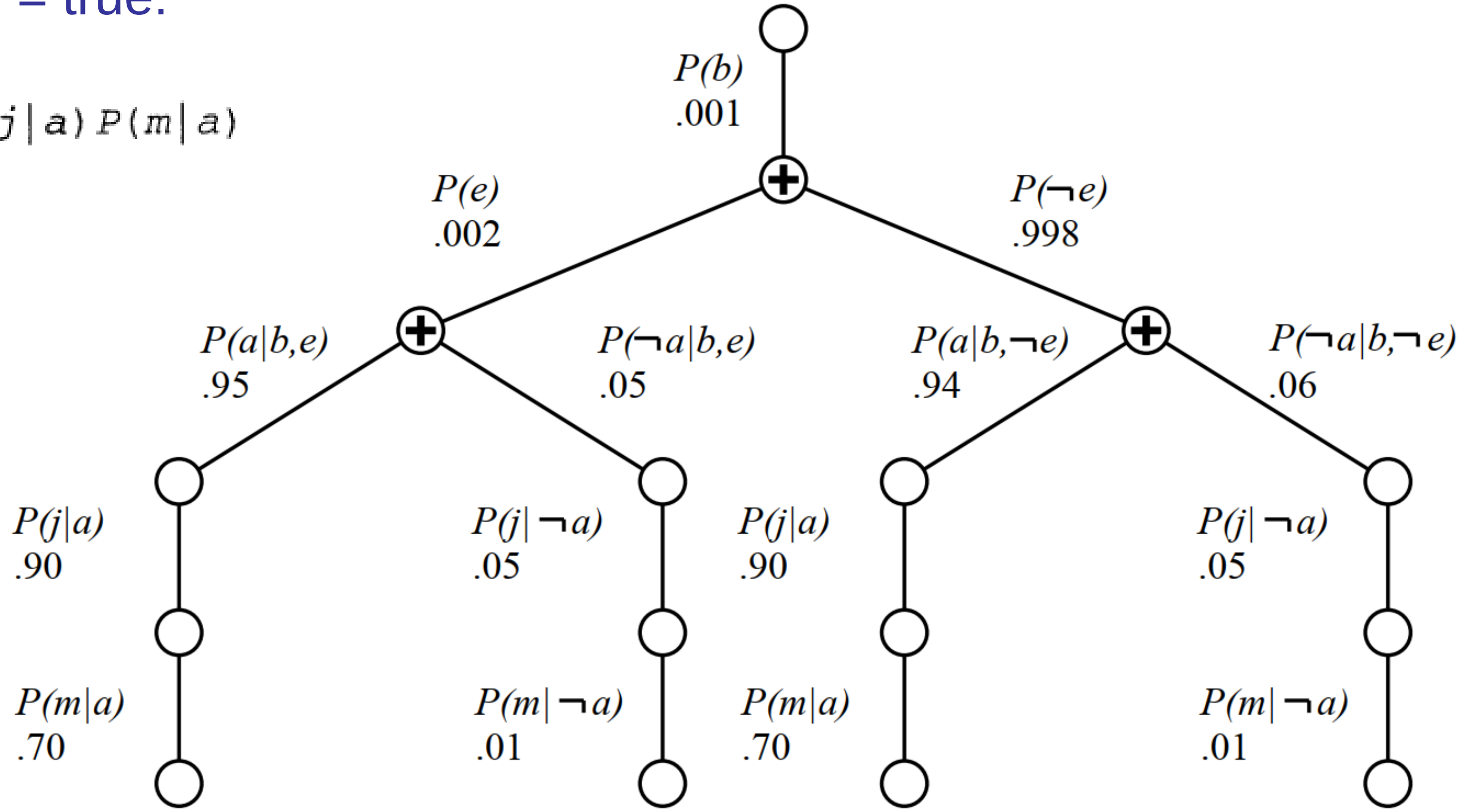
$$P(b|j,m) = \alpha\ P(b)\sum_{e} P(e)\sum_{a} P(a|b,e)P(j|a)P(m|a)$$

Из дерева оценивания видно, что вычисления неэффективны: произведения

P(j | a)P(m | a) и
P(j |-a)P(m |-a)

вычисляются дважды, по одному для каждого значения *e*.

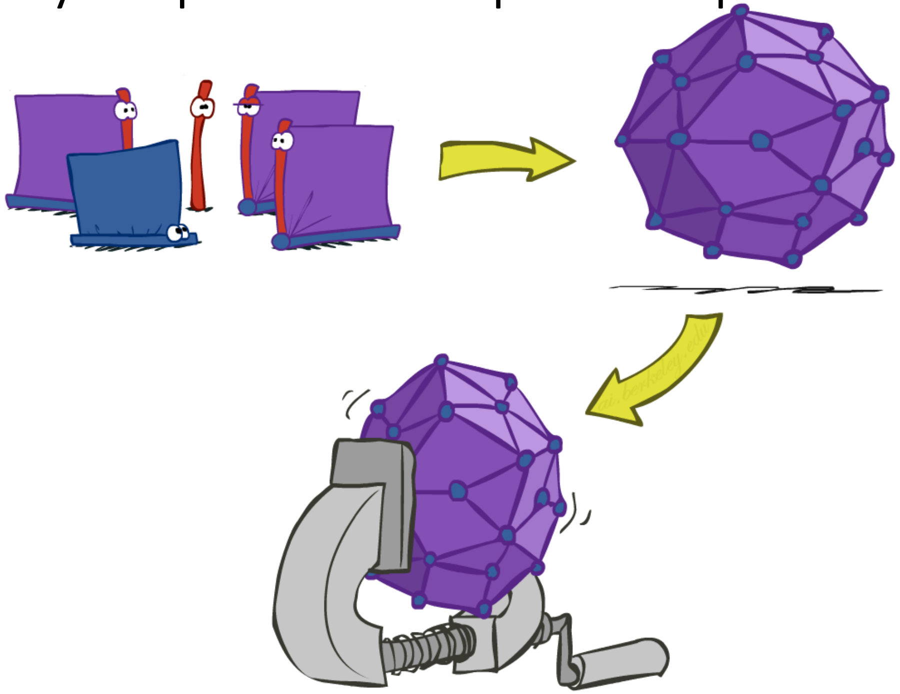
Далее будет рассмотрен **метод исключения переменных,** позволяющий избежать таких избыточных вычислений.



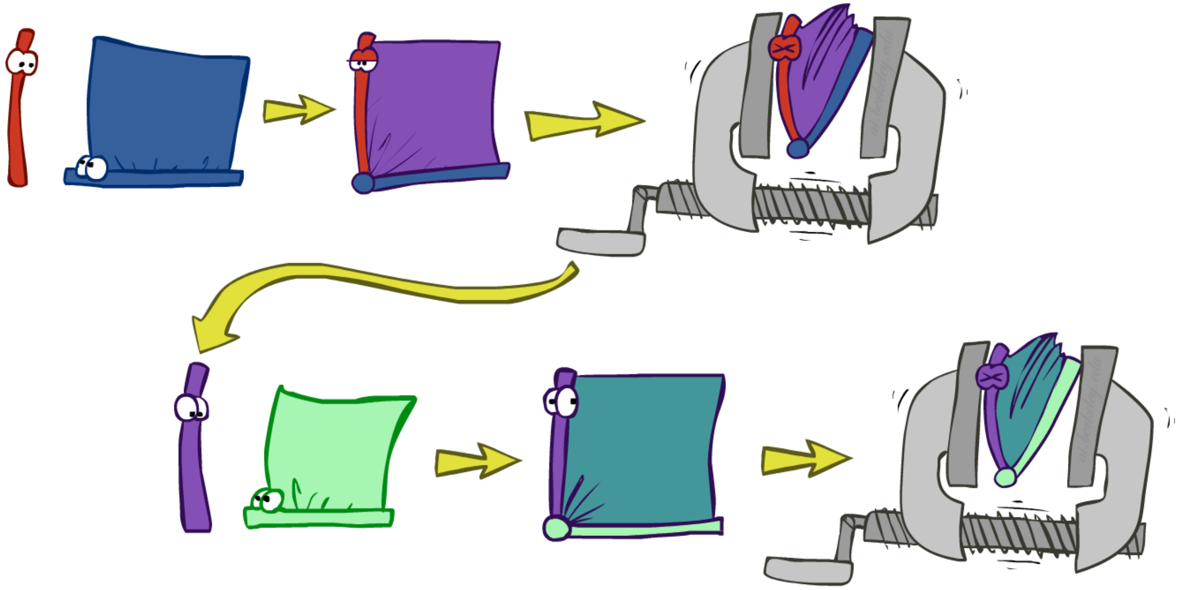
$$\mathbf{P}(B|j,m) = \alpha <0.00059224, 0.0014919> \approx <0.284, 0.716>$$

Исключение переменных (VE- variable elimination)

- Почему вывод перебором такой медленный?
  - Объединяем факторы в полное совместное распределение до выполнения суммирования по скрытым переменным

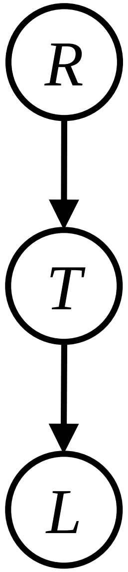


- Идея: разделить объединение и маргинализацию на этапы!
  - Метод называется "Исключение переменных"
  - Все еще NP-трудная задача, но обычно намного быстрее, чем вывод перебором



Пример: Траффик (вывод перебором)

- Случайные переменные
  - R: Дождь
  - T: Траффик
  - L: Опоздание!



$$P(L) = ?$$

$$= \sum_{r,t} P(r,t,L)$$

$$= \sum_{r,t} P(r)P(t|r)P(L|t)$$

$P(R)$

+r	0.1
-r	0.9

$P(T|R)$

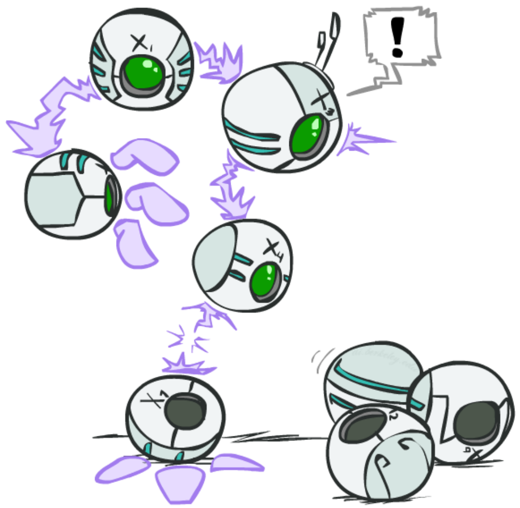
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Вывод перебором: процедура

- Выделить факторы совместного распределения;
- Определить начальные факторы - локальные CPT ;



$P(R)$

+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

- Использовать все известные значения переменных;
  - Например, если мы знаем $L = +\ell$, то начальные факторы можно упростить

$P(R)$

+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

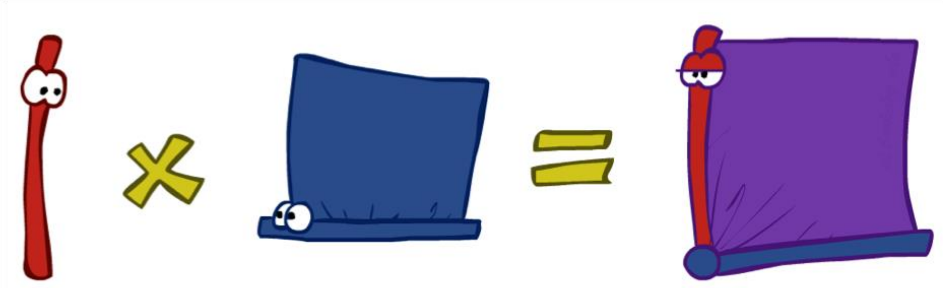
$P(+\ell|T)$

+t	+l	0.3
-t	+l	0.1

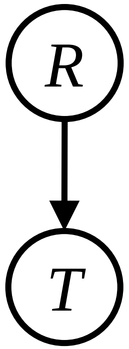
- **Процедура:** Найти произведение всех факторов, затем выполнить суммирование по всем скрытым переменным

Операция 1: Объединение факторов

- Первая базовая операция: объединение факторов
- Комбинирование факторов:
  - Взять все факторы, содержащие объединяемые переменные;
  - Построить новый фактор путём объединения вовлеченных переменных.



- Пример: Объединение по R



$$P(R) \times P(T|R) \rightarrow P(R,T)$$

$P(R)$:

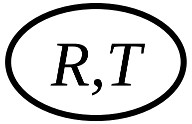
+r	0.1
-r	0.9

$P(T|R)$:

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

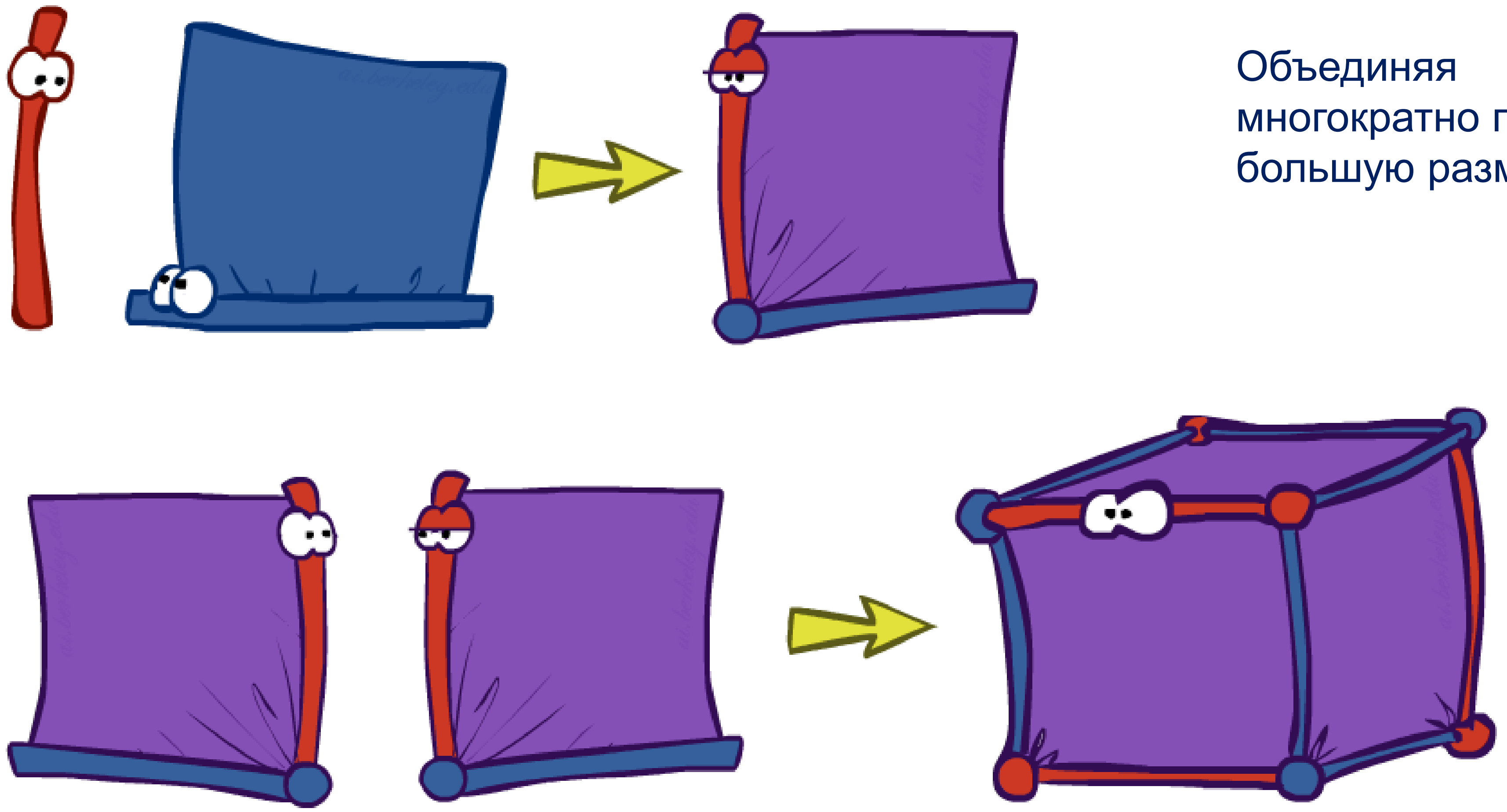
$P(R,T)$:

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81



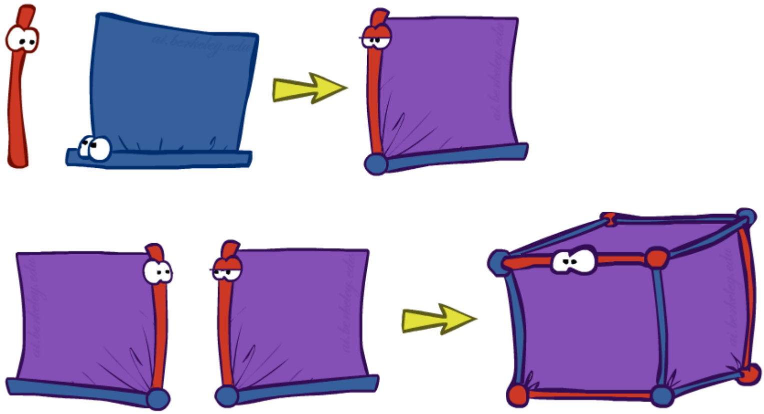
- Вычисление для всех входов точечного произведения: $\forall r, t : \quad P(r,t) = P(r) \cdot P(t|r)$

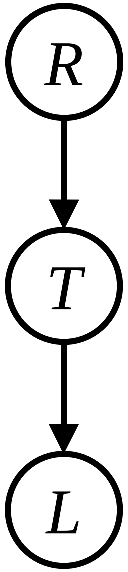
Пример: Многократные объединения



Объединяя многократно получаем большую размерность

Пример: Многократные объединения





$P(R)$

+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

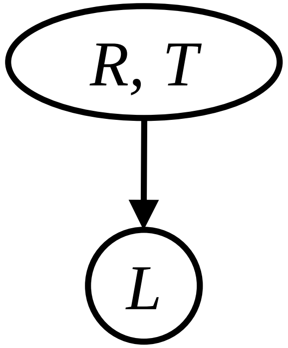
Объединение по R

$P(R,T)$

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9



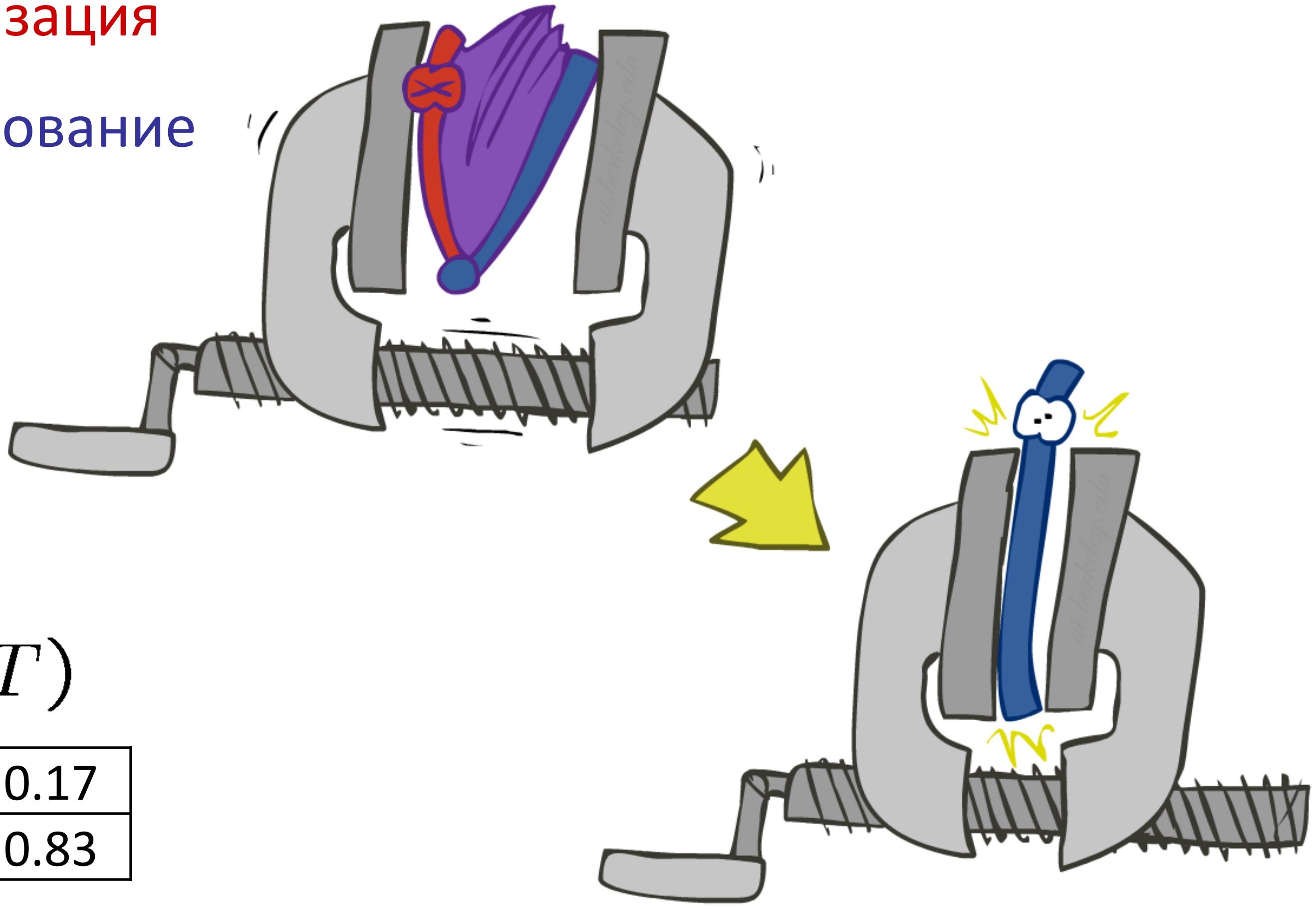
Объединение по T

$P(R,T,L)$

+r	+t	+l	0.024
+r	+t	-l	0.056
+r	-t	+l	0.002
+r	-t	-l	0.018
-r	+t	+l	0.027
-r	+t	-l	0.063
-r	-t	+l	0.081
-r	-t	-l	0.729

Операция 2: исключение путем суммирования

- Вторая базовая операция: маргинализация
- Выбрать фактор и выполнить суммирование по переменным:
  - сжатие фактора ;
  - операция проекции.
- Пример:



$$P(R,T)$$

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

sum R

$$P(T)$$

+t	0.17
-t	0.83

Многократное исключение

R, T, L

$P(R, T, L)$

+r	+t	+l	0.024
+r	+t	-l	0.056
+r	-t	+l	0.002
+r	-t	-l	0.018
-r	+t	+l	0.027
-r	+t	-l	0.063
-r	-t	+l	0.081
-r	-t	-l	0.729

Сумма по R →

T, L

$P(T, L)$

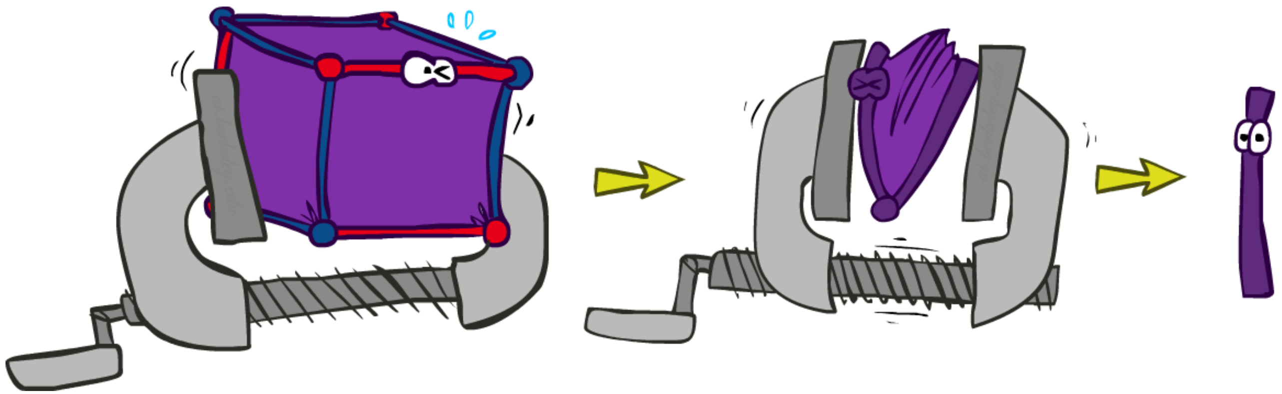
+t	+l	0.051
+t	-l	0.119
-t	+l	0.083
-t	-l	0.747

Сумма по T →

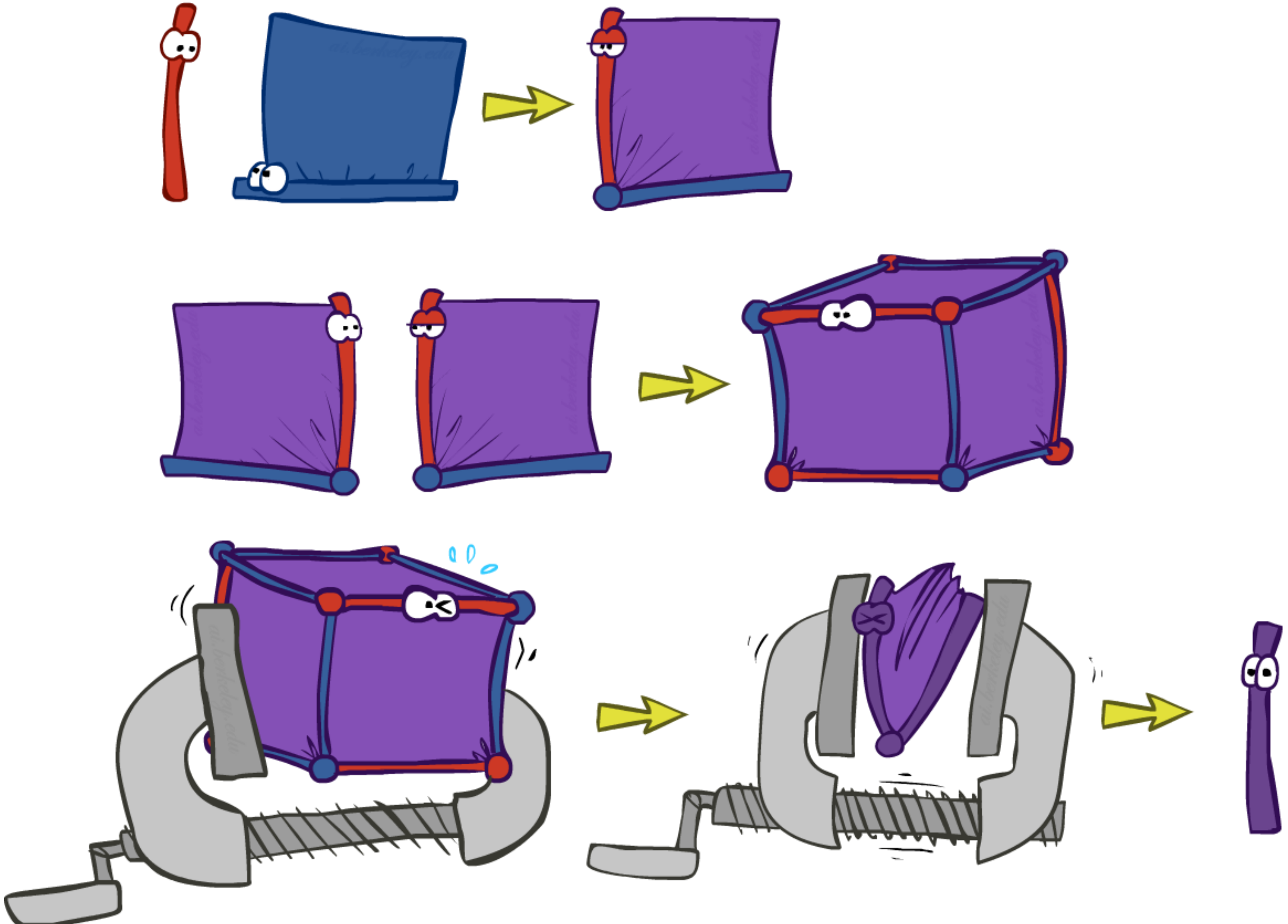
L

$P(L)$

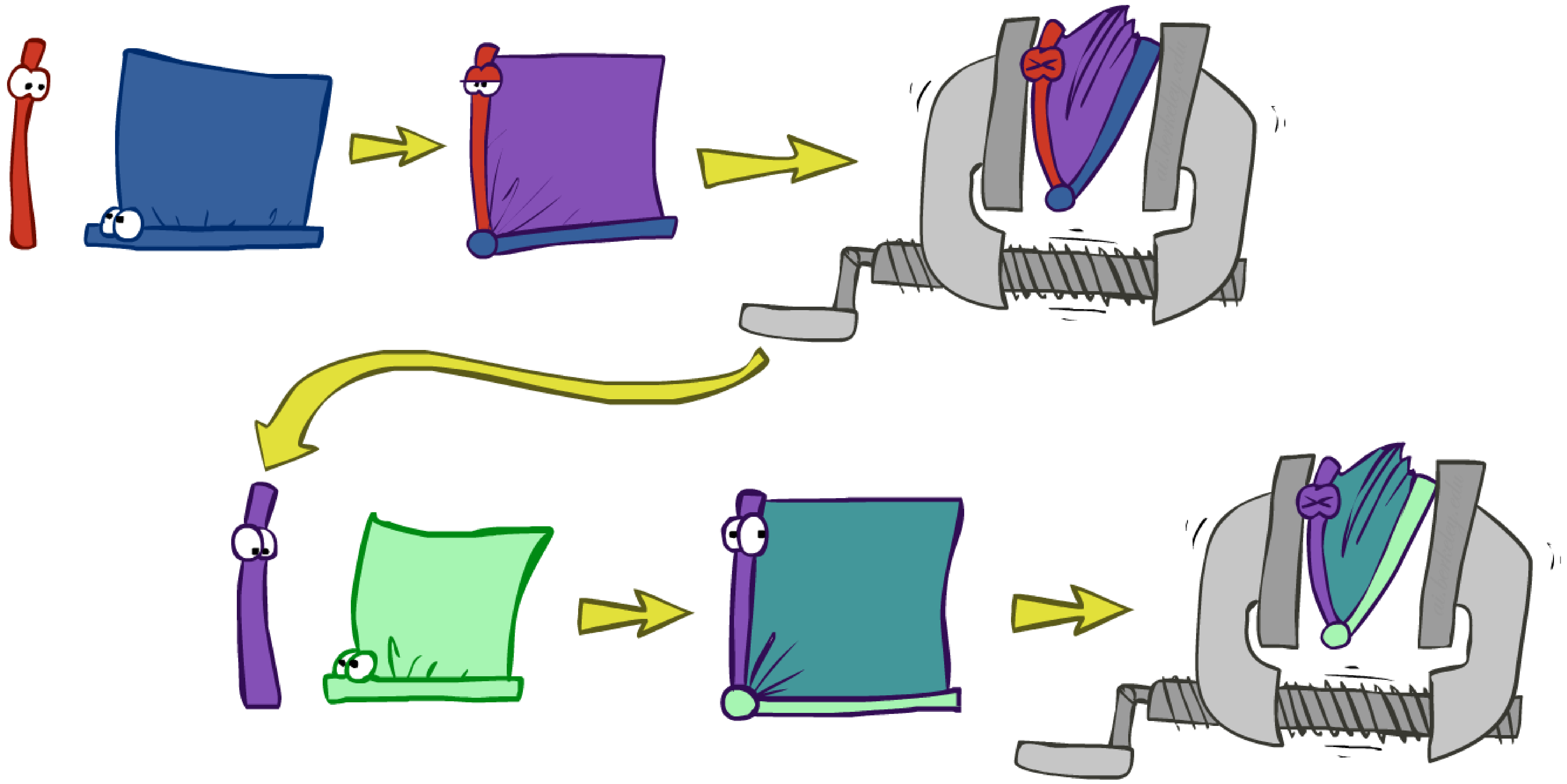
+l	0.134
-l	0.866



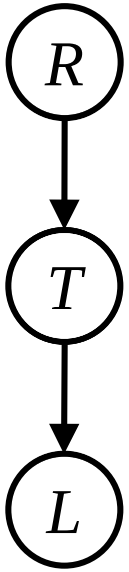
Итог: Вывод перебором = Многократное объединение + Многократное исключение



Ранняя маргинализация (= Исключение переменной)



Исключение переменных: пример Траффик



$$P(L) = ?$$

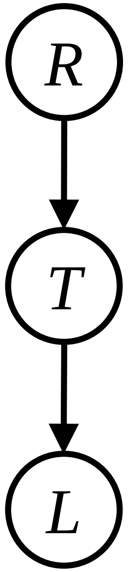
- Вывод перебором

$$=\underbrace{\sum_t \underbrace{\sum_r \underbrace{P(L|t)\underbrace{P(r)P(t|r)}_{\text{Объединение по r}}}_{\text{Объединение по t}}}_{\text{Исключение r}}}_{\text{Исключение t}}$$

- Исключение переменных

$$=\underbrace{\sum_t \underbrace{P(L|t) \underbrace{\sum_r \underbrace{P(r)P(t|r)}_{\text{Объединение r}}}_{\text{Исключение r}}}_{\text{Объединение t}}}_{\text{Исключение t}}$$

Ранняя маргинализация! (она же VE)



$P(R)$

+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

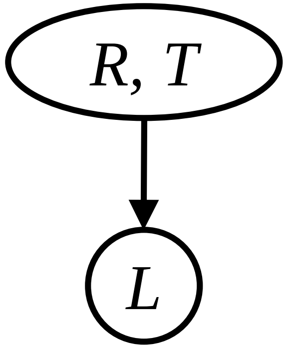
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Объед. по R

$P(R,T)$

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81



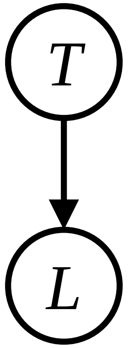
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Сумма по R

$P(T)$

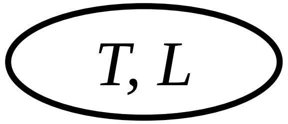
+t	0.17
-t	0.83



$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

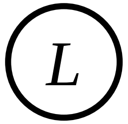
Объед. по T



$P(T,L)$

+t	+l	0.051
+t	-l	0.119
-t	+l	0.083
-t	-l	0.747

Сумма по T



$P(L)$

+l	0.134
-l	0.866

Исключение переменных: операции

1. **Суммирование по переменной** в выражении в виде произведения факторов: вынести любые постоянные множители за пределы суммирования, добавить подматрицу суммы в точечное произведение оставшихся множителей:

$$\Sigma_x f_1 \times \cdots \times f_k = f_1 \times \cdots \times f_i \, \Sigma_x \, f_{i+1} \times \cdots \times f_k = f_1 \times \cdots \times f_i \times f_{\bar{X}}$$

Предполагаем, что $f_1, \ldots, f_i$ не зависит от X.

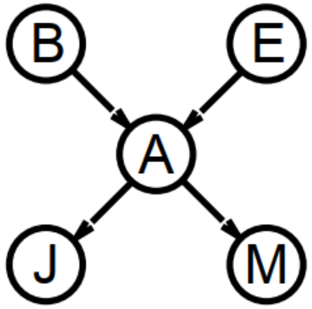
2. **Точечное произведение факторов** f_1 и f_2:

$$f_1(x_1, \ldots, x_j, y_1, \ldots, y_k) \times f_2(y_1, \ldots, y_k, z_1, \ldots, z_l) \\ = f(x_1, \ldots, x_j, y_1, \ldots, y_k, z_1, \ldots, z_l)$$

Например:

$$f_1(a, b) \times f_2(b, c) = f(a, b, c)$$

Исключение переменных: нерелевантные переменные



Рассмотрим запрос: $P(JohnCalls|Burglary = true)$

$$P(J|b) = \alpha P(b) \sum_e P(e) \sum_a P(a|b,e) P(J|a) \sum_m P(m|a)$$

Сумма по *m* даёт 1. *M* не имеет отношения (нерелевантная) к запросу.

Теорема: *Y* нерелевантна запросу, если она не принадлежит $Y \in Ancestors(\{X\} \cup \mathbf{E})$

Здесь $X = JohnCalls$ -- переменная запроса, $\mathbf{E} = \{Burglary\}$ -- свидетельство, множество предков:

$$Ancestors(\{X\} \cup \mathbf{E}) = \{Alarm, Earthquake\}$$

Следовательно $MaryCalls$ нерелевантная переменная.

К запросу не относится любая переменная, которая не является предком переменной запроса или переменной свидетельства. Поэтому алгоритм исключения переменных должен удалять все эти переменные, прежде чем приступит к вычислению ответа на запрос.

Учет свидетельств I

- Для учета свидетельств начните с факторов, которые содержат эти свидетельства
  - Если нет свидетельств, то начальные факторы соответствуют CPT:

$$P(R)$$

+r	0.1
-r	0.9

$$P(T|R)$$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$$P(L|T)$$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

  - При запросе $P(L|+r)$, начальные факторы станут проще:

$$P(+r)$$

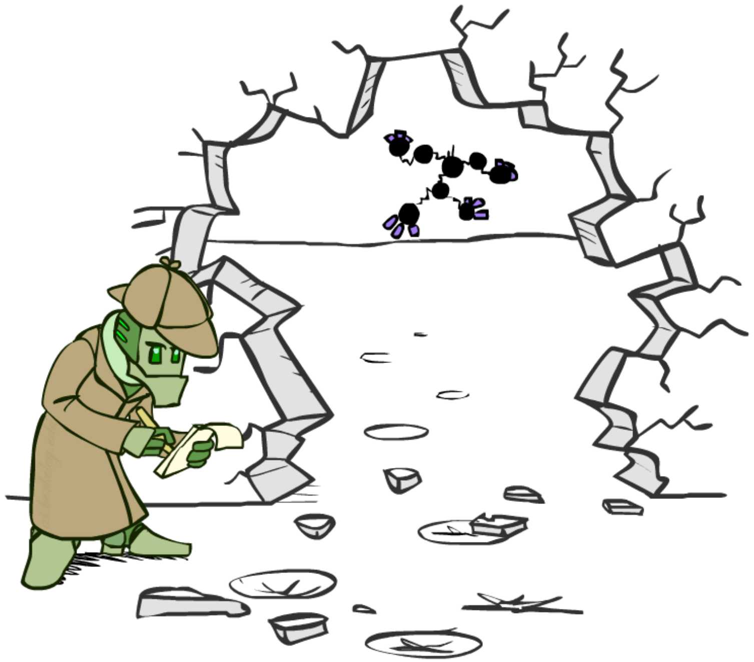
+r	0.1

$$P(T|+r)$$

+r	+t	0.8
+r	-t	0.2

$$P(L|T)$$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9



- В ходе вывода исключаются все переменные, кроме тех которые входят в множество {запрос + свидетельства}

Учет свидетельств II

- Результат получается объединением запроса и свидетельств
  - Например для P(L | +r) мы получим:

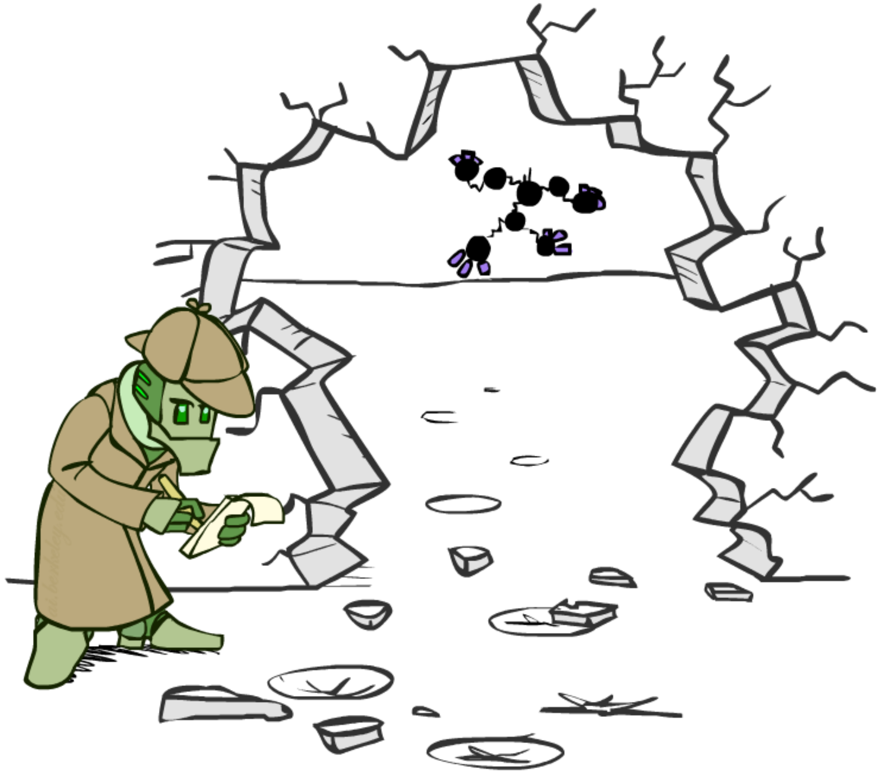
$$P(+r, L)$$

+r	+l	0.026
+r	-l	0.074

$\xrightarrow{\text{Нормализация}}$

$$P(L| + r)$$

+l	0.26
-l	0.74

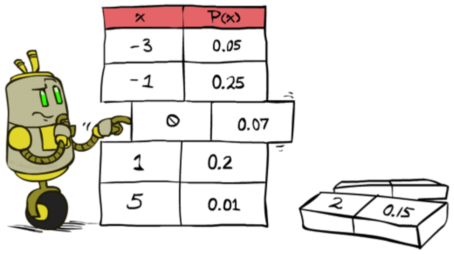


- Для получения ответа просто нормализуйте его!

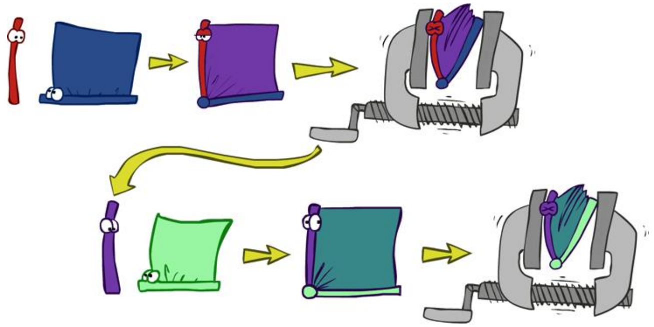
Исключение переменных: обобщение

- Запрос: $P(Q|E_1 = e_1, \dots E_k = e_k)$

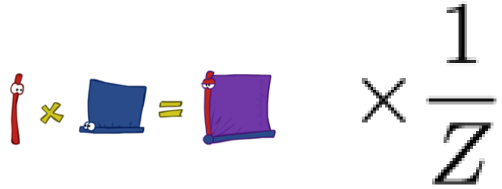
- Начните с начальных факторов:
  - Локальные CPT (без инициализации свидетельств)



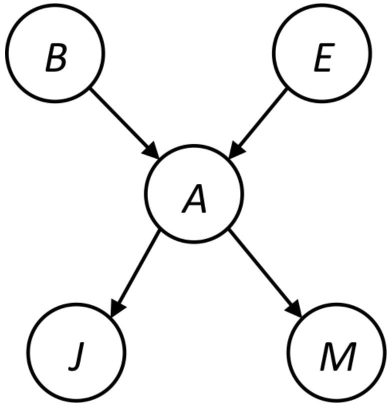
- Пока еще остаются скрытые переменные (не Q и не свидетельства):
  - Возьмите скрытую переменную H;
  - Объедините все факторы, в которых упоминается H;
  - Исключите H суммированием.



- Объедините все оставшиеся факторы и нормализуйте



Пример: сеть Тревоги (исключение переменных)



$$P(B|j,m) \propto P(B,j,m)$$

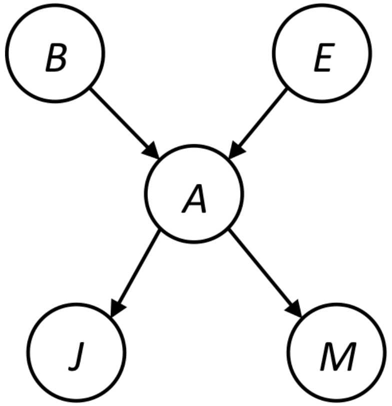
$$\boxed{P(B) \qquad P(E) \qquad P(A|B,E) \qquad P(j|A) \qquad P(m|A)}$$

Выберите A

$$\begin{array}{l} P(A|B,E) \\ P(j|A) \\ P(m|A) \end{array} \xrightarrow{\times} P(j,m,A|B,E) \xrightarrow{\sum} P(j,m|B,E)$$

$$\boxed{P(B) \qquad P(E) \qquad P(j,m|B,E)}$$

Пример: сеть Тревоги (исключение переменных)



$$\boxed{P(B) \qquad P(E) \qquad P(j,m|B,E)}$$

Выберите E

$$P(E) \quad P(j,m|B,E) \xrightarrow{\times} P(j,m,E|B) \xrightarrow{\sum} P(j,m|B)$$

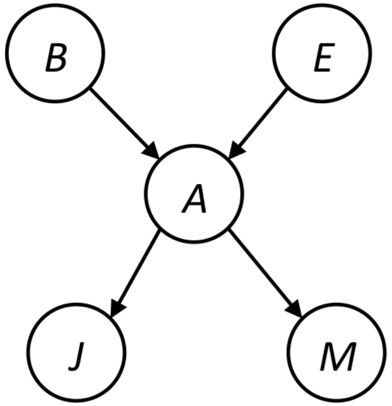
$$\boxed{P(B) \qquad P(j,m|B)}$$

Объединяем осставшиеся факторы и нормализуем:

$$P(B) \quad P(j,m|B) \xrightarrow{\times} P(j,m,B) \xrightarrow{\text{Нормализация}} P(B|j,m)$$

Пример: сеть Тревоги (VE с использованием выражений)

$$P(B|j,m) \propto P(B,j,m)$$



$P(B)$	$P(E)$	$P(A\|B,E)$	$P(j\|A)$	$P(m\|A)$

$$\begin{aligned} P(B|j,m) &\propto P(B,j,m) \\ &= \sum_{e,a} P(B,j,m,e,a) \\ &= \sum_{e,a} P(B)P(e)P(a|B,e)P(j|a)P(m|a) \\ &= \sum_{e} P(B)P(e)\sum_{a} P(a|B,e)P(j|a)P(m|a) \\ &= \sum_{e} P(B)P(e)f_1(j,m|B,e) \\ &= P(B)\sum_{e} P(e)f_1(j,m|B,e) \\ &= P(B)f_2(j,m|B) \end{aligned}$$

Маргинальное может быть получено по СР суммированием

Используйте выражение для совместного распределения сети Байеса

используйте x*(y+z) = xy + xz

Объединяя по a и затем суммируя, получаем f_1

Используйте x*(y+z) = xy + xz

Объединяя по e и затем суммируя, получаем f_2

Другой пример исключения переменных

Запрос: $P(X_3|Y_1 = y_1, Y_2 = y_2, Y_3 = y_3)$

Используя свидетельства, получим следующие начальные факторы

$$P(Z), P(X_1|Z), P(X_2|Z), P(X_3|Z), P(y_1|X_1), P(y_2|X_2), P(y_3|X_3)$$

Удаляем X_1 и вводим фактор $f_1(y_1|Z) = \sum_{x_1} P(x_1|Z)P(y_1|x_1)$, получаем

$$P(Z), P(X_2|Z), P(X_3|Z), P(y_2|X_2), P(y_3|X_3), f_1(y_1|Z)$$

Удаляем X_2 и вводим фактор $f_2(y_2|Z) = \sum_{x_2} P(x_2|Z)P(y_2|x_2)$, получаем

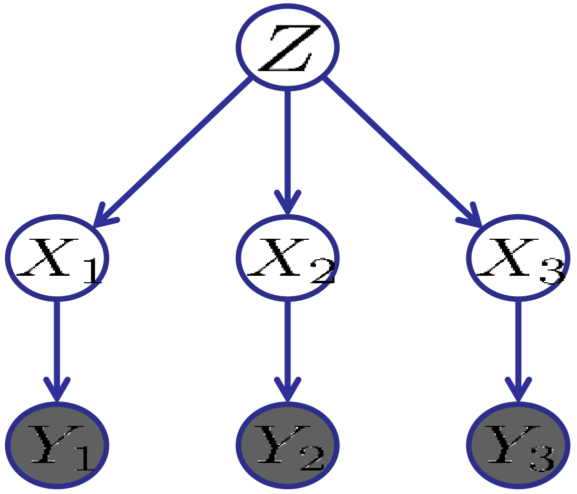
$$P(Z), P(X_3|Z), P(y_3|X_3), f_1(y_1|Z), f_2(y_2|Z)$$

Удаляем Z и вводим фактор $f_3(y_1, y_2, X_3) = \sum_z P(z)P(X_3|z)f_1(y_1|Z)f_2(y_2|Z)$, получаем

$$P(y_3|X_3), f_3(y_1, y_2, X_3)$$

Скрытых переменных нет. Объединяем оставшиеся факторы:

$$f_4(y_1, y_2, y_3, X_3) = P(y_3|X_3), f_3(y_1, y_2, X_3)$$

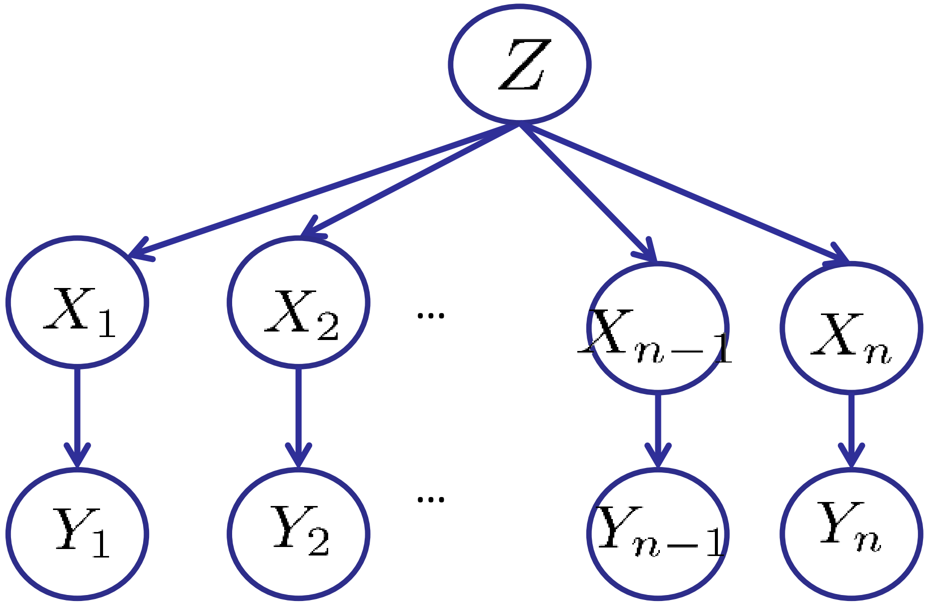


Временная (вычислительная) сложность зависит от наибольшего фактора. **Размер фактора = числу входов таблицы.** В примере (бинарные переменные) размер всех факторов (f1, f2, f3) равен 2 – так как все они зависят только от одной переменной (соответственно от Z и X_3).

Нормализация по X_3 даёт $P(X_3|y_1, y_2, y_3) = f_4(y_1, y_2, y_3, X_3) / \sum_{x_3} f_4(y_1, y_2, y_3, x_3)$

Упорядочение удаляемых переменных

- Для запроса $P(X_n|y_1,\dots,y_n)$ в соответствии с примером на рисунке возможны 2 варианта упорядочения (объединения):
- $X_1, \dots, X_{n-1}, Z$ и $Z, X_1, \dots, X_{n-1}$.
- Какой будет размер сгенерированного максимального фактора для каждого варианта упорядочения, где X_n - переменная запроса, $y_1, \dots, y_n$ – это свидетельства?



Мы хотим знать какое распределение, у X_n , т.к. X_n – это переменная запроса. Все остальные переменные – это скрытые переменные, которые будут исключаться.

Начнем с объединения по X1: $P(y_1 | X_1)*P(X_1 |Z) = P(y_1,X_1|Z)$. Затем суммируем по X_1 и получаем $P(y_1 |Z)$ и т.д.

Если начнем с объединения по Z, то $P(Z)$, $P(X_1 |Z),\dots, P(X_n|Z)$ дают распределение $P(Z,X_1, \dots, X_n)$ – получаем монстра.

- Ответ: 2^n против 2 (для бинарных переменных)
- В общем: упорядочение может существенно влиять на эффективность.

VE: Временная и пространственная сложность

- Временная и пространственная сложности метода исключения переменных определяются размером наибольшего фактора

- Порядок исключения оказывает существенное влияние на размер наибольшего фактора.
  - Например, сложность для предыдущего слайда 2^n против 2

- Всегда ли существует такое упорядочение, которое обеспечивает малый размер фактора?
  - Нет!

Сложность точного вероятностного вывода

Сеть Тревоги принадлежит к семейству сетей, в которых имеется самое большее один неориентированный путь между любыми двумя вершинами в сети. Такие сети называются **односвязными** (singly connected) сетями или **полидеревьями** (polytree), и обладают одним привлекательным свойством: временная и пространственная сложности точного вероятностного вывода в полидеревьях линейно зависят от размера сети: $O(d^k n)$, где *k* – размерность наибольшего фактора (кол-во родителей для каждого узла).

Для **полидеревьев** можно всегда найти упорядочение, которое будет эффективным. Для приведения многосвязных ((multiply connected)) сетей к полидереву можно использовать удаление подмножества переменных, образующих подмножество разрыва цикла (аналогично CSP задачам).

Для **многосвязных** сетей процедура исключения переменных в наихудшем случае может иметь **экспоненциальную временную и пространственную сложность**, даже если количество родительских вершин в расчете на каждую вершину ограниченно.

Сети Байеса

✓ Представление

✓ Условная независимость

✓ Вероятностный вывод:

- вывод простым перебором (точный, экспоненциальная сложность);
- исключение переменных (точный, в наихудшем случае экспоненциальная сложность);
- вывод относится к NP-трудным задачам;
- выборочный метод (аппроксимация);

- Обучение сетей Байеса на данных