Уравнения апостериорного вывода в фрагментах знаний над идеалом дизъюнктов[1]

Золотин А. А., аспирант мат.-мех. фак-та кафедры информатики СПбГУ, стажер-исследователь Санкт-Петербургского института информатики и автоматизации РАН (СПИИРАН), andrey.zolotin@gmail.com
Мальчевская Е. А., студентка мат.-мех. фак-та кафедры информатики СПбГУ, м.н.с. Санкт-Петербургского института информатики и автоматизации РАН (СПИИРАН), katerina.malch@gmail.com
Тулупьев А. Л., профессор кафедры информатики мат.-мех. фак-та СПбГУ, зав. лаб. ТиМПИ Санкт-Петербургского института информатики и автоматизации РАН (СПИИРАН), alexander.tulupyev@gmail.com

Аннотация

В данном докладе рассмотрены задачи локального апостериорного логико-вероятностного вывода в алгебраических сетях для фрагментов знаний, построенных над идеалом дизъюнктов. Для случая пропагации стохастического и неточного свидетельств во фрагмент знаний над идеалом дизъюнктов с интервальными оценками вероятности истинности построены задачи линейного программирования.

Введение

Алгебраические байесовские сети являются классом вероятностных графических моделей (ВГМ). Они позволяют обрабатывать данные с неопределенностью, которая может возникать как по причине необходимости перехода от высказываний на естественном языке к вероятностным оценкам, так и из-за ”пробелов” в данных, возникающих в следствие неполноты имеющейся информации. В отличие от других классов ВГМ, в частности байесовских сетей доверия, АБС позволяют работать не только с точечными (скалярными) оценками вероятности истинности, но и с интервальными оценками вероятности истинности.

Существует не только стандартное представление алгебраических байесовских сетей, но и альтернативные, использование которых в некоторых случаях могут быть удобней не только для экспертов, но и для обработки сети.

[1] Часть публикуемых материалов получена в рамках проекта, выполненного при финансовой поддержке Российского фонда фундаментальных исследований (грант № 15-01-09001-а).

Основным аппаратом для обработки имеющихся данных и получения новых знаний на основе имеющихся оценок является логико-вероятностный вывод. По уровню влияния на сеть ЛВВ делится на локальный и глобальный. Локальные операции являются базовыми, поскольку они используются в глобальных. По этой причине приоритетом исследований является развитие, формализация и программное представление именно операций локального логико-вероятностного вывода.

В локальный логико-вероятностный вывод входят поддержание непротиворечивости, решение задачи априорного вывода и двух задач апостериорного выводов. В некоторых случаях возникает необходимость построения и решения задачи линейного программирования для вычисления результата.

Цель данной статьи – построить задачи линейного программирования для ситуация пропагации стохастического и неточного свидетельств (апостериорного вывода) во фрагмент знаний с интервальными оценками вероятности истинности, построенного над идеалом дизъюнктов.

Теоретическая часть про АБС и ФЗ

Алгебраические байесовские сети (АБС) представляются ненаправленными графами с идеалами конъюнктов в узлах. Конъюнктам приписывается либо скалярная (точечная), либо интервальная оценка вероятности истинности. Назовем идеал конъюнктов с приписанными его элементам оценками вероятности истинности фрагментом знаний (ФЗ).

Как и другие пропозициональные формулы, идеал конъюнктов строится над некоторым фиксированным множеством атомов – алфавитом.

На Рис. 1 приведено графическое представление ФЗ, построенного над алфавитом $A = \{x_1, x_2, x_3\}$, состоящим из трех атомов x_1, x_2, x_3. Конъюнктам приписаны интервальные оценки вероятности истинности.

Помимо классического множества, над которым строится ФЗ, – идеала конъюнктов – в теории АБС рассматривают еще два: множество квантов и идеал дизъюнктов [4]. ФЗ, заданный над идеалом дизъюнктов приведен на Рис. 2.

Оценки вероятности истинности, приписанные элементам ФЗ, организуются в один или два вектора $\mathbf{P_c}$ и $\{\mathbf{P_c^-}, \mathbf{P_c^+}\}$ соответственно. Вектор $\mathbf{P_c}$ задается при скалярных оценках. Векторы $\{\mathbf{P_c^-}, \mathbf{P_c^+}\}$ задаются при интервальных оценках. В последнем случае в один из векторов ($\mathbf{P_c^-}$) помещаются нижние границы оценок, а в другой ($\mathbf{P_c^+}$) – верхние границы.

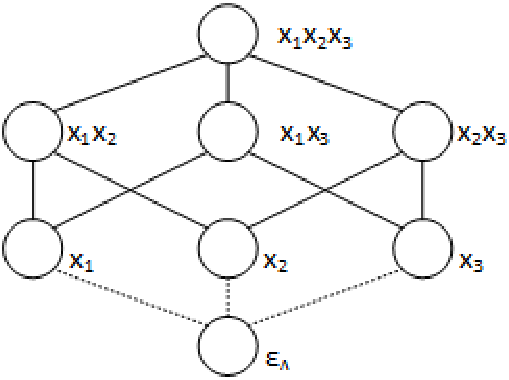

Рис. 1: Фрагмент знаний для идеала конъюнктов.

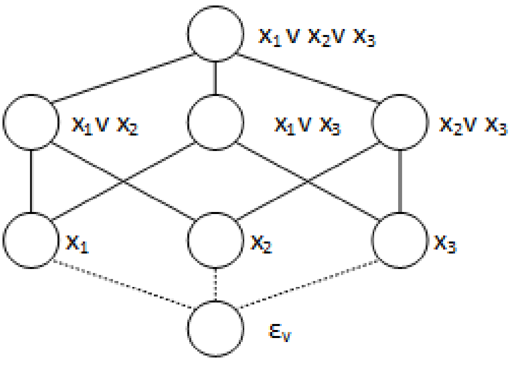

Рис. 2: Фрагмент знаний для идеала дизъюнктов.

$$\mathbf{P_c} = \begin{pmatrix} 1 \\ p(c_1) \\ \vdots \\ p(c_{2^n-1}) \end{pmatrix}$$

$$\mathbf{P_c^-} = \begin{pmatrix} 1 \\ p^-(c_1) \\ \vdots \\ p^-(c_{2^n-1}) \end{pmatrix}; \mathbf{P_c^+} = \begin{pmatrix} 1 \\ p^+(c_1) \\ \vdots \\ p^+(c_{2^n-1}) \end{pmatrix}$$

Для альтернативной модели фрагмента знаний, построенного над идеалом дизъюнктов, вероятности организуются в вектор $\mathbf{P_d}$ аналогичный по структуре случаю с конъюнктами:

$$\mathbf{P_d} = \begin{pmatrix} 0 \\ p(d_1) \\ \vdots \\ p(d_{2^n-1}) \end{pmatrix}$$

Рассмотрим ФЗ над идеалом дизъюнктов, заданный на алфавите $A = \{x_1, x_2, x_3\}$. Тогда вектор вероятностей будет выглядеть следующим образом:

$$\mathbf{P_d} = \begin{pmatrix} 0 \\ p(x_1) \\ p(x_2) \\ p(x_2 \vee x_1) \\ p(x_3) \\ p(x_3 \vee x_1) \\ p(x_3 \vee x_2) \\ p(x_3 \vee x_2 \vee x_1) \end{pmatrix}$$

В теории АБС выделяют два вида вывода: глобальный [1] и локальный[2]. В случае глобального вывода операции осуществляются над всей сетью и изменения в одном фрагменте знаний приводят к общим изменениям в сети. При локальном выводе мы рассматриваем и оперируем только с одним конкретным фрагментом знаний, не затрагивая при этом остальные.

Локальный логико-вероятностный вывод делится на следующие задачи: проверка и поддержание непротиворечивости ФЗ, задача априорного вывода и первая и вторая задачи апостериорного вывода.

Задачи апостериорного вывода

Предположим, что в ФЗ поступает новая информация (свидетельство), обуславливающая существующие оценки вероятностей элементов ФЗ. Как было указано ранее, отличительной особенностью АБС является возможность работы с неточными(интервальными) оценками вероятностей, что позволяет представить свидетельство как ФЗ со стохастическими оценками, так и с интервальными в зависимости от данных, над которыми данное свидетельство построено.

Теперь, зная структуру фрагмента знаний, над которым проводятся операции логико-вероятностного вывода, рассмотрим распространение по ФЗ (пропагацию) стохастического и неточного свидетельств. Так как каждая из компонент вектора вероятностей $\mathbf{P}_{\mathrm{d}}$ принимает значения из интервала (то есть $\mathbf{P}_{\mathrm{d}}^{-}[i] \leq \mathbf{P}_{\mathrm{d}}[i] \leq \mathbf{P}_{\mathrm{d}}^{+}[i]$) то решение первой задачи апостериорного вывода сводится к решению серии задач линейного программирования по нахождению максимума и минимума целевой функции при определенных ограничениях.

Стохастическое свидетельство

Сперва рассмотрим решение первой и второй задач апостериорного вывода для стохастического свидетельства. Пусть в ФЗ поступило свидетельство, представленное ФЗ над идеалом дизъюнктов с оценками вероятностей $\mathbf{P}_{\mathrm{d}}^{ev}$. Как было показано ранее [4] в работе, описывающей пропагацию детерминированного свидетельства, при работе с ФЗ над идеалом дизъюнктов оказывается удобнее оперировать вектором $\mathbf{P}_{\mathrm{d}}^{'}$, связанным с $\mathbf{P}_{\mathrm{d}}$ через соотношение $\mathbf{P}_{\mathrm{d}}^{'} = \mathbf{1} - \mathbf{P}_{\mathrm{d}}$. В той же работе было получено соотношение для решения первой и второй задач апостериорного вывода в случае скалярных оценок вероятностей в идеале дизъюнктов. Приведем их ниже:

$$p\langle i,j\rangle = (\mathbf{d}^{\langle i,j\rangle}, \mathbf{P}_{\mathrm{d}}^{'}) \text{ и } \mathbf{P}_{\mathrm{d}}^{'\,\langle i,j\rangle} = \frac{\mathbf{M}^{\langle i,j\rangle}\mathbf{P}_{\mathrm{d}}^{'}}{(\mathbf{d}^{\langle i,j\rangle}, \mathbf{P}_{\mathrm{d}}^{'})}, \tag{1}$$

где $\mathbf{d} = \bigotimes_{k=0}^{k=n-1} \widetilde{\mathbf{d}}_k^{\langle i,j\rangle}$, $\widetilde{\mathbf{d}}_k^{\langle i,j\rangle} = \begin{cases} \mathbf{d}^+, \text{ если } x_k \text{ входит в } c_i; \\ \mathbf{d}^-, \text{ если } x_k \text{ входит в } c_j; \\ \mathbf{d}^\circ, \text{ иначе}; \end{cases}$,

а $\mathbf{d}^+ = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mathbf{d}^- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\mathbf{d}^\circ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

и $\mathbf{M} = \bigotimes_{k=0}^{k=n-1} \widetilde{\mathbf{M}}_k^{\langle i,j\rangle}$, $\widetilde{\mathbf{M}}_k^{\langle i,j\rangle} = \begin{cases} \mathbf{M}^+, \text{ если } x_k \text{ входит в } c_i; \\ \mathbf{M}^-, \text{ если } x_k \text{ входит в } c_j; \\ \mathbf{M}^\circ, \text{ иначе}; \end{cases}$, причем

$\mathbf{M}^+ = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{M}^- = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$, $\mathbf{M}^\circ = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Как и в предыдущих работах по смежной тематике [1, 2, 3] воспользуемся представлением стохастического свидетельства в качестве серии детерминированных. В данном случае решениями задач апостериорного вывода будут накрывающие интервальные оценки вероятностей. Они получатся как линейная комбинация интервальных оценок, появляющихся при разложении стохастического свидетельства на детерминированные свидетельства и их последующей пропагации.

Стоит также отметить, что на вектор переменных $\mathbf{P}_\mathrm{d}^{'}$ накладываются линейные ограничения для поддержания непротиворечивости фрагмента знаний: ограничения, вытекающие из аксиоматики теории вероятностей ($\mathcal{E}$) и ограничения, вытекающие из предметной области ($\mathcal{D}$). Ограничение $\mathcal{E}$ в случае фрагмента знаний над идеалом дизъюнктов вытекает из уравнения, связывающего вектор вероятностей дизъюнктов и квантов: $\mathbf{L}_\mathrm{n}\mathbf{P}_\mathrm{d}^{'} = \mathbf{P}_\mathrm{q}$. Таким образом условие $\mathbf{P}_\mathrm{q} \geq 0$ можно трансформировать в условие на матрично-векторном языке: $\mathbf{L}_\mathrm{n}\mathbf{P}_\mathrm{d}^{'} \geq 0$. В свою очередь ограничение $\mathcal{D}$ может быть записано как совокупность условий: $\forall i \in [0..n-1] \mathbf{P}_\mathrm{d}^{'-}[i] \leq \mathbf{P}_\mathrm{d}^{'}[i] \leq \mathbf{P}_\mathrm{d}^{'+}[i]$, либо, как это принято в математических дисциплинах, посвященным экстремальным задачам, $\mathbf{P}_\mathrm{d}^{'-} \leq \mathbf{P}_\mathrm{d}^{'} \leq \mathbf{P}_\mathrm{d}^{'+}$, где n — это мощность множества квантов, над которым построен данный фрагмент знаний. То есть, множество ограничений $\mathcal{D} = \{\mathbf{P}_\mathrm{d}^{'-} \leq \mathbf{P}_\mathrm{d}^{'} \leq \mathbf{P}_\mathrm{d}^{'+}\}$. Таким образом, решением первой задачи апостериорного вывода является замкнутый промежуток, ограниченный минимумом и максимумом следующего выражения при условии ограничений $\mathcal{D} \cup \mathcal{E}$:

$$\sum_{i=0}^{2^{n'}-1} (\mathbf{d}^{\langle \mathrm{GInd}(i,m), \mathrm{GInd}(2^{n'}-1-i,m)\rangle}, \mathbf{P}_\mathrm{d}^{'})\mathbf{P}_\mathrm{d}^{'\,ev}[i] \qquad (2)$$

Так как в данном фрагменте знаний заданы интервальные оценки вероятности истинности дизъюнктов, то мы имеем дело с семейством распределений вероятностей; следовательно, искомые во второй подзадаче апостериорные вероятности могут быть оценены лишь интервально. Поиск верхней и нижней границы ${\mathbf{P}_{\mathrm{d}}^{'}}^{\langle i,j\rangle}$ сводится к поиску минимума и максимума выражения ${\mathbf{P}_{\mathrm{d}}^{'}}^{\langle i,j\rangle} = \sum_{i=0}^{2^{n'}-1} \frac{\mathbf{M}^{\langle \mathrm{GInd}(i,m),\mathrm{GInd}(2^{n'}-1-i,m)\rangle}\mathbf{P}_{\mathrm{d}}^{'}}{(\mathbf{d}^{\langle \mathrm{GInd}(i,m),\mathrm{GInd}(2^{n'}-1-i,m)\rangle},\mathbf{P}_{\mathrm{d}}^{'})}{\mathbf{P}_{\mathrm{d}}^{'}}^{ev}[i]$ при ограничениях $\mathcal{D}\cup\mathcal{E}$. Стоит отметить, что здесь и далее задача о поиске минимума и максимума решается независимо и отдельно для каждой целевой функции – компоненты вектора ${\mathbf{P}_{\mathrm{d}}^{'}}^{\langle i,j\rangle}$ — относительно переменных вектора $\mathbf{P}_{\mathrm{d}}^{'}$.

Экстремальные задачи по нахождению минимума и максимума указанной целевой функции относятся к известному классу задач дробно-линейного программирования. Для сведения их к задаче линейного программирования нам нужно избавиться от множителя $\frac{1}{(\mathbf{d}^{\langle \mathrm{GInd}(i,m),\mathrm{GInd}(2^{n'}-1-i,m)\rangle},\mathbf{P}_{\mathrm{d}}^{'})}$. Воспользовавшись тем же приемом и обоснованием, что и в [9, 20, 37], зададим новый вектор переменных (и, таким образом, произведем замену переменных) $\mathbf{D} = \lambda\mathbf{P}_{\mathrm{d}}^{'}$, тогда $\frac{\mathbf{M}^{\langle i,j\rangle}\mathbf{P}_{\mathrm{d}}^{'}}{(\mathbf{d}^{\langle i,j\rangle},\mathbf{P}_{\mathrm{d}}^{'})} = \frac{\mathbf{M}^{\langle i,j\rangle}\lambda\mathbf{P}_{\mathrm{d}}^{'}}{(\mathbf{d}^{\langle i,j\rangle},\lambda\mathbf{P}_{\mathrm{d}}^{'})} = \frac{\mathbf{M}^{\langle i,j\rangle}\mathbf{D}}{(\mathbf{d}^{\langle i,j\rangle},\mathbf{D})}$. Теперь положим $\frac{1}{(\mathbf{d}^{\langle i,j\rangle},\mathbf{D})} = \mathbf{1}$. При новых переменных и введенных ограничениях исходные экстремальные задачи сведутся к задачам линейного программирования — поиску минимума и максимума компонент вектора $\mathbf{M}^{\langle i,j\rangle}\mathbf{P}_{\mathrm{d}}^{'}$ при условиях $\mathcal{R} = \mathcal{D}^{'}\cup\mathcal{E}^{'}\cup(\mathbf{d}^{\langle i,j\rangle},\mathbf{D}) = \mathbf{1}$, причем $\mathcal{D}^{'} = \{\lambda{\mathbf{P}_{\mathrm{d}}^{'}}^{+} \leq \mathbf{D}\}\cup\{\lambda{\mathbf{P}_{\mathrm{d}}^{'}}^{-} \geq \mathbf{D}\}, \mathcal{E}^{'} = \{\mathcal{D} \geq 0\}\cup\{\lambda \geq 0\}$. Исходя из вышеизложенного, окончательное решение второй задачи апостериорного вывода можно будет записать в виде интервала, ограниченного минимумом и максимумом следующего выражения при ограничении $\mathcal{R}$:

$$\sum_{i=0}^{2^{n'}-1} \mathbf{M}^{\langle \mathrm{GInd}(i,m),\mathrm{GInd}(2^{n'}-1-i,m)\rangle}\mathbf{D}{\mathbf{P}_{\mathrm{d}}^{'}}^{ev}[i] \qquad (3)$$

Неточное свидетельство

Однако, как правило стохастическая оценка вероятности истинности является лишь случаем, рассматриваемым в теории и редко встречается на практике, поэтому теперь предположим, что в фрагмент знаний, указанный ранее, поступило неточное свидетельство, выраженное фрагментом знаний с интервальными оценками вероятности истинности элементов $[\mathbf{P}_{\mathrm{d}}^{ev,-}, \mathbf{P}_{\mathrm{d}}^{ev,+}]$.

Рассмотрим интервальное свидетельство как множество всех стохастических свидетельств, оценки которых лежат в заданном интервале вероятностей $[\mathbf{P}_{\mathrm{d}}^{ev,-}, \mathbf{P}_{\mathrm{d}}^{ev,+}]$.

В данном случае решением первой и второй задач апостериорного вывода будут накрывающие интервальные оценки вероятностей [10]. К сожалению, при решении первой задачи апостериорного вывода в случае фрагмента знаний над идеалом дизъюнктов с интервальными оценками и неточных свидетельств (как и в случае стохастических) мы сталкиваемся с выходом за пределы задач линейного программирования; решение же последних приводит лишь к накрывающим оценкам [9, 37]. Следуя преобразованиям, описанным для случае стохастического свидетельства, получим, что решением первой задачи будет интервал, ограниченный следующими значениями:

$$\min_{\mathbf{P}_{\mathrm{d}}^{\prime\, ev,-} \leq \mathbf{P}_{\mathrm{d}}^{\prime\, ev} \leq \mathbf{P}_{\mathrm{d}}^{\prime\, ev,+}} F_{first}(\mathbf{P}_{\mathrm{d}}^{\prime\, ev}, \mathbf{P}_{\mathrm{d}}^{\prime}) \text{ и } \max_{\mathbf{P}_{\mathrm{d}}^{\prime\, ev,-} \leq \mathbf{P}_{\mathrm{d}}^{\prime\, ev} \leq \mathbf{P}_{\mathrm{d}}^{\prime\, ev,+}} F_{first}(\mathbf{P}_{\mathrm{d}}^{\prime\, ev}, \mathbf{P}_{\mathrm{d}}^{\prime}), \tag{4}$$

где

$$F_{first}(\mathbf{P}_{\mathrm{d}}^{\prime\, ev}, \mathbf{P}_{\mathrm{d}}^{\prime}) = \sum_{i=0}^{2^{n^{\prime}}-1} \min_{\mathcal{D} \cup \mathcal{E}} (\mathbf{d}^{\langle \mathrm{GInd}(i,m), \mathrm{GInd}(2^{n^{\prime}}-1-i,m) \rangle}, \mathbf{P}_{\mathrm{d}}^{\prime}) \mathbf{P}_{\mathrm{d}}^{\prime\, ev}[i]$$

Решение второй задачи можно найти воспользовавшись следующими формулами:

$$\min_{\mathbf{P}_{\mathrm{d}}^{\prime\, ev,-} \leq \mathbf{P}_{\mathrm{d}}^{\prime\, ev} \leq \mathbf{P}_{\mathrm{d}}^{\prime\, ev,+}} F_{sec}(\mathbf{P}_{\mathrm{d}}^{\prime\, ev}, \mathbf{P}_{\mathrm{d}}^{\prime}) \text{ и } \max_{\mathbf{P}_{\mathrm{d}}^{\prime\, ev,-} \leq \mathbf{P}_{\mathrm{d}}^{\prime\, ev} \leq \mathbf{P}_{\mathrm{d}}^{\prime\, ev,+}} F_{sec}(\mathbf{P}_{\mathrm{d}}^{\prime\, ev}, \mathbf{P}_{\mathrm{d}}^{\prime}), \tag{5}$$

где

$$F_{sec}(\mathbf{P}_{\mathrm{d}}^{\prime\, ev}, \mathbf{D}) = \sum_{i=0}^{2^{n^{\prime}}-1} \min_{\mathcal{D} \cup \mathcal{E}} (\mathbf{M}^{\langle \mathrm{GInd}(i,m), \mathrm{GInd}(2^{n^{\prime}}-1-i,m) \rangle}, \mathbf{D}) \mathbf{P}_{\mathrm{d}}^{\prime\, ev}[i],$$

а ограничение $\mathcal{R}$ является множеством ограничений, описанных при рассмотрении пропагации стохастического свидетельства.

Заключение

В работе рассмотрены случаи пропагации стохастического и неточного свидетельств во фрагмент знаний с интервальными оценками вероятности над идеалом дизъюнктов. На основе полученных ранее матрично-векторных уравнений построены задачи линейного программирования для решения первой и второй задач апостериорного вывода в обоих случаях. Результаты, полученные в рамках данной работы создают задел для развития программного комплекса, реализующего указанные алгоритмы, за счет возможности использования подходящих структур и их свойств, а также ясности и законченности сформулированных задач. Кроме того, результаты, полученные в работе для локальных структур АБС могут быть в дальнейшем переиспользованы в рамках работ по развитию алгоритмов глобального вывода.

Литература

[1] Тулупьев А. Алгебраические байесовские сети: глобальный логико-вероятностный вывод в деревьях смежности: Учеб. пособие. СПб.: СПбГУ; ООО Издательство «Анатолия», 2007. 40 с. (Сер. Элементы мягких вычислений).

[2] Тулупьев А. Алгебраические байесовские сети: локальный логико-вероятностный вывод: Учеб. пособие. СПб.: СПбГУ; ООО Издательство «Анатолия», 2007. 80 с. (Сер. Элементы мягких вычислений).

[3] Тулупьев А., Николенко С., Сироткин А. Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.

[4] Золотин А., Мальчевская Е. Матрично-векторные алгоритмы локального апостериорного вывода в алгебраических байесовских сетях над идеалами дизъюнктов. XIX Международная конференция по мягким вычислениям и измерениям (SCM-2016). Сборник докладов в 2-х томах. Санкт-Петербург. 2016. Т.1. 79-82с.