Поддержание согласованности в альтернативных фрагментах знаний алгебраических байесовских сетей [1]

Пак С. Т. И. В., студент кафедры информатики СПбГУ, стажёр лаборатории ТиМПИ СПИИРАН DeifyTheGod@gmail.com

Данная статья посвящена поддержанию согласованности в альтернативных фрагментах знаний в алгебраических байесовских сетях. Также рассмотрена реализация поддержания согласованности в альтернативных фрагментах знаний над идеалами дизъюнктов и множествами квантов на основе комплекса программ AlgBNModeller.

Введение

Существует множество вероятностно-графических моделей (ВГМ). Среди них Марковские сети (МС), байесовские сети доверия (БСД) и множество других [1, 2]. Объектом данного исследования являются, в частности, алгебраические байесовские сети (АБС).

АБС, как один из классов ВГМ, состоят из переменных и связей между ними. Такие сети можно декомпозировать, получив набор фрагментов знаний (ФЗ), которые часто являются идеалами конъюнктов над множеством подалфавитов некоторого алфавита с заданными для них оценками [3]. Чаще всего АБС используются для того, чтобы сделать понятным для компьютера видение эксперта некоторой области знаний [4]. Оценки могут быть как скалярными, так и интервальными (последние возникают, чаще всего, при недостаточной информированности системы [4, 5]).

При рассмотрении такой системы особый интерес вызывает поддержание её согласованности при поступлении в систему новых данных. Для этих целей были разработаны два вида логико-вероятностного вывода (ЛВВ): априорный и апостериорный [6, 8]. Необходимость априорного вывода возникает, когда требуется оценить вероятность истинности некоторой формулы для заданной АБС. Необходимость апостериорного вывода возникает при поступлении в систему свидетельств извне. Оба вида вывода могут быть как локальными, то есть работающими над конкретным ФЗ, так и глобальными — над всей АБС [4].

[1] Статья содержит материалы исследований, частично поддержанных грантом РФФИ 15-01-09001 — «Комбинированный логико-вероятностный графический подход к представлению и обработке систем знаний с неопределенностью: алгебраические байесовские сети и родственные модели».

Для идеалов конъюнктов есть уравнения на матрично-векторном языке, которые описывают оба вида ЛВВ. Однако, можно рассматривать и другие структуры в основе ФЗ: например, можно заменить идеал конъюнктов на идеал дизъюнктов или множество квантов [6]. Данные структуры тоже достаточно хорошо изучены, и цель данной работы — представить уравнения локального логико-вероятностного вывода для ФЗ с указанной структурой, а также показать пример их реализации на примере комплекса программ AlgBNModeller.

Фрагменты знаний и логико-вероятностный вывод

АБС, как и любая ВГМ, является математической моделью, представлющей собой набор данных с заданными вероятностями и взаимосвязями.

Вообще говоря, в теории АБС рассматриваются ФЗ над идеалами конъюнктов, но с оговоркой, что данная модель допускает альтернативные ФЗ, построенные над иными множествами пропозициональных формул [7]. Иногда такие ФЗ могут оказаться более удобными для рассмотрения (в силу особенностей задачи, для решения которой используется АБС).

В данной статье рассматривается локальный логико-вероятностный вывод.

Априорный вывод, как следует из названия, использует предшествующий опыт (то есть данные в ФЗ) для синтеза согласованных оценок поступающих на вход пропозиций над алфавитом ФЗ.

Апостериорный вывод, напротив, использует поступающие свидетельства (новые, «обуславливающие» данные) для решения двух важных задач:

1. Какова оценка вероятности поступающего свидетельства при условиях данного ФЗ?

2. Как изменятся оценки вероятностей данного ФЗ при условии поступающего свидетельства?

Достаточно богатая теория была разработана, помимо ФЗ над идеалами конъюнктов, также для ФЗ над множествами квантов и идеалами дизъюнктов [4–8].

В случае с множествами квантов большим подспорьем является [5], в которой рассматриваются ФЗ, построенные над ними.

В случае с идеалами дизъюнктов можно произвести ряд операций, которые сведут задачи ЛВВ к аналогичным задачам для ФЗ над множеством квантов. Данные операции подробно описаны в [4].

Решение задачи априорного вывода в альтернативных ФЗ

Когда мы работаем с квантами, нам удобно представлять их вероятностные оценки в виде вектора

$$\mathbf{P}_q = \begin{pmatrix} p(q_1) \\ \vdots \\ p(q_{2^n}) \end{pmatrix},$$

где n - мощность алфавита, над которым построен данный ФЗ. При выборе такого представления встаёт вопрос о нумерации квантов. Возможным решением является следующее [5]: алфавит, над которым построен ФЗ, нумеруется, а номер кванта вычисляется таким образом: элемент алфавита с номером t, получивший положительное означивание, добавляет к номеру кванта 2^t, а получивший отрицательное означивание - нуль.

Для дизъюнктов дело обстоит несколько иначе: их нумерация аналогична нумерации конъюнктов, а вектор оценок вероятностей выглядит как

$$\mathbf{P}_d = \begin{pmatrix} p(\epsilon_\vee) \\ p(x_1) \\ \vdots \\ p(x_n \vee x_{n-1} \vee \ldots \vee x_1) \end{pmatrix},$$

причём $p(\epsilon_\vee) = 0$. Как было показанко в [4],

$$\overline{\mathbf{P}_d} = \mathbb{1} - \mathbf{P}_d = \begin{pmatrix} p(\epsilon_\wedge) \\ p(\overline{x_1}) \\ \vdots \\ p(\overline{x_n} \wedge \overline{x_{n-1}} \wedge \ldots \wedge \overline{x_1}) \end{pmatrix}.$$

Тогда, взяв вектор $\overline{\mathbf{P}_d}$ и умножив матрицу

$$\mathbf{L}_m = \overline{\overline{\mathbf{E}}}_m I_m, \text{ где } \overline{\overline{\mathbf{E}}}_m = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{[m]}, \mathbf{I}_m = \begin{pmatrix} 1 & \text{-1} \\ 0 & 1 \end{pmatrix}^{[m]},$$

на него, можно получить вектор вероятностей квантов данного ФЗ.

Таким образом, реализовав необходимые матрицы, можно свести задачу априорного вывода для ФЗ над идеалом дизъюнктов к задаче априорного вывода для ФЗ над множеством квантов, что значительно проще.

Пример программной реализации

Поддержание согласованности в альтернативных ФЗ было реализовано в рамках комплекса программ «AlgBNModeller». Данный комплекс программ уже предлагал реализацию поддержания согласованности для ФЗ над идеалами конъюнктов в виде библиотек Inferrer и Propagator, потому было принято решение расширять данные библиотеки.

Был написан ряд классов, которые схожи по поведению с аналогичными для конъюнктов. В частности, были реализованы классы для решения задачи априорного вывода во фрагментах знаний с бинарными, скалярными и интервальными оценками вероятностей, построенными над идеалами дизъюнктов и множествами квантов.

В случае квантов всё оказывается весьма тривиально: проверка непротиворечивости сводится к проверке того, что сумма вероятностей всех квантов равна единице, а все вероятности находятся на отрезке от нуля до единицы. Также достаточно легко решается проблема априорного вывода: нужная нам пропозиция выражается в совершенной дизъюнктивной нормальной форме (сокращённо СДНФ), после чего вероятности квантов, входящих в данную СДНФ, суммируются.

В случае дизъюнктов мы можем перевести [4] вектор вероятностей дизъюнктов в вектор вероятностей соответствующих квантов. Тогда вектор дизъюнктов будет непротиворечив и согласован лишь в том случае, если будет непротиворечив и согласован построенный вектор квантов [4]. Тогда поступающие задачи априорного вывода мы можем решать на данном векторе квантов, опираясь на то, что это те же оценки вероятности, только записанные в другой форме. Данное утверждение корректно, так как во всех операциях применялись только линейные переходы. Очевидно, что тогда решение задачи априорного вывода сведётся к аналогичной для квантов.

Так как реализовывать хранение данных матриц «в лоб» достаточно трудно (в виду растущего экспоненциально размера последних), они были реализованы процедурно. Неоценимым здесь является тот факт, что единственное, от чего зависят данные матрицы — это мощность алфавита ФЗ, для которого такая матрица понадобилась. Можно получить функцию, которая по номеру строки и столбца будет выдавать число, стоящее в такой матрице [6]. Пользуясь данными результатами можно хранить лишь процедуры получения данных чисел (так как данные числа ограничены в своём разнообразии нулём и единицами разных знаков, можно даже не бояться переполнения при подсчёте данного числа).

Далее приведена реализация матрицы $\mathbf{L}_m$:

```
public class DQTransformationMatrix : TransformationMatrix
  {
      public DQTransformationMatrix(int order) : base(
          order)
      {
      }

      public override int getElementValue(int i, int j)
      {
          var k = getDimension() - i - 1;
          if ((k & j) != k)
              return 0;
          return (Util.BitCount(j ^ k) & 1) == 1 ? -1 : 1;
      }
  }
```

Такое название позволяет легко понять назначение данной матрицы даже человеку, плохо запомнившему наименования матриц преобразования: эта матрица, при умножении на неё вектора оценок вероятностей дизъюнктов $\mathbf{D}$ трансформирует его в вектор оценок вероятностей квантов $\mathbf{Q}$.

Литература

[1] Cowell, Robert G. Probabilistic networks and expert systems. Berlin: Springer, 1999. 81 p.

[2] Pearl J. Probabilistic Reasoning in Intelligent Systems. 2nd revised. San Mateo, CA: Morgan Kaufmann, 1988. 552 p.

[3] Городецкий В. И. Байесовский вывод. Препринт № 149. Л.: ЛИИАН, 1991. 38 с.

[4] Тулупьев А. Л. Алгебраические байесовские сети: логико-вероятностная графическая модель баз фрагментов знаний с неопределённостью: дисс. ... д-р. физ.-мат. наук. СПб.: СПбГУ, 2009. 670 с.

[5] Сироткин А. В. Алгебраические байесовские сети: вычислительнаая сложность алгоритмов логико-вероятностного вывода в условиях неопределённости: дисс. ... канд. физ.-мат. наук. ССПб.: СПбГУ, 2011. 218 с.

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

[7] Тулупьев А. Л., Сироткин А. В., Николенко С. И. Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах. СПб.: Издательство С.- Петербургского университета, 2009. 400 с. (Элементы мягких вычислений).

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