УДК 004.8

М.П. Момзикова, О.И. Великодная, М.Я. Пинский,
А.В. Сироткин, А.Л. Тулупьев, А.А. Фильченков

ПРЕДСТАВЛЕНИЕ БИНАРНЫХ ЛИНЕЙНЫХ ПО СТРУКТУРЕ СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ В ВИДЕ АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЕЙ

Момзикова М.П., Великодная О.И., Пинский М.Я., Сироткин А.В., Тулупьев А.Л., Фильченков А.А. **Представление бинарных линейных по структуре скрытых марковских моделей в виде алгебраических байесовских сетей.**

Аннотация. Для моделирования различных процессов в таких областях, как распознавание речи, теория информации, машинный перевод, молекулярная биология, широко используются вероятностно-графические модели в том числе скрытые марковские модели и байесовские сети. Цель данной работы — исследовать взаимосвязь между скрытой марковской моделью и алгебраической байесовской сетью. Предложен алгоритм представления бинарной линейной по структуре скрытых марковских моделей в виде алгебраических байесовских сетей. Доказана теорема о совпадении вероятностных семантик скрытых марковских моделей и алгебраических байесовских сетей.

Ключевые слова: скрытые марковские модели, алгебраические байесовские сети, бинарные линейные по структуре СММ, вероятностно-графические модели.

Momzikova M.P., Velikodnaya O.I., Pinsky M.I., Sirotkin A.V., Tulupyev A.L., Filchenkov A.A. **Representation of binary linear hidden Markov models in the form of algebraic Bayesian networks.**

Abstract. Probabilistic graphical models including hidden Markov models and Bayesian networks are widespread in process modeling in such fields as speech recognition, information theory, machine translation and molecular biology. The goal of this work is toresearch of mutual relations between hidden Markov models and algebraic Bayesian networks. An algorithm to design a binary linear hidden Markov models as an algebraic Bayesian networks is suggested. The theorem about coincidence between probabilistic semantics is proven.

Keywords:hidden Markov models, algebraic Bayesian networks, binary linear hidden Markov models, probabilistic graphical models

1. Введение.Для моделирования различных процессов в таких областях, как распознавание речи, теория информации, машинный перевод, молекулярная биология, широко используются вероятностно-графические модели[1, 4–7]. Поскольку механизмы работы и область применения моделей одни и те же (взаимоотношение событий и вероятностей), то большую роль в их развитии играют исследования, посвященные представлению связей между их различными вариантами. Теоретические и алгоритмические успехи одних моделей можно применять к другим. Целью данной работы является исследование взаимоотношений между такими разновидностями вероятностно-графических моделей, как скрытая марковская модель (СММ) и алгебраическая байесовская сеть (АБС).

Скрытые марковские модели — инструмент для моделирования информации временного ряда. СММ используются во многих современных системах распознавания речи, в большинстве приложений вычислительной молекулярной биологии, в сжатии информации, в системах статистического машинного перевода, приложениях компьютерного зрения.

Алгебраическая байесовская сеть — это одна из математических моделей баз фрагментов знаний (ФЗ) с неопределенностью. Она формализует знания (с неопределенностью) эксперта с помощью вероятностной логики. Развитие аппарата АБС осуществлялось с начала 1980-х гг., и на сегодняшний день в теории АБС существуют алгоритмы для решения различных задач, однако АБС все еще редко используются для практических целей[3].

Ранее было доказано[8], что СММ сводятся к байесовским сетям доверия (БСД), а БСД, в свою очередь, сводятся к АБС. Однако алгоритм непосредственного сведения СММ к АБС не был ни формализован, ни исследован. В данной работе будет изложен алгоритм представления бинарной линейной по структуре СММ в виде АБС. Благодаря такому представлению появляется возможность решать задачи для скрытых марковских моделей посредством алгоритмов логико-вероятностного вывода для алгебраических байесовских сетей, а, кроме того, решать задачи, сформулированные в терминах СММ, с помощью методов теории АБС.

Кроме представления СММ в виде АБС, в данной статье будет приведено доказательство корректности этого представления с помощью теоремы о совпадении вероятностных семантик СММ и АБС, полученной из этой СММ.

2. Теоретические основы. Изложение теории СММ будет вестись по работам[4, 6, 8], а теории АБС — по[2–3].

Скрытая марковская модель (СММ) — модель, включающая в себя набор наблюдений, набор скрытых состояний и удовлетворяющая следующим свойствам:

1. $P\left(q_{t+1} \middle| q_t, q_{t-1}, q_{t-2}, \ldots, q_1\right) = P\left(q_{t+1} \middle| q_t\right)$ —*марковское свойство*.

2. $P\left(o_t \middle| o_1, \ldots o_T, q_1, \ldots, q_T\right) = P\left(o_t \middle| q_t\right)$ — зависимость текущего наблюдения только от текущего состояния,где

$O=\{o_1,o_2,\ldots o_T\}$ — последовательность наблюдений, $Q=\{q_1,q_2,\ldots q_T\}$ — последовательность состояний.

Скрытая марковская модель состоит из следующих объектов:

1. Набор возможных значений скрытых состояний: $S=\{s_1,s_2,\ldots s_N\}$;
2. Последовательность скрытых состояний во времени: $Q=\{q_1,q_2,\ldots q_T\}$;
3. Матрица переходных вероятностей: $A=\{a_{ij}\}$, $a_{ij}=P\left(q_{t+1}=s_j \middle| q_t=s_i\right)$, $1\le i,j\le N$;
4. Вектор начального распределения: $\pi=\{\pi_i\}$, $\pi_i=P(q_1=s_i)$, $1\le i\le N$;
5. Алфавит возможных значений наблюдений: $V=\{v_1,v_2,\ldots v_M\}$;
6. Последовательность наблюдений во времени: $O=\{o_1,o_2,\ldots o_T\}$;
7. Матрица вероятностей наблюдений: $B=\{b_j(k)\}$, $b_j(k)=P\left(v_k=o_t \middle| s_j=q_t\right)$, $1\le j\le N$, $1\le k\le M$.

В расчетах, связанных со СММ, последние представляются в виде набора матриц вероятностей: $\mu=(A,B,\pi)$.

В теории СММ сформулированы три основных задачи:

1. *Правдоподобие наблюдений.*

Дана СММ с известными матрицами вероятностей. Определить вероятность поступающей последовательности наблюдений во времени относительно этой СММ.

Формальная постановка задачи: даны последовательность наблюдений $O=\{o_1,o_2,\ldots o_T\}$ и модель $\mu=(A,B,\pi)$. $P\left(O\middle|\mu\right)=?$

Для данной задачи существуют различные решения, например, алгоритм «вперед–назад».

2. *Декодирование скрытой последовательности.*

Даны СММ с оценками вероятности и поступившая последовательность наблюдений. Требуется определить наиболее вероятную последовательность скрытых состояний.

Формальная постановка задачи: даны последовательность наблюдений $O = \{o_1, o_2, \ldots o_T\}$ и модель $\mu = (A, B, \pi)$. Найти наиболее вероятную последовательность скрытых состояний $Q = \{q_1, q_2, \ldots q_T\}$, соответствующую данной последовательности наблюдений.

Данная задача решается с помощью алгоритма Витерби.

3. *Обучение СММ.*

Изменить (настроить) матрицы вероятностей СММ так, чтобы максимизировать вероятности поступающего набора последовательностей наблюдений. То есть требуется обучить СММ на наборе тренировочных последовательностей наблюдений.

Формальная постановка задачи: настроить параметры модели $\mu = (A, B, \pi)$ так, чтобы максимизировать $P(O|\mu)$.

Данную задачу можно решать с помощью алгоритма Баума–Вэлха.

Для преобразования в АБС в данной работе будем использовать бинарные линейные по структуре СММ. Это такие СММ, у которых могут быть только два скрытых состояния $(S = \{s_1, s_2\})$ и два вида наблюдений $(V = \{v_1, v_2\})$. В последнем будет иметь место $S = V = \{0,1\} = \{true,\ false\}$.

Алгебраическая байесовская сеть — математическая модель базы фрагментов знаний с неопределенностью[1]. Фрагмент знаний представляется в виде идеала конъюнктов с оценками их истинности.

Пусть T — конечный набор элементарных пропозиций, $S = \{s_1, s_2, \ldots, s_k\}$ — непустое подмножество T, $S^{\Diamond} = \{v_1 v_2 \ldots v_r : v_1, v_2 \ldots v_r \in S, r = 0 \ldots k\}$ — конъюнкты (множество положительно означенных конъюнкций над S); $S^{\Delta} = S^{\Diamond} \setminus \{e\}$ — идеал конъюнкта (множество положительно означенных конъюнктов без пустого конъюнкта). В идеале существует один максимальный элемент (максимальная по длине конъюнкция) и множество минимальных элементов (одноатомных конъюнкций).

Квант $Q = \{\tilde{x}_0 \tilde{x}_1 \ldots \tilde{x}_{n-1}\}$ — конъюнкция, которая для любой атомарной переменной из алфавита содержит либо ее формулу, либо отрицание.

Нумерация конъюнктов и квантов. Каждому конъюнкту из идеала $\left\{x_{i_1}x_{i_2}\ldots x_{i_k} \,\middle|\, 0 \le i_1 < \ldots < i_k \le n-1, k \le n\right\}$ можно сопоставить номер вида $2^{i_1} + 2^{i_2} + \ldots + 2^{i_k}$. Обозначим через c_i конъюнкт с порядковым номером i.

Выделим из кванта положительную часть. Номер получившегося конъюнкта будет номером кванта. Обозначим через q_i квант с порядковым номером i.

Вероятности конъюнктов и квантов. Вероятности, относящиеся к фрагменту знаний, удобно упорядочивать по номерам конъюнктов и квантов и представлять в виде векторов:

$$\mathbf{P_c} = \begin{pmatrix} 1 \\ p(c_1) \\ \vdots \\ p(c_{2^n-1}) \end{pmatrix}; \ \mathbf{P_q} = \begin{pmatrix} p(q_0) \\ p(q_1) \\ \vdots \\ p(q_{2^n-1}) \end{pmatrix}$$

Выделяют две структуры АБС: первичную и вторичную.

Первичная структура АБС — это база ФЗ.

Вторичная структура АБС — это связи между ФЗ. Она определяется набором вершин и набором ребер между ними. Вершины — это ФЗ, они однозначно задаются глобальным индексом минимального конъюнкта. Ребра однозначно задаются двумя вершинами.

Виды свидетельств. Свидетельства делятся на детерминированные и недетерминированные, среди последних выделяют стохастические.

Детерминированным свидетельством называют свидетельство, сообщающее о том, какой набор атомов означен положительно, а какой — отрицательно.

Стохастическим называется свидетельство, представимое в виде ФЗ с точечными оценками.

АБС позволяют производить два вида выводов: априорный и апостериорный. Задача априорного вывода в ФЗ: по известным вероятностным оценкам истинности заданных пропозициональных формул построить вероятностную оценку формулы, не вошедшей в число заданных.

Первая задача апостериорного вывода в ФЗ: оценить вероятности появления свидетельства при заданных оценках.

Вторая задача апостериорного вывода в ФЗ: оценить условные вероятности всех конъюнктов АБС относительно поступившего свидетельства.

3. Представление бинарных линейных по структуре СММ в виде АБС. Начнем с примера. Рассмотрим простейшую линейную бинарную СММ в четыре момента времени (рис. 1).

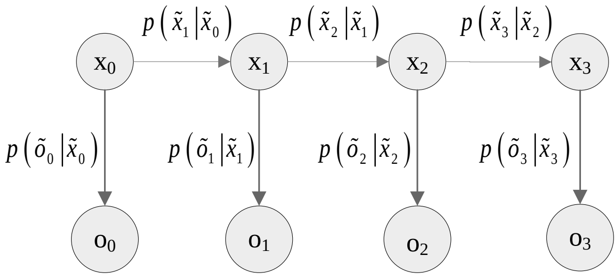

Рис. 1. Линейная бинарная СММ в четыре момента времени.

Матрицы данной модели $\mu=\left(A,B,\pi\right)$ будут иметь следующий вид:

$$A=\begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix}=\begin{pmatrix} p(x_i|x_i) & p(x_i|\overline{x}_i) \\ p(\overline{x}_i|x_i) & p(\overline{x}_i|\overline{x}_i) \end{pmatrix}$$

$$B=\begin{pmatrix} b_0(v_0) & b_0(v_1) \\ b_1(v_0) & b_1(v_1) \end{pmatrix}=\begin{pmatrix} p(o_i|x_i) & p(\overline{o}_i|x_i) \\ p(o_i|\overline{x}_i) & p(\overline{o}_i|\overline{x}_i) \end{pmatrix}$$

$$\pi=\begin{pmatrix} \pi_0 \\ \pi_1 \end{pmatrix}=\begin{pmatrix} p(x_i) \\ p(\overline{x}_i) \end{pmatrix}$$

Теперь построим АБС, соответствующую рассмотренной СММ (рис. 2).

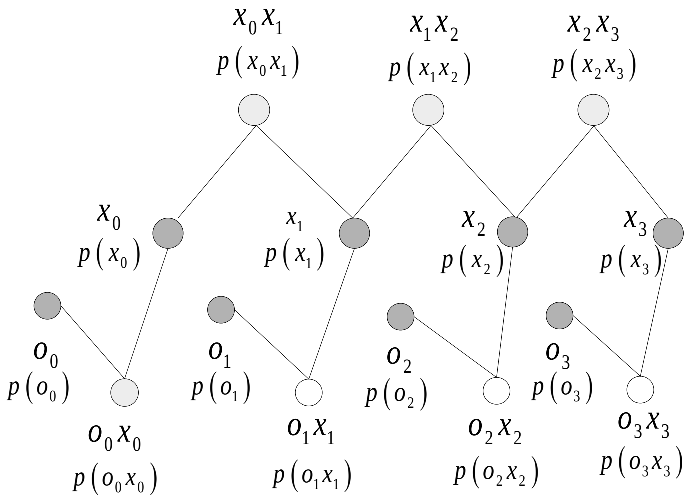

Рис. 2. АБС, соответствующая рассмотренной СММ.
Серым отмечены узлы, соответствующие одноатомным конъюнктам, белым — соответствующие двухатомным конъюнктам.

Для нумерации вершин необходимо ввести алфавит. В алфавите соответствующей АБС будем чередовать o_i и x_i, начиная с $i=0$: $\{o_0,x_0,o_1,x_1,\ldots o_{N-1},x_{N-1}\}$.

Проиндексируем алфавит следующим образом:

o_0 : $2^0=1$; x_0 : $2^1=2$; o_1 : $2^2=4$; x_1: $2^3=8$; o_2 : $2^4=16$; x_2 : $2^5=32$; o_3: $2^6=64$; x_3: $2^7=128$.

Тогда глобальные индексы ФЗ будут следующими (в соответствии с правилами нумерации квантов и конъюнктов):

o_0x_0: $2^0+2^1=3$

x_0x_1: $2^1+2^3=10$

o_1x_1: $2^2+2^3=12$

x_1x_2: $2^3+2^5=40$

o_2x_2: $2^4+2^5=48$

x_2x_3: $2^5+2^7=160$

$o_3x_3:\ 2^6+2^7=192$

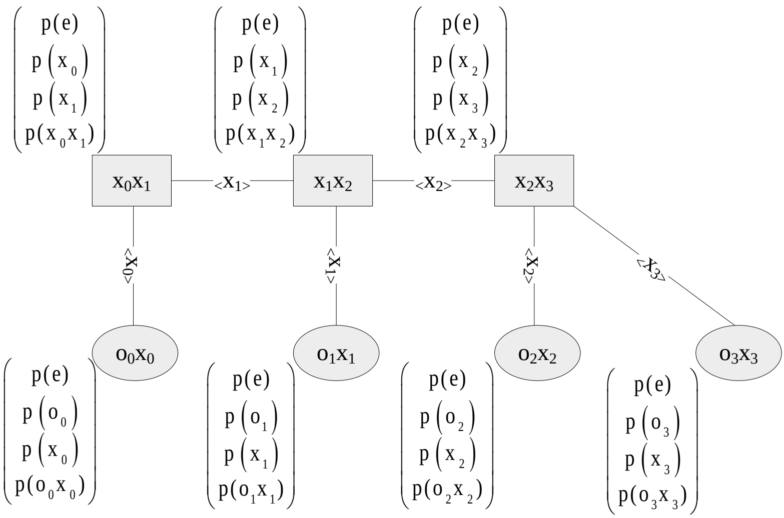

Рис. 3. Соответствующая АБС, изображенная как база ФЗ, изображены узлы и векторы вероятностей, хранящиеся в них.

Данная АБС будет содержать ФЗ (см. рис. 3) вида $\left\{\left(\left\{o_i,x_i\right\}^{\Delta},\left\{x_i,x_{i+1}\right\}^{\Delta}\right)_{i=0}^{N-2},\left\{o_{N-1},x_{N-1}\right\}^{\Delta}\right\}$ (рис. 3) и соответствующие вероятностные вектора

$$\left\{\left(\begin{pmatrix} p(e) \\ p(o_i) \\ p(x_i) \\ p(o_i x_i) \end{pmatrix},\begin{pmatrix} p(e) \\ p(x_i) \\ p(x_{i+1}) \\ p(x_i x_{i+1}) \end{pmatrix}\right)_{i=0}^{N-2},\begin{pmatrix} p(e) \\ p(o_{N-1}) \\ p(x_{N-1}) \\ p(o_{N-1}x_{N-1}) \end{pmatrix}\right\}.$$

Вершинами вторичной структуры будут ФЗ со своими глобальными индексами, ребра будут иметь вид

$$\left\{\left(\left(o_i x_i, x_i x_{i+1}\right),\left(x_i x_{i+1}, o_{i+1}x_{i+1}\right)_{i=0}^{N-2}\right),\left(o_{N-1},x_{N-1}\right)\right\}.$$

Для преобразования матриц вероятностей СММ в векторы вероятностей для ФЗ АБС, будем двигаться по набору ФЗ в АБС в следующем порядке: o_0x_0, x_0x_1, o_1x_1, x_1x_2, o_2x_2, x_2x_3, o_3x_3, и в порядке продвижения заполнять соответствующие вероятностные вектора.

1) Рассмотрим ФЗ o_0x_0:

$$p(e)=1;$$
$$p(o_0)=p(o_0x_0)+p(o_0\overline{x}_0)=$$
$$=p(o_0|x_0)\cdot p(x_0)+p(o_0|\overline{x}_0)\cdot p(\overline{x}_0)=$$
$$=b_{x_0}(o_0)\cdot\pi(x_0)+b_{\overline{x}_0}(o_0)\cdot\pi(\overline{x}_0);$$
$$p(x_0)=\pi(0);$$
$$p(o_0x_0)=p(o_0|x_0)\cdot p(x_0).$$

Так как x_0 — начальное состояние по времени, поэтому $p(x_0)=\pi(x_0)$, $p(x_0)$ — маргинальная вероятность.

2) Рассмотрим ФЗ x_0x_1:

$$p(e)=1;$$
$$p(x_0)=\pi(0);$$
$$p(x_1)=p(x_1|x_0)\cdot p(x_0)+p(x_1|\overline{x}_0)\cdot p(\overline{x}_0)=$$
$$=a_{x_0x_1}\cdot\pi(x_0)+a_{\overline{x}_0x_1}\cdot\pi(\overline{x}_0);$$
$$p(x_0x_1)=p(x_1|x_0)\cdot p(x_0)=a_{x_0x_1}\cdot\pi(x_0).$$

3) Рассмотрим o_1x_1. Вероятность $p(x_1)$ была вычислена на предыдущем шаге:

$$p(e)=1;$$
$$p(o_1)=p(o_1x_1)+p(o_1\overline{x}_1)=$$
$$=p(o_1|x_1)\cdot p(x_1)+p(o_1|\overline{x}_1)\cdot p(\overline{x}_1)=$$
$$=b_{x_1}(o_1)\cdot p(x_1)+b_{\overline{x}_1}(o_1)\cdot p(\overline{x}_1);$$
$$p(x_1);$$
$$p(o_1x_1)=b_{x_1}(o_1)\cdot p(x_1).$$

4) Рассмотрим x_1x_2:

$$p(e)=1;$$
$$p(x_1);$$

$$p(x_2) = p(x_2|x_1)\cdot p(x_1) + p(x_2|\overline{x}_1)\cdot p(\overline{x}_1) =$$
$$= a_{x_1x_2}\cdot p(x_1) + a_{\overline{x}_1x_2}\cdot p(\overline{x}_1);$$
$$p(x_1x_2) = p(x_2|x_1)\cdot p(x_1) = a_{x_1x_2}\cdot p(x_1).$$

Далее для всех временных шагов i для ФЗ вида o_ix_i вероятностные векторы будут иметь вид

$$\begin{pmatrix} p(e) \\ p(o_i) \\ p(x_i) \\ p(o_ix_i) \end{pmatrix} = \begin{pmatrix} 1 \\ b_{x_i}(o_i)\cdot p(x_i) + b_{\overline{x}_i}(o_i)\cdot p(\overline{x}_i) \\ p(x_i) \\ b_{x_i}(o_i)\cdot p(x_i) \end{pmatrix}.$$

Для ФЗ вида x_ix_{i+1} векторы будут иметь вид

$$\begin{pmatrix} p(e) \\ p(x_i) \\ p(x_{i+1}) \\ p(x_ix_{i+1}) \end{pmatrix} = \begin{pmatrix} 1 \\ p(x_i) \\ a_{x_ix_{i+1}}\cdot p(x_i) + a_{\overline{x}_ix_{i+1}}\cdot p(\overline{x}_i) \\ a_{x_ix_{i+1}}\cdot p(x_i) \end{pmatrix}.$$

Замечание: Каждая $p(x_i)$ в этих векторах будет вычислена на шаге $i-1$ в векторе $\begin{pmatrix} p(e) \\ p(x_{i-1}) \\ p(x_i) \\ p(x_{i-1}x_i) \end{pmatrix}$.

Будем рассматривать события: $\tilde{o}_0\tilde{x}_0\tilde{o}_1\tilde{x}_1\ldots\tilde{o}_{n-1}\tilde{x}_{n-1}$, пространство элементарных событий: $\Omega = \{\tilde{o}_0\tilde{x}_0\tilde{o}_1\tilde{x}_1\ldots\tilde{o}_{n-1}\tilde{x}_{n-1}\}$ (множество всех возможных означиваний), вероятностное пространство: $\left(\Omega,\ 2^{\Omega},\ p\right)$, вероятностную семантику (распределение): $\left\{p\left(\tilde{o}_0\tilde{x}_0\tilde{o}_1\tilde{x}_1\ldots\tilde{o}_{n-1}\tilde{x}_{n-1}\right)\right\}$.

Теорема о совпадении вероятностных семантик данных СММ и АБС.Вероятностные семантики СММ и АБС, полученной из этой скрытой марковской модели, совпадают.

Пример: рассмотрим СММ при времени $T=1$. Тогда событие, определенное временем, выглядит следующим образом: $\tilde{o}_0\tilde{x}_0$. Вектор:

$\begin{pmatrix} p\left(\bar{o}_0\bar{x}_0\right) \\ p\left(o_0\bar{x}_0\right) \\ p\left(\bar{o}_0 x_0\right) \\ p\left(o_0 x_0\right) \end{pmatrix}$ будет являться вероятностным распределением или вероятностной семантикой. То есть в каждой строке вектора находится вероятность конкретно-означенной временной последовательности. Для СММ эти вероятности будут считаться одним способом, а для АБС — другим. Нужно доказать, что вектора для СММ и АБС совпадают.

Доказательство. По определению условной вероятности:

$$\begin{aligned} & p\left(\tilde{o}_0\tilde{x}_0\tilde{o}_1\tilde{x}_1\ldots\tilde{o}_{n-1}\tilde{x}_{n-1}\right) = p\left(\tilde{o}_{n-1}\tilde{x}_{n-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right) = \\ & = p\left(\tilde{o}_{n-1}\left|\tilde{x}_{n-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right.\right)\cdot p\left(\tilde{x}_{n-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right) = \\ & = p\left(\tilde{o}_{n-1}\left|\tilde{x}_{n-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right.\right)\cdot p\left(\tilde{x}_{n-1}\left|\tilde{o}_{n-2}\tilde{x}_{n-2}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right.\right)\cdot \\ & \cdot p\left(\tilde{o}_{n-2}\tilde{x}_{n-2}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right) = \ldots = \\ & = p\left(\tilde{o}_{n-1}\left|\tilde{x}_{n-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right.\right)\cdot p\left(\tilde{x}_{n-1}\left|\tilde{o}_{n-2}\tilde{x}_{n-2}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right.\right)\cdot \\ & \cdot p\left(\tilde{o}_{n-2}\left|\tilde{x}_{n-2}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right.\right)\ldots\cdot p\left(\tilde{o}_0\left|\tilde{x}_0\right.\right)\cdot p\left(\tilde{x}_0\right) \end{aligned}$$

Получили:

$$\begin{aligned} & p\left(\tilde{o}_{n-1}\tilde{x}_{n-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right) = \\ & = p\left(\tilde{o}_{n-1}\left|\tilde{x}_{n-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right.\right)\cdot p\left(\tilde{x}_{n-1}\left|\tilde{o}_{n-2}\tilde{x}_{n-2}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right.\right)\cdot\ldots \\ & \ldots\cdot p\left(\tilde{o}_0\left|\tilde{x}_0\right.\right)\cdot p\left(\tilde{x}_0\right) \end{aligned} \quad (1)$$

I. Для СММ воспользуемся свойствами условной независимости, которые были даны в определении:

1. Марковское свойство
$$P\left(x_{t+1}\left|x_t, x_{t-1}, x_{t-2},\ldots, x_1\right.\right) = P\left(x_{t+1}\left|x_t\right.\right);$$
2. Зависимость текущего наблюдения только от текущего состояния (по времени):
$$P\left(o_t\left|o_0,\ldots o_T, x_0,\ldots, x_T\right.\right) = P\left(o_t\left|x_t\right.\right),$$

где $O = \{o_0, o_1,\ldots o_T\}$ — последовательность наблюдений, $X = \{x_0, x_1,\ldots x_T\}$ — последовательность состояний.

Тогда

$$p\left(\tilde{o}_{n-1}\tilde{x}_{n-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)=$$
$$=p\left(\tilde{o}_{n-1}\middle|\tilde{x}_{n-1}\right)\cdot p\left(\tilde{x}_{n-1}\middle|\tilde{x}_{n-2}\right)\cdot\ldots\cdot p\left(\tilde{x}_1\middle|\tilde{x}_0\right)\cdot p\left(\tilde{o}_0\middle|\tilde{x}_0\right)\cdot p\left(\tilde{x}_0\right)$$

II. В формуле (1) рассмотрим множители типа $p\left(\tilde{x}_k\middle|\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)$, $p\left(\tilde{o}_k\middle|\tilde{x}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)$. И, пользуясь вновь определением условной вероятности и свойством условной независимости ФЗ, распишем их:

$$p\left(\tilde{x}_k\middle|\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)\underset{усл.в.}{=}\frac{p\left(\tilde{x}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}{p\left(\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}\underset{усл.в.}{=}$$
$$=\frac{p\left(\tilde{x}_k\tilde{o}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\middle|\tilde{x}_{k-1}\right)\cdot p\left(\tilde{x}_{k-1}\right)}{p\left(\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}\underset{у.нез.}{=}$$
$$=\frac{p\left(\tilde{x}_k\middle|\tilde{x}_{k-1}\right)\cdot p\left(\tilde{o}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\middle|\tilde{x}_{k-1}\right)\cdot p\left(\tilde{x}_{k-1}\right)}{p\left(\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}\underset{усл.в.}{=}$$
$$=\frac{p\left(\tilde{x}_k\middle|\tilde{x}_{k-1}\right)\cdot p\left(\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}{p\left(\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}=p\left(\tilde{x}_k\middle|\tilde{x}_{k-1}\right)$$

$$p\left(\tilde{o}_k\middle|\tilde{x}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)\underset{усл.в.}{=}\frac{p\left(\tilde{o}_k\tilde{x}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}{p\left(\tilde{x}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}\underset{усл.в.}{=}$$
$$=\frac{p\left(\tilde{o}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\middle|\tilde{x}_k\right)\cdot p\left(\tilde{x}_k\right)}{p\left(\tilde{x}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}\underset{у.нез.}{=}$$
$$=\frac{p\left(\tilde{o}_k\middle|\tilde{x}_k\right)\cdot p\left(\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\middle|\tilde{x}_k\right)\cdot p\left(\tilde{x}_k\right)}{p\left(\tilde{x}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}\underset{усл.в.}{=}$$
$$=\frac{p\left(\tilde{o}_k\middle|\tilde{x}_k\right)\cdot p\left(\tilde{x}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}{p\left(\tilde{x}_k\tilde{o}_{k-1}\tilde{x}_{k-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)}=p\left(\tilde{o}_k\middle|\tilde{x}_k\right)$$

Получили, что у данной АБС, как и у данной СММ,

$$p\left(\tilde{o}_{n-1}\tilde{x}_{n-1}\ldots\tilde{o}_1\tilde{x}_1\tilde{o}_0\tilde{x}_0\right)=$$
$$=p\left(\tilde{o}_{n-1}\middle|\tilde{x}_{n-1}\right)\cdot p\left(\tilde{x}_{n-1}\middle|\tilde{x}_{n-2}\right)\cdot\ldots\cdot p\left(\tilde{x}_1\middle|\tilde{x}_0\right)\cdot p\left(\tilde{o}_0\middle|\tilde{x}_0\right)\cdot p\left(\tilde{x}_0\right),$$

что и требовалось доказать.

4.Заключение. Известно, что скрытые марковские модели (СММ) могут быть представлены как частный случай динамических байесовских сетей доверия (БСД), которые, в свою очередь, могут быть преобразованы в алгебраические байесовские сети (АБС). Отсюда возник естественный вопрос о связи СММ и АБС.

В данной работе была рассмотрена одна из разновидностей СММ — бинарная линейная по структуре СММ (такая СММ, у которой могут быть только два скрытых состояния и два вида наблюдений). Для остальных разновидностей СММ алгоритмыих преобразования в АБС исследуются, но еще не получили своего описания.

Для указанного выше вида СММ был разработан алгоритм для построенияее представления в виде АБС, а также доказана теорема о совпадении вероятностных семантик соответствующих СММ и АБС, подтверждающая корректность данного алгоритма.

Таким образом, в данной статье был изложен способ представления одной из разновидностей СММ в виде АБС, что может позволить решать типичные задачи для данного вида СММ с помощью алгоритмов теории АБС.

Для использования полученного теоретического результата в практических целях необходимо проведение дальнейших исследований. Чтобы решать задачи, сформулированные для СММ, с помощью АБС, необходимо представить классические задачи теории СММ в виде задач логико-вероятностного вывода в теории АБС. Также необходимо отметить, что предложенный в данной работе алгоритм носит частный характер, и одним из направлений дальнейших исследований может являться поэтапное обобщение алгоритма преобразования СММ в АБС на все разновидности СММ.

Литература

1. *Николенко С.И., Тулупьев А Л.* Самообучающиеся системы. М.: МНЦМО, 2009. 288 с.
2. *Тулупьев А.Л.* Алгебраические байесовские сети. Логико-вероятностный подход к моделированию баз знаний с неопределенностью. СПб.: СПИИРАН, 2000. 292 с.
3. *Тулупьев А.В., Николенко С.И., Сироткин А.В.* Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006.
4. *Cowell R. G., Dawid A. P., Lauritzen S. L., Spiegelhalter D. J.* Probabilistic Networks and Expert Systems. NY.: Springer-Verlag, 1999.
5. *Huang X., Acero A., Hsiao-Wuen Hon* Spoken Language Processing. Prentice Hall, 2001. ISBN 0-13-022616-5.
6. *Huang X., Jack M., and Ariki Y.* Hidden Markov Models for Speech Recognition. Edinburgh University Press, 1990. ISBN 0748601627
7. *Jurafsky D., Martin J.H.* Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics. 2nd edition. Prentice-Hall, 2009.
8. *Stengel M.*Introduction to Graphical Models, Hidden Markov Models and Bayesian Networks. Department of Information and Computer Sciences Toyohashi University of Technology Toyohashi, 441-8580 Japan, 2003.

Момзикова Мария Петровна — студент философского факультета С.-Петербургского государственного университета (СПбГУ). Область научных интересов: вероятностно-графические модели в автоматизированных информационных системах. Число научных публикаций — 1. masya.momz@gmail.com, СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450. Научныйруководитель — А.Л. Тулупьев.

Momzikova Maria Petrovna — student of Faculty of Philosophy ,SPbSU. Research area: probabilistic graphical models in automated informational systems. The number of publications — 1. masya.momz@gmail.com, SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450. Scientificadvisor — A.L. Tulupyev.

Великодная Ольга Игоревна — студент кафедры компьютерных технологий факультета информационных технологий и программирования СПб ГУ ИТМО. velikola@yandex.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. Научныйруководитель — А.Л. Тулупьев.

Velikodnaya Olga Igorevna — student of Computer Technologies Department, SPbIT-MO.velikola@yandex.ru, SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450. Scientificadvisor — A.L. Tulupyev.

Пинский Михаил Яковлевич — студент кафедры компьютерных технологий факультета информационных технологий и программирования СПб ГУ ИТМО. mikhailpinsky@gmail.com, СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450. Научныйруководитель — А.Л. Тулупьев.

Pinsky Mikhail Iakovlevich — student of Computer Technologies Department, SPbITMO. mikhailpinsky@gmail.com, SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450. Scientificadvisor — A.L. Tulupyev.

Сироткин Александр Владимирович — младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики СПИИРАН. Область научных интересов: алгебраические байесовские сети: вычислительные аспекты логико-вероятностного вывода в условиях неопределенности. Число научных публикаций — 40. avs@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.

Sirotkin Alexander Vladimirovich — junior researcher, Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS.Research interests: algebraic Bayesian networks, algorithms of probabilistic-logic inference under uncertainty. The number of publications — 40.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, доцент 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. Число научных публикаций — 210. 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.

Tulupyev Alexander Lvovich — PhD in Computer Science, Dr. of Sc..Associate Professor; Head of Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS, Associate 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 — 210. 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.

Фильченков Андрей Александрович — аспирант кафедры информатики математико-механического факультета С.-Петербургского государственного университета (СПбГУ), младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики СПИИРАН. Область научных интересов: автоматическое обучение вероятностных графических моделей. Число научных публикаций — 6. aaafil@mail.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. Научныйруководитель — А.Л. Тулупьев.

Filchenkov AndreyAlexandrovich — PhD student of Computer Science Department, SPbGU, junior researcher, Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS Research area: machine learning of probabilistic graphical models. The number of publications — 6. aaafil@mail.ru, SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450. Scientificadvisor — A.L. Tulupyev.

Поддержка исследования. Работа выполнена при финансовой поддержке РФФИ, проект № **09-01-00861**-а «Методология построения интеллектуальных систем поддержки принятия решений на основе баз фрагментов знаний с вероятностной неопределенностью».

РЕФЕРАТ

Момзикова М.П., Великодная О.И., Пинский М.Я., Сироткин А.В., Тулупьев А.Л., Фильченков А.А. **Представление бинарных линейных по структуре скрытых марковских моделей в виде алгебраических байесовских сетей.**

Для моделирования различных процессов в таких областях, как распознавание речи, теория информации, машинный перевод, молекулярная биология, широко используются вероятностно-графические модели в том числе скрытые марковские модели и байесовские сети. Цель работы — исследование взаимоотношений между такими разновидностями вероятностно-графических моделей, как скрытая марковская модель и алгебраическая байесовская сеть.

В работе была обоснована необходимость проведения исследований в указанном направлении, были приведены области практического использования скрытых марковских моделей и алгебраических байесовских сетей. В теоретическом введении были изложены основы теории скрытых марковских моделей иалгебраических байесовских сетей, включая основные задачи, сформулированные для этих разновидностей вероятностно-графических моделей, необходимые для понимания основного содержания работы.

Был реализован следующий пример: преобразование бинарной линейной по структуре скрытой марковской модели в четыре момента времени в алгебраическую байесовскую сеть, построена диаграмма Хассе полученной алгебраической байесовской сети, представлены соответствующие иллюстрации.

По результатам примера был формализован вид полученной алгебраической байесовской сети, представлен алгоритм заполнения соответствующих вероятностных векторов для примера бинарной линейной по структуре скрытой марковской модели в четыре момента времени. Далее алгоритм заполнения вероятностных векторов алгебраической байесовской сети был формализован для общего случая алгебраической байесовской сети, полученных в ходе преобразования бинарных линейных по структуре скрытых марковских моделей.

Для обоснования корректности приведенного алгоритма преобразования бинарной линейной по структуре скрытой марковской модели в алгебраическую байесовскую сеть была доказана теорема о совпадении вероятностных семантик соответствующих моделей.

Таким образом, в данной статье был изложен способ представления одной из разновидностей скрытых марковских моделей в виде алгебраических байесовских сетей, что может позволить решать типичные задачи для данного вида скрытых марковских моделей с помощью алгоритмов теории алгебраических байесовских сетей.

Однако необходимо отметить, что для использования полученного теоретического результата в практических целях требуется проведение дальнейших исследований. Для того чтобы решать задачи, сформулированные для скрытых марковских моделей, с помощью алгебраических байесовских сетей, необходимо представить классические задачи теории скрытых марковских моделей в виде задач логико-вероятностного вывода в теории алгебраических байесовских сетей. Также необходимо отметить, что предложенный в данной работе алгоритм носит частный характер, и одним из направлений дальнейших исследований может являться поэтапное обобщение алгоритма преобразования скрытой марковской модели в алгебраическую байесовскую сеть на все разновидности скрытых марковских моделей.

SUMMARY

Momzikova M.P., Velikodnaya O.I., Pinsky M.I., Sirotkin A.V., Tulupyev A.L., Filchenkov A.A. **Representation of binary linear hidden Markov models in the form of algebraic Bayesian networks.**

Probabilistic graphical models including hidden Markov models and Bayesian networks are widespread in process modeling in such fields as speech recognition, information theory, machine translation and molecular biology. The goal of this work is to research mutual relations between such varieties of probabilistic graphic models as hidden Markov models and algebraic Bayesian networks.

A necessity ofresearches for the specified direction has been proved in the work. Also there were given areas of practical use of hidden Markov models. In theoretical introduction bases of theories of algebraic Bayesian networks and hidden Markov models were given. The main objects of hidden Markov model theory and algebraic Bayesian network theory were formulated. It is useful to mention that the onlynecessary minimum of the theories was given.

The following example was implemented: binary linear hidden Markov model transformation during four moments of time into an algebraic Bayesian network. Also the Hasse diagram of the obtained network was constructed, corresponding illustrations were presented.

According results of an example the kind of the obtained algebraic Bayesian network was formalized. An algorithm to fill corresponding algebraic Bayesian network probabilistic vectors was presented. Further the algorithm of filling probabilistic vectors was formalized for the algebraic Bayesian network general case which was received during binary linear hidden Markov model transformation.

For a substantiation of a correctness of the resulted algorithm of binary linear hidden Markov model transformation into algebraic Bayesian network, the theorem of coincidence of probabilistic semantics of corresponding hidden Markov models and algebraic Bayesian networks was proved.

As a result the way of representation of one of hidden Markov models varieties in the form of algebraic Bayesian networks was stated in this article. This representation presumes to solve typical problems for this variety of hidden Markov models by means of algebraic Bayesian network theory algorithm.

However it is necessary to notice that for use of the obtained theoretical result in the practical purposes carrying out of the further researches is required. To solve the problems formulated for hidden Markov models by means of algebraic Bayesian network theory it is necessary to present classical problems of hidden Markov model theory in the form of a logical probabilistic conclusion problems in algebraic Bayesian network theory. Also it is necessary to notice that the algorithm offered in this work has particular character, and one of directions of the further researches the generalization of algorithm of transformation hidden Markov model into algebraic Bayesian network for all thehidden Markov models varieties can be.