УДК 004.8

А.В. Сироткин

ВЫЧИСЛИТЕЛЬНАЯ СЛОЖНОСТЬ АЛГОРИТМОВ ЛОКАЛЬНОГО АПОСТЕРИОРНОГО ВЫВОДА В АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЯХ

Сироткин А.В. **Вычислительная сложность алгоритмов локального апостериорного вывода в алгебраических байесовских сетях.**
Аннотация. В статье рассматривается локальный апостериорный вывод в алгебраических байесовских сетях. Для трёх видов свидетельств (детерминированного, стохастического и неточного) описывается способ проведения вывода и доказываются оценки сложности предлагаемых вычислений. В случае, когда вывод сводится к решению задач линейного программирования, оценка сложности даётся в виде числа таких задач, а также оценки числа переменных и ограничений в них. В остальных случаях сложность описывается числом арифметических операций.
Ключевые слова: алгебраическая байесовская сеть, апостериорный вывод, фрагмент знаний.

Sirotkin A.V. **Computational complexity of local posteriori inference algorithms in algebraic Bayesian networks.**
Abstract. The paper presents algorithmical complexity estimates for local posteriori inference in algebraic Bayesian networks. We consider the ways of implementing the inference for three types of evidence (deterministic, stochastic, and imprecise). If we need to solve linear programming tasks for inference, the comlexity estimations are given in numbers of such tasks and numbers of variables and constraints in each task. In other cases, complexity estimates are given in numbers of arithmetic operations.
Keywords: algebraic Bayesian network, posteriori inference, knowledge pattern.

1. Введение. Алгебраические байесовские сети (АБС) [2,5–10, 14,15,17], предложенные В.И. Городецким, являются одной из парадигм представления знаний с неопределённостью; эта парадигма основана на представлении данных в виде совместных распределений вероятности над небольшими наборами пропозициональных переменных. Достаточно глубокое и подробное описание алгебраических байесовских сетей можно найти в монографиях [6,14,17]. В данной статье мы остановимся подробно на локальном апостериорном выводе. Его общее описание можно найти в [9,13], однако в статье будет дано достаточно подробное описание возникающих задач и способов их решения с конечной целью предложить и доказать набор оценок сложности для задач, возникающих в апостериорном локальном логико-вероятностном выводе в АБС.

Мы заимствуем базовые обозначения из статьи [4], публикуемой в этом же выпуске. Кроме того, в [4] даются базовые определения, такие как фрагмент знаний, квант, конъюнкт. В дальнейшем тексте мы предполагаем, что читатель знаком с этими понятиями по другим публикациям.

В целом система подходов, доказуемых фактов, результатов, терминов и обозначений в настоящей работе основывается на [3, 6, 11, 12, 14, 17].

2. Локальный апостериорный вывод и свидетельства. Для описания логико-вероятностной модели свидетельств в теории АБС нам потребуется ряд вспомогательных обозначений.

Определение 1. *Под свидетельством мы понимаем новые «обусловливающие» данные, которые поступили во фрагмент знаний, и с учётом которых нам требуется пересмотреть все (или некоторые) оценки. Для обозначения свидетельства будут использоваться угловые скобки — $\langle\ldots\rangle$.*

Определение 2. *Первая задача апостериорного вывода заключается в том, чтобы оценить вероятность истинности свидетельства при уже заданных оценках вероятности истинности элементов фрагмента знаний.*

Определение 3. *Вторая задача апостериорного вывода заключается в том, чтобы оценить условные вероятности истинности элементов фрагмента знаний при предположении, что имеет место быть свидетельство.*

В теории АБС рассматриваются три вида свидетельств:

- детерминированные свидетельства;
- стохастические свидетельства;
- неточные свидетельства.

Если задан фрагмент знаний с носителем C, тогда в рамках локального апостериорного вывода действуют следующие (4–6) определения.

Определение 4. *Детерминированное свидетельство — это предположение, что один или несколько атомов получили* *конкретные означивания. Данное свидетельство обозначается* $\langle c_i, c_j \rangle$. *Здесь* c_i *и* c_j *— конъюнкты, состоящие из атомов, получивших положительные и отрицательные означивания соответстенно. Для сокращения записи может также использоваться обозначение* $\langle i, j \rangle$, *где* i *и* j *— индексы соответствующих конъюнктов.*

Определение 5. *Стохастическое свидетельство — это предположение о том, что над* C' *— подыдеале* C *— задан непротиворечивый фрагмент знаний со скалярными оценками, который определяет вероятности истинности элементов соответствующего подыдеала. Данное свидетельство обозначается* $\langle (C', \mathbf{P_c}) \rangle$.

Определение 6. *Неточное свидетельство — это предположение о том, что над* C' *— подыдеале* C, *задан непротиворечивый фрагмент знаний с интервальными оценками, который определяет вероятности истинности элементов соответствующего подыдеала. Данное свидетельство обозначается* $\langle (C', \mathbf{P_c^-}, \mathbf{P_c^+}) \rangle$.

В первую очередь рассмотрим процесс обработки детерминированного свидетельства, поступающего во фрагмент знаний со скалярными оценками $(C, \mathbf{P_c})$. Пусть поступило свидетельство $\langle i, j \rangle$.

Определение 7. *Мы говорим, что квант* q_m *согласован со свидетельством* $\langle c_i, c_j \rangle$, *если все атомы конъюнкта* c_i *положительно означены в* q_m, *а все атомы конъюнкта* c_j *— означены отрицательно.*

Замечание 1. *В рамках индексации квантов и конъюнктов, описанной в [4], можно формализовать определение 7 следующим образом* $(m \dot{\&} i = i) \& (\dot{\neg} m \dot{\&} j = j)$, *здесь и далее мы будем использовать символ логической операции с точкой над ней для обозначения соответствующей побитовой операции. В частности,* $\dot{\&}$ *используется для операции побитового «и», а* $\dot{\neg}$ *— для операции побитового отрицания.*

Определение 8. *Введём вспомогательные обозначения:*

$$\mathbf{H}^+ = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}; \quad \mathbf{H}^- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad \mathbf{H}^\circ = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Определение 9. *Определим матрицу* $\mathsf{H}^{\langle i,j\rangle}$ *как следующее тензорное произведение:*

$$\mathsf{H}^{\langle i,j\rangle} = \widetilde{\mathsf{H}}^{\langle i,j\rangle}_{n-1} \otimes \widetilde{\mathsf{H}}^{\langle i,j\rangle}_{n-2} \otimes \ldots \widetilde{\mathsf{H}}^{\langle i,j\rangle}_{0},$$

где

$$\widetilde{\mathsf{H}}^{\langle i,j\rangle}_{k} = \begin{cases} \mathsf{H}^{+}, & \textit{если } x_k \textit{ входит в } c_i; \\ \mathsf{H}^{-}, & \textit{если } x_k \textit{ входит в } c_j; \\ \mathsf{H}^{\circ}, & \textit{иначе.} \end{cases}$$

Лемма 1. *Матрица* $\mathsf{H}^{\langle i,j\rangle}$ *состоит из нулей за исключением некоторого числа единиц на главной диагонали. Диагональные элементы могут быть заданы следующей формулой:*

$$\mathsf{H}^{\langle i,j\rangle}[m,m] = \begin{cases} 1, & \textit{если } (m \dot{\&} i = i) \textit{ и } (\dot{\neg} m \dot{\&} j = j), \\ 0, & \textit{иначе.} \end{cases} \tag{1}$$

Доказательство. По определению матрица $\mathsf{H}^{\langle i,j\rangle}$ представляет собой тензорное произведение матриц $\widetilde{\mathsf{H}}^{\langle i,j\rangle}_{k}$. Любой элемент матрицы $\mathsf{H}^{\langle i,j\rangle}$ является произведением элементов из матриц $\widetilde{\mathsf{H}}^{\langle i,j\rangle}_{k}$. Если этот элемент не лежит на главной диагонали, то по крайней мере один из множителей элементов матриц $\widetilde{\mathsf{H}}^{\langle i,j\rangle}_{k}$ тоже не будет лежать на главной диагонали, а следовательно будет равен нулю, что означает, что и все произведение будет равно нулю. Рассмотрим элемент $\mathsf{H}^{\langle i,j\rangle}[m,m]$. Рассмотрим k-тый разряд числа m. Если это единица, то это соответствует нижнему элементу главной диагонали матрицы $\widetilde{\mathsf{H}}^{\langle i,j\rangle}_{k}$. Соответственно, если элемент не равен нулю, то это или матрица H^{+}, или H°. Аналогично, если k-тый разряд числа m равен нулю, а элемент равен единице, то k-тая матрица или H^{-}, или H°. Рассмотрим индекс положительного конюънкта i. Если k-тый разряд числа i единица, то k-тая матрица H^{+} и единица может быть только, если k-тый разряд числа m равен единице. Аналогично, если k-тый разряд числа j единица, то k-тая матрица H^{-} и единица может быть только, если k-тый разряд числа m равен нулю. А именно эти два утверждения и соответствуют формулам $(m \dot{\&} i = i)$ и $(\dot{\neg} m \dot{\&} j = j)$.

$\square$

Замечание 2. *По лемме 1 все ненулевые элементы лежат на главной диагонали, а следовательно* $\mathbf{H}^{\langle i;j\rangle^{T}} = \mathbf{H}^{\langle i;j\rangle}$.

Лемма 2. *Матрица* $\mathbf{H}^{\langle i,j\rangle}$ *действует на вектор* $\mathbf{P_q}$ *следующим образом: вероятности квантов, которые согласованы со свидетельством* $\langle i,j\rangle$*, остаются неизменными, а вероятности всех остальных квантов приравниваются к нулю.*

Доказательство. Согласно лемме 1 матрица $\mathbf{H}^{\langle i,j\rangle}$ содержит в m-той строке единицу, только если $(m\dot{\&}i = i)$ и $(\dot{\neg}m\&j = j)$, но согласно замечанию 1, это значит, что квант q_m согласован со свидетельством $\langle i,j\rangle$. Так как единица находится на диагонали, а все остальные элементы нули, то при умножении на матрицу m-тый элемент полученного вектора будет совпадать с m-тым элементом $\mathbf{P_q}$, то есть с $p(q_m)$. Если m не удовлетворяет указанным выше условиям, то это значит, с одной стороны, что квант q_m, не согласован с нашим свидетельством, а с другой, что все элементы m-той строки матрицы $\mathbf{H}^{\langle i,j\rangle}$ нулевые, а значит в произведении на m-той строке будет ноль. □

Определение 10. *Пусть имеется свидетельство* $\langle c_i, c_j\rangle$*, а* c_j *представляет собой конъюнкт* $x_{j_1}x_{j_2}\ldots x_{j_k}$*, тогда этому свидетельству соответствует пропозициональная формула*

$$q_{\langle i,j\rangle} = c_i\bar{x}_{j_1}\bar{x}_{j_2}\ldots\bar{x}_{j_k}.$$

Предложение 1. *Пусть задан фрагмент знаний со скалярными оценками* $(C, \mathbf{P_c})$*, тогда вероятность истинности свидетельства* $\langle i,j\rangle$ *при условии заданных фрагментом знаний может быть вычислена по формуле*

$$p(\langle i,j\rangle|(C,\mathbf{P_q})) = (\mathbf{1}, \mathbf{H}^{\langle i;j\rangle} \times \mathbf{P_q}).$$

Доказательство. Рассмотрим пропозициональную формулу $q_{\langle i,j\rangle}$, соответствующую свидетельству. Вероятность истинности этой пропозициональной формулы может быть выражена на основе характеристического вектора $\chi_{q_{\langle i,j\rangle}}$ (см. [4]). Таким образом, нам остаётся показать, что

$$(\mathbf{1}, \mathbf{H}^{\langle i;j\rangle} \times \mathbf{P_q}) = (\chi_{q_{\langle i,j\rangle}}, \mathbf{P_q}).$$

Воспользуемся свойствами скалярного произведения и тем, что по построению ${\mathbf{H}^{\langle i;j\rangle}}^{\mathsf{T}} = \mathbf{H}^{\langle i;j\rangle}$; получим, что надо проверить, что

$$(\mathbf{H}^{\langle i;j\rangle} \times \mathbf{1}, \mathbf{P_q}) = (\chi_{q_{\langle i,j\rangle}}, \mathbf{P_q}).$$

Но матрица $\mathbf{H}^{\langle i;j\rangle}$ строилась нами таким образом, что будучи применена к вектору $\mathbf{1}$, она «оставляет» только согласованные со свидетельством кванты. Согласованные со свидетельством кванты, это именно те кванты, которые входят в $S_{q_{\langle i,j\rangle}}$. Таким образом

$$\mathbf{H}^{\langle i;j\rangle^{\mathsf{T}}} \times \mathbf{1} = \chi_{q_{\langle i,j\rangle}}.$$

А значит

$$p(\langle i,j\rangle|(C,\mathbf{P_q})) = (\mathbf{1}, \mathbf{H}^{\langle i;j\rangle} \times \mathbf{P_q}).$$

□

Рассмотрим квант q_m и свидетельство $\langle c_i, c_j\rangle$. Вычислим $p(q_m|\langle c_i, c_j\rangle)$. Как мы уже отметили, при вычислениях можно заменить свидетельство $\langle c_i, c_j\rangle$ на соответствующую ему формулу $q_{\langle i,j\rangle}$. Воспользуемся определением условной вероятности:

$$p(q_m|q_{\langle i,j\rangle})p(q_{\langle i,j\rangle}) = p(q_m \wedge q_{\langle i,j\rangle}). \tag{2}$$

Особенности квантов и $q_{\langle i,j\rangle}$ гарантируют нам, что $q_m \wedge q_{\langle i,j\rangle}$ может либо принимать значение q_m, если квант согласован со свидетельством, либо значение $\mathbb{F}$, если не согласован. Теперь рассмотрим вектор

$$\mathbf{P}_{\mathbf{q}}^{\langle i,j\rangle} = \begin{pmatrix} p(q_0|q_{\langle i,j\rangle}) \\ \vdots \\ p(q_{2^n-1}|q_{\langle i,j\rangle}) \end{pmatrix}.$$

Данный вектор состоит из условных вероятностей квантов при заданном свидетельстве.

Воспользовавшись определением матрицы $\mathbf{H}$, можно записать (2) в векторном виде следующим образом:

$$\mathbf{P}_{\mathbf{q}}^{\langle i,j\rangle} p(\langle c_i, c_j\rangle) = \mathbf{H}^{\langle i;j\rangle} \times \mathbf{P_q}. \tag{3}$$

По аналогии с вектором $\mathbf{P}_{\mathbf{q}}^{\langle i,j\rangle}$ построим вектор

$$\mathbf{P}_{\mathbf{c}}^{\langle i,j\rangle} = \begin{pmatrix} p(c_0|q_{\langle i,j\rangle}) \\ \vdots \\ p(c_{2^n-1}|q_{\langle i,j\rangle}) \end{pmatrix}.$$

Предложение 2. *Векторы* $\mathbf{P}_\mathbf{q}^{\langle i,j\rangle}$ *и* $\mathbf{P}_\mathbf{c}^{\langle i,j\rangle}$ *связаны соотношениями*

$$\mathbf{P}_\mathbf{c}^{\langle i,j\rangle} = \mathbf{J}_n \times \mathbf{P}_\mathbf{q}^{\langle i,j\rangle} \text{ и } \mathbf{P}_\mathbf{q}^{\langle i,j\rangle} = \mathbf{I}_n \times \mathbf{P}_\mathbf{c}^{\langle i,j\rangle}.$$

Доказательство. Условные вероятности являются вероятностями. Таким образом, $\mathbf{P}_\mathbf{c}^{\langle i,j\rangle}$ соответствует одному из распределений, которое могло бы быть задано на $\mathbf{P}_\mathbf{c}$, а, следовательно, соотношения между $\mathbf{P}_\mathbf{c}$ и $\mathbf{P}_\mathbf{q}$, распространяется и на частный случай $\mathbf{P}_\mathbf{q}^{\langle i,j\rangle}$ и $\mathbf{P}_\mathbf{c}^{\langle i,j\rangle}$. □

Воспользовавшись соотношением между $\mathbf{P}_\mathbf{q}$ и $\mathbf{P}_\mathbf{c}$, можно переписать (3) как

$$\mathbf{I}_n \times \mathbf{P}_\mathbf{c}^{\langle i,j\rangle} p(\langle c_i, c_j\rangle) = \mathbf{H}^{\langle i;j\rangle} \times \mathbf{I}_n \times \mathbf{P}_\mathbf{c}. \quad (4)$$

Домножив обе части равенства на $\mathbf{J}_n$ и воспользовавшись тем, что

$$\mathbf{J}_n \times \mathbf{I}_n = \mathbf{E},$$

получаем

$$\mathbf{P}_\mathbf{c}^{\langle i,j\rangle} p(\langle c_i, c_j\rangle) = \mathbf{J}_n \times \mathbf{H}^{\langle i;j\rangle} \times \mathbf{I}_n \times \mathbf{P}_\mathbf{c}. \quad (5)$$

Рассмотрим матрицу

$$\mathbf{T}^{\langle i,j\rangle} = \mathbf{J}_n \times \mathbf{H}^{\langle i,j\rangle} \times \mathbf{I}_n.$$

При заданных i, j и n матрица $\mathbf{T}^{\langle i,j\rangle}$ задаётся однозначно, при этом

$$\mathbf{P}_\mathbf{c}^{\langle i,j\rangle} p(\langle c_i, c_j\rangle) = \mathbf{T}^{\langle i,j\rangle} \times \mathbf{P}_\mathbf{c}. \quad (6)$$

Теорема 1. *В случае ФЗ со скалярными оценками решением первой задачи апостериорного вывода для свидетельства* $\langle i, j\rangle$ *является* $(\mathbf{T}^{\langle i,j\rangle} \times \mathbf{P}_\mathbf{c})[0]$.

Доказательство. По построению $\mathbf{P}_\mathbf{c}^{\langle i,j\rangle}[0] = 1$, таким образом нулевой элемент произведения $\mathbf{T}^{\langle i,j\rangle} \times \mathbf{P}_\mathbf{c}$ — это в точности вероятность $p(\langle c_i, c_j\rangle)$. То есть

$$p(\langle c_i, c_j\rangle) = (\mathbf{T}^{\langle i,j\rangle} \times \mathbf{P}_\mathbf{c})[0]. \quad (7)$$

□

Если $p(\langle c_i, c_j \rangle) \neq 0$, то получаем

$$\mathbf{P_c}^{\langle i,j \rangle} = \frac{1}{(\mathbf{T}^{\langle i;j \rangle} \times \mathbf{P_c})[0]} \mathbf{T}^{\langle i;j \rangle} \times \mathbf{P_c} = \frac{\mathbf{T}^{\langle i;j \rangle} \times \mathbf{P_c}}{(\mathbf{T}^{\langle i;j \rangle} \times \mathbf{P_c})[0]}. \qquad (8)$$

Теорема 2. *Если $p(\langle c_i, c_j \rangle) \neq 0$, то формула (8) даёт решение второй задачи апостериорного вывода в случае детерминированного свидетельства.*

Доказательство. По приведённому выше построению. □

Замечание 3. *Если $p(\langle c_i, c_j \rangle) = 0$, то поступившее свидетельство имело вероятность 0, а значит, произойти не могло, так как рассматриваемое нами пространство событий является дискретным. В таком случае мы говорим, что свидетельство невероятное. Если же требуются апостериорные оценки, то решением второй задачи апостериорного вывода являются интервальные оценки $0 \leq \mathbf{P_c}^{\langle i,j \rangle}[k] \leq 1$ для всех элементов вектора $\mathbf{P_c}^{\langle i,j \rangle}$, кроме нулевого, чья вероятность будет равна единице.*

Опишем структуру матрицы

$$\mathsf{T}^{\langle i,j \rangle} = \mathbf{J}_n \times \mathbf{H}^{\langle i,j \rangle} \times \mathbf{I}_n.$$

Теорема 3. *Матрица $\mathsf{T}^{\langle i,j \rangle}$ равна тензорному произведению:*

$$\mathsf{T}^{\langle i,j \rangle} = \widetilde{\mathsf{T}}_{n-1}^{\langle i,j \rangle} \otimes \widetilde{\mathsf{T}}_{n-2}^{\langle i,j \rangle} \otimes \ldots \widetilde{\mathsf{T}}_{0}^{\langle i,j \rangle},$$

где

$$\widetilde{\mathsf{T}}_{k}^{\langle i,j \rangle} = \begin{cases} \mathsf{T}^{+}, & \text{если } x_k \text{ входит в } c_i; \\ \mathsf{T}^{-}, & \text{если } x_k \text{ входит в } c_j; \\ \mathsf{T}^{\circ}, & \text{иначе;} \end{cases}$$

причём

$$\mathsf{T}^{+} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}; \qquad \mathsf{T}^{-} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}; \qquad \mathsf{T}^{\circ} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Доказательство. Воспользовавшись выражением матриц $\mathbf{J}_n$, $\mathbf{H}^{\langle i,j \rangle}$ и $\mathbf{I}_n$ через тензорное произведение, получаем:

$$\mathsf{T}^{\langle i,j \rangle} = (\mathbf{J}_1 \otimes \ldots \otimes \mathbf{J}_1) \times (\widetilde{\mathbf{H}}_{n-1}^{\langle i,j \rangle} \otimes \widetilde{\mathbf{H}}_{n-2}^{\langle i,j \rangle} \otimes \ldots \widetilde{\mathbf{H}}_{0}^{\langle i,j \rangle}) \times (\mathbf{J}_1 \otimes \ldots \otimes \mathbf{J}_1).$$

Воспользовавшись свойствами тензорного произведения, получаем

$$\mathsf{T}^{\langle i,j\rangle} = (\mathbf{J}_1 \times \widetilde{\mathsf{H}}_{n-1}^{\langle i,j\rangle} \times \mathbf{I}_1) \otimes (\mathbf{J}_1 \times \widetilde{\mathsf{H}}_{n-2}^{\langle i,j\rangle} \times \mathbf{I}_1) \otimes \ldots \otimes (\mathbf{J}_1 \times \widetilde{\mathsf{H}}_{0}^{\langle i,j\rangle} \times \mathbf{I}_1).$$

Введем обозначение $\widetilde{\mathsf{T}}^{\langle i,j_k\rangle} = (\mathbf{J}_1 \times \widetilde{\mathsf{H}}_{k}^{\langle i,j\rangle} \times \mathbf{I}_1)$ получаем

$$\mathsf{T}^{\langle i,j\rangle} = \widetilde{\mathsf{T}}_{n-1}^{\langle i,j\rangle} \otimes \widetilde{\mathsf{T}}_{n-2}^{\langle i,j\rangle} \otimes \ldots \otimes \widetilde{\mathsf{T}}_{0}^{\langle i,j\rangle}.$$

Матрица $\widetilde{\mathsf{T}}$ может принимать одно из трёх возможных значений:

$$\mathsf{T}^{+} = (\mathbf{J}_1 \times \mathsf{H}_k^{+} \times \mathbf{I}_1) = \\ = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$

$$\mathsf{T}^{-} = (\mathbf{J}_1 \times \mathsf{H}_k^{-} \times \mathbf{I}_1) = \\ = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \times \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix},$$

$$\mathsf{T}^{\circ} = (\mathbf{J}_1 \times \mathsf{H}_k^{\circ} \times \mathbf{I}_1) = \\ = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

□

Замечание 4. *Подобное представление говорит нам о том, что ненормированный априорный вывод является линейным преобразованием. Это, в свою очередь, позволяет применять методы линейной алгебры для исследования устойчивости и чувствительности данного процесса.*

3. Апостериорный вывод над интервальными оценками. Пусть задан ФЗ с оценками $\mathbf{P_c}^{-}$ и $\mathbf{P_c}^{+}$, тогда решение первой и второй задач апостериорного вывода для детерминированного свидетельства можно представить как решение серии задач линейного программирования. Начнём с первой задачи апостериорного вывода.

Предложение 3. *Для детерминированного свидетельства* $\langle i, j\rangle$ *существует вектор* $\mathcal{L}^{\langle i,j\rangle}$ *такой, что для решения первой задачи апостериорного вывода при данном свидетельстве и интервальных оценках во фрагменте знаний необходимо решить две задачи линейного программирования, соответствующие поиску величин:*

$$\min_{\mathcal{D}\cup\mathcal{E}}\{(\mathcal{L}^{\langle i,j\rangle}, \mathbf{P_c})\} \quad \text{\textit{и}} \quad \max_{\mathcal{D}\cup\mathcal{E}}\{(\mathcal{L}^{\langle i,j\rangle}, \mathbf{P_c})\}. \tag{9}$$

Доказательство. В случае детерминированного свидетельства первая задача апостериорного вывода может быть решена как задача априорного вывода для формулы, соответствующей свидетельству. Согласно формуле (7) $p(\langle c_i, c_j\rangle) = (\mathbf{T}^{\langle i,j\rangle} \times \mathbf{P_c})[0]$. На основе этой формулы можно построить вектор $\mathcal{L}^{\langle i,j\rangle}$, соответствующий первой строке матрицы $\mathbf{T}^{\langle i,j\rangle}$, такой что

$$p(\langle c_i, c_j\rangle) = (\mathcal{L}^{\langle i,j\rangle}, \mathbf{P_c}). \tag{10}$$

Из чего и следует утверждение. □

Перейдём к решению второй задачи апостериорного вывода. Для этого требуется решить задачи по поиску:

$$\min_{\mathcal{D}\cup\mathcal{E}}\{\mathbf{P_c}^{\langle i,j\rangle}[i]\} \quad \text{и} \quad \max_{\mathcal{D}\cup\mathcal{E}}\{\mathbf{P_c}^{\langle i,j\rangle}[i]\}. \tag{11}$$

Замечание 5. *Следует отметить, что указанные задачи могут решаться абсолютно независимо. Это может быть полезно, например при использовании параллельных вычислений.*

Теорема 4. *Решением второй задачи локального апостериорного вывода в случае детерминированного свидетельства и фрагмента знаний с интервальными оценками является решение серии задач линейного программирования по поиску:*

$$\min\{\mathbf{T}^{\langle i;j\rangle} \times \mathbf{D}\} \quad \text{\textit{и}} \quad \max\{\mathbf{T}^{\langle i;j\rangle} \times \mathbf{D}\}, \tag{12}$$

при условиях:

$$\{\lambda\mathbf{P_c^-} \le \mathbf{D}\} \cup \{\mathbf{D} \le \lambda\mathbf{P_c^+}\} \cup \{\mathbf{I_n} \times \mathbf{D} \geqslant \mathbf{0}\} \cup \{(\mathcal{L}^{\langle i,j\rangle}, \mathbf{D}) = 1\} \cup \{\lambda \ge 0\}$$

Задачи соответствующие разным элементам вектора $\mathbf{T}^{\langle i;j\rangle} \times \mathbf{D}$ *решаются независимо. Построенные таким образом оценки задают непротиворечивый ФЗ.*

Доказательство. Из рассмотренного выше случая скалярных оценок (формула (8)) известно, что

$$\mathbf{P_c}^{\langle i,j\rangle} = \frac{1}{(\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c})[0]}\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c}.$$

Из-за наличия множителя $\frac{1}{(\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c})[0]}$ построенные задачи не являются задачами линейного программирования, а относятся к классу задач дробно-линейного программирования. В работе [**1**] доказывается, что при соблюдении ряда условий можно свести решение задачи дробно-линейного программирования к решению задачи линейного программирования. Мы не будем подробно останавливаться на общем случае, а рассмотрим частную реализацию подобного перехода от задачи дробно-линейного программирования к задаче линейного программирования, основанную на работе [**14**]. Введём замену переменных

$$\mathbf{D} = \lambda \mathbf{P_c}.$$

Тогда

$$\frac{1}{(\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c})[0]}\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c} = \frac{\lambda}{\lambda(\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c})[0]}\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c} =$$

$$= \frac{1}{(\mathbf{T}^{\langle i;j\rangle} \times (\lambda\mathbf{P_c}))[0]}\mathbf{T}^{\langle i;j\rangle} \times (\lambda\mathbf{P_c}) = \frac{1}{(\mathbf{T}^{\langle i;j\rangle} \times \mathbf{D})[0]}\mathbf{T}^{\langle i;j\rangle} \times \mathbf{D}.$$

То есть:

$$\mathbf{P_c}^{\langle i,j\rangle} = \frac{1}{(\mathbf{T}^{\langle i;j\rangle} \times \mathbf{D})[0]}\mathbf{T}^{\langle i;j\rangle} \times \mathbf{D}. \tag{13}$$

Так как мы произвольным образом выбирали λ, то можно подобрать его таким образом, чтобы $(\mathbf{T}^{\langle i;j\rangle} \times \mathbf{D})[0] = 1$. Данное условие можно переписать в виде

$$(\mathcal{L}^{\langle i,j\rangle}, \mathbf{D}) = 1 \tag{14}$$

Но если λ выбрана таким образом, что выполняется (14), тогда (13) принимает вид

$$\mathbf{P_c}^{\langle i,j\rangle} = \mathbf{T}^{\langle i;j\rangle} \times \mathbf{D}. \tag{15}$$

Теперь, объединяя все выше полученное, можно сформулировать серию задач линейного программирования, решение которых будет решением второй задачи апостериорного вывода в случае детерминированного свидетельства и интервальных оценок во фрагменте знаний.

Опираясь на (15), можно переформулировать (11) как поиск величин, выраженных через переменные $\mathbf{D}$:

$$\min_{\mathcal{D}\cup\mathcal{E}\cup\{(\mathcal{L}^{\langle i,j\rangle},\mathbf{D})=1\}}\{\mathbf{T}^{\langle i;j\rangle}\times\mathbf{D}\}\quad\text{и}\quad\max_{\mathcal{D}\cup\mathcal{E}\cup\{(\mathcal{L}^{\langle i,j\rangle},\mathbf{D})=1\}}\{\mathbf{T}^{\langle i;j\rangle}\times\mathbf{D}\}. \tag{16}$$

Теперь остаётся только выразить ограничения $\mathcal{D}$ и $\mathcal{E}$ через переменные $\mathbf{D}$. Вспомним, что

$$\mathbf{D}=\lambda\mathbf{P_c}.$$

Тогда ограничение 14 принимает вид

$$\lambda(\mathcal{L}^{\langle i,j\rangle},\mathbf{P_c})=1,$$

то есть

$$\lambda=\frac{1}{(\mathcal{L}^{\langle i,j\rangle},\mathbf{P_c})}.$$

Но мы знаем, что $(\mathcal{L}^{\langle i,j\rangle},\mathbf{P_c})=p(\langle c_i,c_j\rangle)$. Так как $p(\langle c_i,c_j\rangle)$ — вероятность, то $0\leq p(\langle c_i,c_j\rangle)\leq 1$, а следовательно

$$\lambda\geq 1.$$

Согласно [4]:

$$\mathcal{D}=\{\mathbf{P_c^-}\leq\mathbf{P_c}\}\cup\{\mathbf{P_c}\leq\mathbf{P_c^+}\}.$$

Так как λ заведомо положительно, данные ограничения преобразуются в

$$\mathcal{D}=\{\lambda\mathbf{P_c^-}\leq\mathbf{D}\}\cup\{\mathbf{D}\leq\lambda\mathbf{P_c^+}\}.$$

Аналогично

$$\mathcal{E}=\{\mathbf{I_n}\times\mathbf{D}\geqslant\mathbf{0}\}.$$

□

4. Недетерминированные свидетельства.

Определение 11. *Пусть задан идеал конъюнктов C, тогда будем использовать обозначение $\mathcal{A}_C$ для сужения алфавита $\mathcal{A}$ до элементов, над которыми задан C, а Q_C — для множества квантов, заданных над этим алфавитом.*

Пусть нам задан фрагмент знаний со скалярными оценками $(C, \mathbf{P_c})$. И согласно определению 5 задан непротиворечивый фрагмент знаний $(C', \mathbf{P_c}^a)$, такой что $C' \subseteq C$. Так как $(C', \mathbf{P_c}^a)$ — непротиворечивый, то на основе оценок вероятностей для конъюнктов $\mathbf{P_c}^a$, можно построить оценки вероятностей для квантов $Q_{C'}$:

$$\mathbf{P_q}^a = \mathbf{I} \times \mathbf{P_c}^a.$$

Элементы множества $Q_{C'}$ являются взаимоисключающими, то есть при любом зафиксированном означивании всех атомов истинным будет один и только один элемент множества $Q_{C'}$. Кроме того, каждый элемент $Q_{C'}$ можно трактовать как соответствующее детерминированное свидетельство для фрагмента знаний $(C, \mathbf{P_c})$. Пользуясь этими свойствами, мы можем формализовать стохастическое свидетельство следующим образом: свидетельство $(C', \mathbf{P_c}^a)$ — случайный элемент, принимающий значение детерминированных свидетельств, соответствующих элементам $Q_{C'}$ с вероятностями $\mathbf{P_q}^a = \mathbf{I} \times \mathbf{P_c}^a$. При такой формализации мы можем перейти к описанию двух задач апостериорного вывода. Так как свидетельство $(C', \mathbf{P_c}^a)$ является случайным элементом, а в качестве решений первой и второй задач апостериорного вывода хочется получать численные значения, то решениями мы будем считать соответствующие математические ожидания.

Рассмотрим решение первой задачи апостериорного вывода. Для детерминированного свидетельства $\langle i, j \rangle$ решением первой задачи апостериорного выводя является оценка $(\mathbf{T}^{\langle i,j \rangle} \times \mathbf{P_c})[0]$. Нам необходимо сопоставить индексы детерминированных свидетельств с индексами множества $Q_{C'}$. Для этого определим функцию $\mathrm{GI}(i, m)$, здесь m — это индекс наибольшего элемента C' в алфавите $\mathcal{A}$, i — индекс конъюнкта в алфавите $\mathcal{A}_{C'}$, а функция возвращает индекс соответствующего конъюнкта в алфавите $\mathcal{A}$.

Описать, что при этом происходит, можно, например, следующим образом. Запомним на каких позициях в m стоят нули и «вставим» нули на эти же позиции в i, сдвигая элементы. Таким образом m будет служить маской, определяющей на какие элементы алфавита $\mathcal{A}$ проецируется индекс i.

Обозначим мощность алфавита $\mathcal{A}_{C'}$ через n'. В дальнейших выкладках нам постоянно понадобится «дополнение» i до $2^{n'} - 1$ равное $\neg i$. Это индекс «оставшейся» отрицательно означеной части. Для сокращения записи во всех нижеследующих формулах

$\neg i$ обозначает это дополнение. Выпишем в этих обозначениях математическое ожидание, которое и будет решением первой задачи апостериорного вывода:

$$\sum_{i=0}^{2^{n'}-1} (\mathbf{T}^{\langle GI(i,m),GI(\neg i,m)\rangle} \times \mathbf{P_c})[0](\mathbf{I} \times \mathbf{P_c}^a)[i]. \qquad (17)$$

Напомним, что решением второй задачи апостериорного вывода в случае детерминированного свидетельства являются оценки:

$$\mathbf{P_c}^{\langle i,j\rangle} = \frac{\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c}}{(\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c})[0]}.$$

Математическое ожидание, которое нам необходимо вычислить:

$$\mathbf{P_c}^{\langle (C',\mathbf{P_c}^a)\rangle} = \sum_{i=0}^{2^{n'}-1} \mathbf{P_c}^{\langle GI(i,m),GI(\neg i,m)\rangle}(\mathbf{I} \times \mathbf{P_c}^a)[i].$$

И соответственно, если все детерминированные свидетельства возможны (имеют ненулевую вероятность), то решением второй задачи апостериорного вывода для стохастического свидетельства будет:

$$\mathbf{P_c}^{\langle (C',\mathbf{P_c}^a)\rangle} = \sum_{i=0}^{2^{n'}-1} \frac{\mathbf{T}^{\langle GI(i,m),GI(\neg i,m)\rangle} \times \mathbf{P_c}}{(\mathbf{T}^{\langle GI(i,m),GI(\neg i,m)\rangle} \times \mathbf{P_c})[0]}(\mathbf{I} \times \mathbf{P_c}^a)[i]. \qquad (18)$$

Если какое-либо детерминированное свидетельство невозможно (его вероятность равна нулю), и при этом оценки, задаваемые стохастическим свидетельством для данного детерминированного свидетельства, нулю не равны, то мы говорим, что такое свидетельство *невероятное*.

Есть несколько способов трактовки невероятного свидетельства. В первую очередь, появление таких свидетельств показывает на ошибку либо в исходной базе знаний, либо в свидетельстве. Если мы считаем, что ошибка в свидетельстве, то мы можем его скорректировать, так что соответствующее детерминированное свидетельство будет иметь нулевую вероятность. Другой путь, это сопоставить такому детерминированному свидетельству максимально неопределённый результат. То есть ФЗ, в котором все оценки истинности конъюнктов лежат в диапазоне от нуля до единицы.

Рассмотрим следующий набор исходных данных: фрагмент знаний с интервальными оценками и стохастическое свидетельство.

Для решения первой задачи апостериорного вывода нам необходимо найти максимум и минимум значения

$$\sum_{i=0}^{2^{n'}-1} (\mathbf{T}^{\langle GI(i,m),GI(\neg i,m)\rangle} \times \mathbf{P_c})[0](\mathbf{I} \times \mathbf{P_c}^a)[i]. \tag{19}$$

Заметим, что данная функция линейная и, следовательно, нам достаточно решить две соответствующие задачи линейного программирования.

К сожалению, для второй задачи апостериорного вывода выписанная формула (18) не является линейной, и точно найти соответствующие максимумы и минимумы, не выходя за рамки задач линейного программирования, представляется невозможным. Поэтому мы построим не точное решение, а накрывающее его. Вычислим оценки

$$\mathbf{P_c}^{\langle GI(i,m),GI(\neg i,m)\rangle,\min} = \min_{\{\mathcal{D}\cup\mathcal{E}\}} \left(\frac{\mathbf{T}^{\langle GI(i,m),GI(\neg i,m)\rangle} \times \mathbf{P_c}}{(\mathbf{T}^{\langle GI(i,m),GI(\neg i,m)\rangle} \times \mathbf{P_c})[0]} \right), \tag{20}$$

$$\mathbf{P_c}^{\langle GI(i,m),GI(\neg i,m)\rangle,\max} = \max_{\{\mathcal{D}\cup\mathcal{E}\}} \left(\frac{\mathbf{T}^{\langle GI(i,m),GI(\neg i,m)\rangle} \times \mathbf{P_c}}{(\mathbf{T}^{\langle GI(i,m),GI(\neg i,m)\rangle} \times \mathbf{P_c})[0]} \right) \tag{21}$$

и будем считать решением второй задачи апостериорного вывода в случае стохастического свидетельства и интервальных оценок оценки:

$$\sum_{i=0}^{2^{n'}-1} \mathbf{P_c}^{\langle GI(i,m),GI(\neg i,m)\rangle,\min}(\mathbf{I} \times \mathbf{P_c}^a)[i]; \tag{22}$$

$$\sum_{i=0}^{2^{n'}-1} \mathbf{P_c}^{\langle GI(i,m),GI(\neg i,m)\rangle,\max}(\mathbf{I} \times \mathbf{P_c}^a)[i]. \tag{23}$$

Рассмотрим последний вид свидетельств — неточные. Неточное свидетельство представляет собой ФЗ с интервальными оценками $(C', \mathbf{P_c}^{a,-}, \mathbf{P_c}^{a,+})$. Мы будем трактовать его следующим образом: неточное свидетельство $(C', \mathbf{P_c}^{a,-}, \mathbf{P_c}^{a,+})$ — это семейство всех возможных стохастических свидетельств $(C', \mathbf{P_c}^a)$, таких что $\mathbf{P_c}^{a,-} \leq \mathbf{P_c}^a \leq \mathbf{P_c}^{a,+}$.

К сожалению, в случае неточных свидетельств мы сталкиваемся с выходом за задачи линейного программирования уже на этапе решения первой задачи апостериорного вывода. Это связано с тем, что формула (19) перестаёт быть линейной. Мы снова приблизим решение, зафиксировав максимумы и минимумы значений оценок для каждого детерминированного свидетельства. Получаем две линейных оценки, которые накроют истинные интервальные оценки для решения первой задачи апостериорного вывода.

$$\min_{\mathbf{P_c}^{a,-} \leq \mathbf{P_c}^a \leq \mathbf{P_c}^{a,+}} \left(\sum_{i=0}^{2^{n'}-1} \min_{\mathcal{E} \cup \mathcal{D}} ((\mathbf{T}^{\langle GI(i,m), GI(\neg i,m) \rangle} \times \mathbf{P_c})[0])(\mathbf{I} \times \mathbf{P_c}^a)[i] \right); \tag{24}$$

$$\max_{\mathbf{P_c}^{a,-} \leq \mathbf{P_c}^a \leq \mathbf{P_c}^{a,+}} \left(\sum_{i=0}^{2^{n'}-1} \max_{\mathcal{E} \cup \mathcal{D}} ((\mathbf{T}^{\langle GI(i,m), GI(\neg i,m) \rangle} \times \mathbf{P_c})[0])(\mathbf{I} \times \mathbf{P_c}^a)[i] \right). \tag{25}$$

Заметим, что если исходный ФЗ имел скалярные оценки, то эти формулы точные, а не накрывающие.

Для решения второй задачи апостериорного вывода достаточно найти:

$$\min_{\mathbf{P_c}^{a,-} \leq \mathbf{P_c}^a \leq \mathbf{P_c}^{a,+}} \left(\sum_{i=0}^{2^{n'}-1} \mathbf{P_c}^{\langle GI(i,m), GI(\neg i,m) \rangle, \min} (\mathbf{I} \times \mathbf{P_c}^a)[i] \right); \tag{26}$$

$$\max_{\mathbf{P_c}^{a,-} \leq \mathbf{P_c}^a \leq \mathbf{P_c}^{a,+}} \left(\sum_{i=0}^{2^{n'}-1} \mathbf{P_c}^{\langle GI(i,m), GI(\neg i,m) \rangle, \max} (\mathbf{I} \times \mathbf{P_c}^a)[i] \right). \tag{27}$$

Пример 1. *Разница между точной и накрывающей оценкой апостериорного вывода. Рассмотрим ФЗ, заданный над двумя атомами* $\{x, y\}$*, со следующими оценками:*

$$p(x) \in [0.3, 0.6],$$

$$p(y) \in [0.2, 0.6],$$

$$p(xy) \in [0.2, 0.5].$$

Пусть поступило свидетельство $\langle p(x) = 0.8\rangle$*. Вычислим апостериорную оценку для* $p(y)$*. Для этого, в частности, необходимо вычислить*

$$\min_{\{\mathcal{D}\cup\mathcal{E}\}}\left(\frac{p(xy)}{p(x)}\cdot 0.8 + \frac{p(y)-p(xy)}{1-p(x)}\cdot 0.2\right).$$

Минимум этой функции — 0.64 *достигается при* $p(x) = 0.5$ *и* $p(xy) = 0.2$*, соответственно. Но как было уже отмечено выше, данная задача не является задачей линейного программирования. Рассмотрим теперь два детерминированных свидетельства* $\langle x\rangle$ *и* $\langle \bar{x}\rangle$*. Для них получаем:*

$$\min_{\{\mathcal{D}\cup\mathcal{E}\}}\left(\frac{p(xy)}{p(x)}\right) = \frac{1}{3},$$

$$\min_{\{\mathcal{D}\cup\mathcal{E}\}}\left(\frac{p(y)-p(xy)}{1-p(x)}\right) = \frac{5}{7}.$$

И в соответствии с формулой (22) получаем нижнюю оценку для $p(y)$

$$\frac{1}{3}\cdot 0.8 + \frac{5}{7}\cdot 0.2 \approx 0.4095.$$

Таким образом точной нижней границе 0.64*, мы, не выходя за рамки ЗЛП, можем сопоставить только накрывающую* 0.4095*.*

5. Оценки сложности алгоритмов локального апостериорного вывода. Начнём разбор с пропагации детерминированного свидетельства в случае точечных оценок. Согласно теоремам и для проведения апостериорного вывода нам потребуется вычислить произведение

$$\mathbf{T}^{\langle i;j\rangle} \times \mathbf{P_c},$$

где $\langle i;j\rangle$ индекс поступившего свидетельства. Нам необходимо оценить количество ненулевых элементов матрицы $\mathbf{T}^{\langle i;j\rangle}$. Из описания матрицы $\mathbf{T}^{\langle i;j\rangle}$, приведённого в теореме 3, следует, что количество не нулевых элементов в этой матрице равно 2^n и не зависит от $\langle i;j\rangle$.

Предложение 4. *Для проведения пропагации детерминированного свидетельства требуется сделать не более $2^n - 2^{n-k}$ сложений/вычитаний и $2^{n-k} - 1$ делений на одно и то же число, где n — общее число атомов, а k — число атомов, получивших отрицательное означивание.*

Доказательство. Как мы показали раньше, в матрице $\mathbf{T}^{\langle i;j \rangle}$ ровно 2^n ненулевых элементов, каждый из которых либо один, либо минус один. Следовательно надо сделать не более $2^n - t$ сложений/вычитаний элементов вектора $\mathbf{P_c}$, где t — количество строчек матрицы, содержащее ненулевые элементы, для вычисления произведения

$$\mathbf{T}^{\langle i;j \rangle} \times \mathbf{P_c}.$$

Оценим количество ненулевых строчек. Строчка, состоящая из нулей, возникает для конъюнктов, которые содержат атомы, получившие отрицательное означивание. Следовательно, ненулевые строки соответствуют конъюнктам, не содержащим отрицательно означенные атомы. Очевидно, что таких 2^{n-k}; таким образом мы оценили t и получаем, что необходимо сделать $2^n - 2^{n-k}$ сложений/вычитаний. Только заведомо ненулевые строки потребуют нормализации, то есть потребуется ещё сделать $2^{n-k} - 1$ делений на нормирующий элемент. Минус единица возникает из-за того, что нормирующий делитель равен нулевому элементу, и деление не требуется. □

Для случая интервальных исходных оценок нам придётся решать серию ЗЛП.

Предложение 5. *Для пропагации детерминированного свидетельства в ФЗ с интервальными оценками истинности потребуется решить 2^{n-k+1} задач линейного программирования, содержащих $2(2^n)$ ограничений на переменные из предметной области, и 2^n ограничений из аксиоматики теории вероятностей и одно дополнительное ограничение, где k — количество атомарных пропозиций, получивших отрицательное означивание, а n — общее число атомарных пропозиций, входящих во фрагмент знаний.*

Доказательство. Доказательство этой оценки следует непосредственно из теоремы 4 (оценка на сложность каждой ЗЛП) и определения матрицы $\mathbf{T}^{\langle i;j \rangle}$ (оценка количества ЗЛП, которые необходимо решить). Проведём пропагацию детерминированного свидетельства. Очевидно, что после пропагации вероятность любого конъюнкта, содержащего атомарную пропозицию, равна нулю. Таким образом, вычислять их не потребуется. Оценим количество конъюнктов, которые могут получить ненулевую оценку вероятности. Это те, которые содержат только атомарные пропозиции, получившие положительные означивания или не получившие никакого. Число таких атомарных пропозиций — $n - k$, следовательно, количество возможных конъюнктов оценивается числом 2^{n-k}. Нам необходимо для каждого такого конъюнкта решить 2 ЗЛП для поиска оценки снизу и оценки сверху соответственно. Из чего и следует оценка числа ЗЛП 2^{n-k+1}. □

Два предыдущих утверждения полностью покрыли случай пропагации детерминированного свидетельства. Теперь естественным будет рассмотрение стохастического, а затем неточного свидетельств.

Приняв во внимание (17) и (18), можно оценить сложность соответствующих вычислений.

Предложение 6. *Для проведения пропагации стохастического свидетельства в ФЗ со скалярными оценками истинности требуется сделать не более $2^{n+n'} - 2^{n'+1} + 1$ сложений/вычитаний, $2^{n-n'} \cdot 3^{n'} - 2^{n'}$ делений и $2^{n-n'} \cdot 3^{n'} - 2^{n'}$ умножений, где n — общее число атомов, а n' — число атомов входящих в свидетельство.*

Доказательство. Согласно формулам (17) и (18) нам необходимо провести суммирование по всем означиваниям набора из n' атомов. Как мы доказали (предложение 4), сложность пропагации детерминированного свидетельства зависит от числа отрицательно означенных атомов и не зависит от того, какие именно атомы положительно означены. Следовательно, сумму из формулы (18) можно разбить на следующие части. Во-первых, вычислив количество сложений/вычитаний, необходимое для пропагации всех означиваний детерминированного свидетельства, получаем:

$$\sum_{k=0}^{n'} C_n^k(2^n - 2^{n-k}) = 2^n \cdot \sum_{k=0}^{n'} C_n^k - \sum_{k=0}^{n'} C_n^k \cdot 2^{n-k} =$$
$$= 2^n \cdot 2^{n'} - 2^{n-n'} \cdot \sum_{i=0}^{n'} C_n^i \cdot 2^i = 2^{n+n'} - 2^{n-n'} 3^{n'} = 2^{n-n'}(4^{n'} - 3^{n'}). \tag{28}$$

Аналогично, оценка на число необходимых делений:

$$\sum_{k=0}^{n'} (2^{n-k} - 1) = 2^{n-n'} \cdot 3^{n'} - 2^{n'}.$$

Во-вторых, домножим результаты пропагации детерминированных свидетельств на апостериорные оценки. Количество необходимых умножений, в точности соответствует ранее вычисленному количеству делений. В третьих, для каждого конъюнкта просуммируем его оценки для разных свидетельств. Нет смысла добавлять заведомо нулевые оценки конъюнктов. Чтобы оценить число необходимых сложений, будем рассуждать нижеследующим образом. Мы посчитали количество ненулевых элементов, когда считали количество делений и/или умножений. При этом мы исключили нулевой конъюнкт, так как его оценка неизменна и будет равна единице, а значит вычислять ее нет необходимости. И, наконец, для каждого конъюнкта сложений потребуется ровно на одно меньше, чем число ненулевых значений его оценки после пропагации детерминированного свидетельства, то есть общее число сложений:

$$2^{n-n'} \cdot 3^{n'} - 2^{n'} - (2^{n'} - 1) = 2^{n-n'} \cdot 3^{n'} - 2^{n'+1} + 1.$$

Осталось только объединить оценки на число сложений/вычитаний, соответствующие разным этапам. Итого, сложений/вычитаний необходимо:

$$2^{n-n'}(4^{n'} - 3^{n'}) + 2^{n-n'} \cdot 3^{n'} - 2^{n'+1} + 1 =$$
$$= 2^{n-n'} \cdot 4^{n'} - 2^{n'+1} + 1 = 2^{n+n'} - 2^{n'+1} + 1. \tag{29}$$

□

Теперь перейдём к оценке сложности пропагации стохастического свидетельства в ФЗ с интервальными оценками. Как мы уже отметили раньше, найти точное решение, не выходя за рамки задач линейного программирования, в общем случае невозможно. Поэтому мы дадим оценки сложности построения приближенной накрывающей оценки, задаваемой формулами (22) и (23).

Предложение 7. *Для пропагации стохастического свидетельства в ФЗ с интервальными оценками требуется решить* $2^{n-n'+1}3^{n'}$ *ЗЛП, содержащие* $2(2^n)$ *ограничений на переменные из предметной области,* 2^n *ограничений из аксиоматики теории вероятностей и одно дополнительное ограничение, а также сделать* $2^{n-n'+1}3^{n'}$ *умножений и* $2^{n-n'+1}3^{n'} - 2^{n+1}$ *сложений.*

Доказательство. В предложении 5 приведены оценки на сложность пропагации одного детерминированного свидетельства. Нам же потребуется провести подобную пропагацию $2^{n'}$ раз, после чего просуммировать полученные результаты для нижней и верхней оценки каждого коньюнкта. Заметим, что количество ограничений в каждой задаче не зависит от количества атомов, входящих в детерминированное свидетельство. Таким образом, нам надо оценить, сколько задач линейного программирования потребуется решить. Так как мы перебираем все возможные детерминированные свидетельства над n' атомами, а число задач зависит от числа отрицательно означенных пропозиций, то общее число задач можно представить в виде суммы:

$$\sum_{k=0}^{n'} C_n^k(2^{n-k+1}) = 2^{n-n'+1} \cdot \sum_{i=0}^{n'} C_n^i \cdot 2^i = 2^{n-n'+1}3^{n'}. \qquad (30)$$

Таким образом, требуется решить $2^{n-n'+1}3^{n'}$ ЗЛП, содержащие $2(2^n)$ ограничений на переменные из предметной области, 2^n ограничений из аксиоматики теории вероятностей и одно дополнительное ограничение. После этого остаётся просуммировать результаты. Заметим, что число ЗЛП определяет количество конъюнктов с ненулевыми оценками. Каждую ненулевую оценку потребуется домножить на вероятность соответствующего детерминированного свидетельства. Это даст нам $2^{n-n'+1}3^{n'}$ умножений. После чего, нам потребуется $2^{n-n'+1}3^{n'} - 2^{n+1}$ сложений. Слагаемое 2^{n+1} возникает из-за того, что для каждой оценки (верхней и нижней) каждого конъюнкта сложений на единицу меньше, чем слагаемых (соответствующих ЗЛП), и для каждого конъюнкта, в общем случае, есть хотя бы одна не нулевая оценка. □

Остался последний случай апостериорного вывода — пропагация неточного свидетельства. Как и в случае со стохастическим свидетельством, находясь в рамках задач линейного программирования, мы можем предложить только накрывающую оценку.

Предложение 8. *Для пропагации неточного свидетельства в ФЗ с интервальными оценками требуется решить* $2^{n-n'+1}3^{n'}$ *ЗЛП, содержащие* $2(2^n)$ *ограничений на переменные из предметной области,* 2^n *ограничений из аксиоматики теории вероятностей и одно дополнительное ограничение каждая, а также* $2(2^n - 1)$ *задач линейного программирования. Каждая из этих задач над* $2^{n'} - 1$ *переменной содержит* $2(2^{n'} - 1)$ *ограничений на переменные из предметной области,* $2^{n'}$ *ограничений из аксиоматики теории вероятностей, где* n *— количество атомов в ФЗ, а* n' *— количество атомов в свидетельстве.*

Доказательство. Заметим, что первый этап — это пропагация детерминированных свидетельств. Его сложность мы уже вычислили в доказательстве утверждения 7, формула (30). Теперь мы должны использовать формулы (26) и (27) для каждого конъюнкта. То есть мы должны решить еще $2(2^n - 1)$ задач линейного программирования. Каждая из этих задач над $2^{n'} - 1$ переменной содержит $2(2^{n'} - 1)$ ограничений на переменные из предметной области, и $2^{n'}$ ограничений из аксиоматики теории вероятностей. □

6. Заключение. В работе мы рассмотрели локальный апостериорный вывод в алгебраических байесовских сетях. Затем описали задачи апостериорного вывода и способы их решения для трёх видов свидетельств (детерминированных, стохастических и неточных) и двух видов оценок в ФЗ АБС (скалярных и интервальных). Наконец, построили комплекс оценок сложности для алгоритмов пропагации. Следует отметить, что так как количество исходных данных порядка 2^n, где n — число атомов в ФЗ, то все построенные оценки сложности либо линейны, либо могут быть мажорированы полиномом степени $\log_2 3 \approx 1.585$.

Кроме того, в статье приведён пример, который демонстрирует, что в общем случае для ФЗ с интервальными оценками невозможно построить точные оценки пропагации стохастического свидетельства оставаясь в рамках ЗЛП и не накладывая дополнительных ограничений.

Литература

1. *Гавурин М.К., Малоземов В.Н.* Экстремальные задачи с линейными ограничениями: учебное пособие. Л.: ЛГУ, 1984. 175 с.

2. *Городецкий В.И., Тулупьев А.Л.* Формирование непротиворечивых баз знаний с неопределенностью // Изв. РАН. Сер. Теория и системы управления. 1997. Т. 5. С. 33–42.

3. *Сироткин А.В., Тулупьев А.Л.* Матричные уравнения локального логико-вероятностного вывода в алгебраических байесовских сетях // Труды СПИИРАН. 2008. Вып. 6. СПб.: Наука, 2008. С. 134–143.

4. *Сироткин А.В., Тулупьев А.Л.* Моделирование знаний и рассуждений в условиях неопределенности: матрично-векторная формализация локального синтеза согласованных оценок истинности. // Труды СПИИРАН. 2011. Вып. 3(18). [*В настоящем выпуске*].

5. *Тулупьев А.Л.* Алгебраические байесовские сети: теоретические основы и непротиворечивость. СПб.: СПИИРАН, 1995. 76 с.

6. *Тулупьев А.Л.* Алгебраические байесовские сети: логико-вероятностный подход к моделированию баз знаний с неопределенностью. СПб.: СПИИРАН, 2000. 282 с.

7. *Тулупьев А.Л.* Метод построения и исследования баз фрагментов знаний с неопределенностью // Труды СПИИРАН. Вып. 1. 2002. Т. 1. С. 258–271.

8. *Тулупьев А.Л.* Алгебраические байесовские сети: глобальный логико-вероятностный вывод в деревьях смежности: Учеб. пособие. Элементы мягких вычислений. СПб.: СПбГУ; ООО Издательство «Анатолия», 2007. 40 с.

9. *Тулупьев А.Л.* Алгебраические байесовские сети: локальный логико-вероятностный вывод: Учеб. пособие. Элементы мягких вычислений. СПб.: СПбГУ; ООО Издательство «Анатолия», 2007. 80 с.

10. *Тулупьев А.Л.* Байесовские сети: логико-вероятностный вывод в циклах. СПб.: Изд-во С.-Петербургского ун-та, 2008. 140 с. (Элементы мягких вычислений.)

11. *Тулупьев А.Л.* Алгебраические байесовские сети: система операций локального логико-вероятностного вывода // Информационно-измерительные и управляющие системы. 2009. № 4. С. 41–44.

12. *Тулупьев А.Л.* Непротиворечивость оценок вероятностей в идеалах конъюнктов и дизъюнктов. // Вестник СПбГУ. Сер. 10. 2009. Вып. 2. С. 121–131.

13. *Тулупьев А.Л.* Апостериорные оценки вероятностей в идеале конъюнктов // Вестник СПбГУ. 2010. Серия 10. Вып. 1. С. 95–104.

14. *Тулупьев А.Л., Николенко С.И., Сироткин А.В.* Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 608 с.

15. *Тулупьев А.Л., Сироткин А.В.* Алгебраические байесовские сети: принцип декомпозиции и логико-вероятностный вывод в условиях неопределенности // Информационно-измерительные и управляющие системы. 2008. №10, т. 6. С. 85–87.

16. *Тулупьев А.Л., Сироткин А.В.* Локальный апостериорный вывод в алгебраических байесовских сетях как система матрично-векторных операций // Интегрированные модели и мягкие вычисления в искусственном интеллекте. V-я Международная научно-практическая конференция. Сборник научных трудов. В 2-х т. Т. 1. СПб.: Наука, 2009. С. 425–434.

17. *Тулупьев А.Л., Сироткин А.В., Николенко С.И.* Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах. СПб.: Изд-во С.-Петербургского ун-та, 2009. 400 с.

Сироткин Александр Владимирович — младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики Учреждения Российской академии наук С.-Петербургский институт информатики и автоматизации РАН (СПИИРАН). Область научных интересов: алгебраические байесовские сети: вычислительные аспекты логико-вероятностного вывода в условиях неопределенности, математические методы анализа генома. Число научных публикаций — 64. avs@iias.spb.su; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450. Научный руководитель — Тулупьев А.Л.

Alexander Vladimirovich Sirotkin — junior researcher, Theoretical and Interdisciplinary Computer Science Laboratory, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS).

Research interests: algebraic Bayesian networks, algorithms of probabilistic-logic inference under uncertainty. The number of publications — 64. avs@iias.spb.su, www.tulupyev.spb.ru; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

Тулупьев Александр Львович — д.ф.-м.н., доцент; заведующий лабораторией теоретических и междисциплинарных проблем информатики СПИИРАН, профессор кафедры информатики математико-механического факультета С.-Петербургского государственного университета (СПбГУ). Область научных интересов: представление и обработка данных и знаний с неопределенностью, применение методов математики и информатики в социокультурных исследованиях, применение методов биостатистики и математического моделирования в эпидемиологии, технология разработки программных комплексов с СУБД, методы автоматизированной оценки защищенности персонала информационных систем от социоинженерных атак. Число научных публикаций — 230. ALT@iias.spb.su, www.tulupyev.spb.ru; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450.

Alexander Lvovich Tulupyev — PhD in Computer Science, Dr. of Sc., Associate Professor; Head of Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS, Professor of Computer Science Department, SPbSU. Research area: uncertain data and knowledge representation and processing, mathematics and computer science applications in socio-cultural studies, biostatistics, simulation, and mathematical modeling applications in epidemiology, data intensive software systems development technology. Number of publications — 230. ALT@iias.spb.su, www.tulupyev.spb.ru; SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ, проект № 09-01-00861-а.

Рекомендовано ТиМПИ СПИИРАН, зав. лаб. Тулупьев А.Л., д.ф.-м.н., доцент.

Статья поступила в редакцию 08.07.2011.

РЕФЕРАТ

Сироткин А.В. **Вычислительная сложность алгоритмов локального апостериорного вывода в алгебраических байесовских сетях.**

Статья посвящена анализу вычислительной сложности алгоритмов локального апостериорного логико-вероятностного вывода. Алгоритмы апостериорного вывода позволяют моделировать рассуждения со свидетельствами в условиях неопределённости.

В статье описываются три существующих вида свидетельств: детерминированное, стохастическое и неточное. Даются необходимые определения, в том числе определения двух задач, возникающих в апостериорном выводе. Подробно разбирается процесс распространения детерминированных свидетельств, представляющих собой фиксированное означивание набора атомарных пропозиций. Апостериорный вывод в случае детерминированного свидетельства выражается на матрично-векторном языке. Описывается структура матриц, возникающих при распространении детерминированного свидетельства. Доказывается возможность разложения матрицы, соответствующей оператору ненормированного апостериорного вывода в случае фрагментов знаний со скалярными оценками, на матрицы малой размерности. На основе этих результатов приводится решение задач апостериорного вывода в случае детерминированного свидетельства и фрагментов знаний со скалярными и интервальными оценками.

Стохастические и неточные свидетельства рассматриваются как случайные элементы, принимающие значения детерминированных свидетельств с заданным распределением или семейством распределений вероятности. На основе описания решений задач для детерминированных свидетельств формируются описания решений задач для стохастических и неточных свидетельств, как оценка соответствующих математических ожиданий.

Кроме того, в статье приводится пример, который демонстрирует, что в общем случае для ФЗ с интервальными оценками, оставаясь в рамках ЗЛП и не накладывая дополнительных ограничений, невозможно построить точные оценки пропагации стохастического свидетельства.

На основе формализованных описаний решений указанных выше задач строятся и доказываются оценки сложности алгоритмов локального апостериорного вывода в алгебраических байесовских сетях.

SUMMARY

Sirotkin A.V. **Computational complexity of local posteriori inference algorithms in algebraic Bayesian networks.**

Abstract. The paper covers complexity estimates of local posteriori inference algorithms in algebraic Bayesian networks. A posteriori inference algorithms provide the way for decision making based on evidence upon uncertainty.

We describe three known kinds of evidence: deterministic, stochastic and imprecise. The necessary definitions are given. We demonstrate two types of tasks for posteriori inference. The deterministic evidence is a given assignment of some propositional variables. We carefully describe deterministic evidence propagation and provide description in the form of matrix and vector operations. We specify these matrices and prove that the matrix used in propagation can be represented as tensor product of small-size matrices. Based on this results, we present solution of deterministic evidence propagation task for both scalar and interval estimations in knowledge pattern.

We consider stochastic and imprecise evidence as probabilistic elements that take these values in form of deterministic evidence according to the given distribution or set of distributions. Based on deterministic evidence case, we provide description of solution problem in case of stochastic or imprecise evidence in form of estimation for expected value.

In addition, we provide example demonstrating that it is impossible in general case to build exact values of estimation for stochastic and imprecise evidence tasks using only linear programming problems and not assuming additional restrictions.

Basing on formal description of tasks mentioned above, we provide and prove a system of local posteriori inference algorithms complexity estimates.