ИЗМЕРИМЫЕ МОДЕЛИ ТЕОРИИ ПЕРВОГО ПОРЯДКА В РАССТОЯНИЯХ НА ВЫСКАЗЫВАНИЯХ И ВЕРОЯТНОСТЯХ НА ЗНАНИЯХ*

Александр Викентьев

***Аннотация**: При анализе знаний, заданных в виде высказываний экспертов, для различия содержащейся в них информации и группирования их по похожести, возникает необходимость введения расстояния между высказываниями экспертов и меры опровержимости (информативности) высказываний экспертов. Этой проблемой занимались Загоруйко Н.Г., Лбов Г.С., Викентьев А.А. [1-4]. Вводим расстояние на измеримых формулах с использованием не более чем счетных моделей некоторой, заранее фиксированной теории T языка первого порядка. Такой подход является естественным при изучении некоторой конкретной прикладной проблемы, (поскольку тогда расстояние и информативность не будут искажены моделями, не относящимися к изучаемой аксиоматизированной области знаний) заданной например, некоторыми аксиомами-связями между переменными в ней, далее - теорией. Работа проделана в рамках проекта РФФИ 07-01-00331.*

***Keywords** : базы знаний,высказывания экспертов, теория моделй, метрика.*

***ACM Classification Keywords**: I.2.6. Artificial Intelligence - knowledge acquisition.*

Введение

К настоящему времени достаточно хорошо развиты теория и методы построения решающих функций распознавания образов на основе анализа эмпирической информации, представленной в виде таблиц данных. Параллельно этому проявляется все больший интерес к построению решающих функций на основе анализа экспертной информации, заданной в виде логических «знаний» нескольких экспертов, которые представлены формулами языка первого порядка [2-4]. При этом возникает задача введения расстояния на таких «знаниях» и определения мер опровержимости, информативности и вероятностей этих «знаний» [1-2] с учетом имеющейся теории T и ее моделей.

При решении задач распознавания образов, кластерного и регрессионного анализа важную роль играет информация, полученная от экспертов. В трудноформализуемых областях исследований особую важность приобретают методы обработки эмпирической информации, представленной списком экспертных логических «знаний» (эти «знания» могут быть частично или полностью противоречивы) [1], но совместны с имеющейся теорией T. Очевидно, такие высказывания («знания») могут различаться по количеству содержащейся в них информативности относительно T.

Расстояние на высказываниях экспертов и его свойства

Рассмотрим сигнатуру $\Omega = \left\{ P_1^{n_1}, \ldots, P_f^{n_f} \right\}$, состоящую из конечного числа предикатных символов от конечного числа переменных, которые выбираются для записи и изучения имеющихся связей между

* Работа выполнена при поддержке РРФИ 07-01-00331 и Интеграция НГУ.

переменными в конкретной прикладной области. Пусть выбрано некоторое исходное множество переменных $X=\{x_1,\dots,x_k\}$, среди которых присутствуют переменные разных типов. Обозначим через D_{x_j} множество возможных значений переменной x_j. Пусть $A_n=\bigcup_{j=1}^{k} D_{x_j}$ - непустое множество мощности не превосходящей некоторого числа n, являющееся объединением всех значений рассматриваемых переменных, включенных в предикаты.

Предполагаем, что в Ω для каждой переменной x_j включен сорт - одноместный предикат P_{x_j}, определенный только на области значений переменной x_j в A_n, то есть предикат $P_{x_j}(a)$ истинен в модели A_n на элементах $a\in A_n \Leftrightarrow a\in D_{x_j}$.

Во многих прикладных задачах значения переменных и их число конечны, и поэтому рассматриваются конечные множествам, но мы рассмотрим и более общий измеримый произвольный случай (модели могут быть и бесконечными: произвольной мощности n).

Определение1.[5,6]. Под интерпретацией будем понимать отображение γ, ставящее в соответствие каждому n_i- местному предикатному символу $P_i^{n_i}$ из сигнатуры Ω n_i- местный предикат (отношение) $P_i^{A_n}\subseteq A_n^{n_i}$, заданный на множестве A_n. Это позволяет говорить о модели $\langle A_n,\Omega\rangle$ сигнатуры Ω. В модели $\langle A_n,\Omega\rangle$ истинен предикат $P(x_1,\dots,x_k)$ на элементах $a_1,\dots a_k$ из A_n (записывается $\langle A_n,\Omega\rangle \models P(a_1,\dots,a_k)$) тогда и только тогда, когда $\langle a_1,\dots a_k\rangle\in P^{A_n}$ и $a_j\in D_{x_j}$. Ясно, что достаточно рассматривать модели только конечной сигнатуры, хотя результаты верны и в общей ситуации.

Пусть имеется конечное число s экспертов и области возможных значений всех переменных. Модели (в смысле теории моделей) задаются самими экспертами с условием , что они модели Т. Каждый эксперт задает свою модель. Каждый j-ый эксперт задает свою интерпретацию каждого предикатного символа $P_i^{n_i}$ сигнатуры Ω соответствующим отношением (предикатом) на множестве A_n. В результате имеем множество моделей $\{M_j\}_{j=1}^{s}$.

Считаем, что «Знания» экспертов можно записать в виде формул языка первого порядка, а формулы определяют формульные предикаты (подмножества) в каждой модели M_j по заданной интерпретации j-го эксперта [4].

Пусть F- система подмножеств множества $\mathrm{A}=\bigcup_k A_n^k$, где $A_n^k=\underbrace{A_n\times\ldots\times A_n}_{k\ раз}$, образующая σ-алгебру. Нас будут интересовать только такие подмножества S_j из F, для которых найдется формула ψ_j, отражающая «знания» экспертов, которая и определяет это подмножество S_j (формульное подмножество). То есть S_j - это множество кортежей из A, на которых выполняется формула ψ_j.

Формула ψ_j либо отражает какое-то из «знаний» экспертов, либо является их булевой комбинацией. В дальнейшем каждому рассматриваемому нами множеству S_j соответствует некоторая формула ψ_j.

Ω -множество всех используемых экспертами предикатных символов. Пусть B --замыкание множества Ω относительно логических операций $\neg, \vee, \wedge, \rightarrow$ и кванторов $\forall$ и $\exists$ по переменным. Ясно, что интересующее нас множество формул в B содержится.

Определение2. [8]. Вероятностной мерой μ на множестве B называется отображение $\mu : B \rightarrow [0,1]$, удовлетворяющее для ϕ и $\psi \in B$ условиям:

1) если |- $\phi \equiv \psi$, то $\mu(\phi) = \mu(\psi)$;

2) если |- ϕ, то $\mu(\phi) = 1$;

3) если |- $\neg\phi$, то $\mu(\phi) = 0$;

4) $\mu(\neg\phi) = 1 - \mu(\phi)$;

5) если |- $\neg(\phi \wedge \psi)$, то $\mu(\phi \vee \psi) = \mu(\phi) + \mu(\psi)$.

Далее всегда предполагаем, что на множестве формул B задана вероятностная мера μ. Тем самым, вероятностная мера μ задана на множествах из F.

По заданным «знаниям» экспертов мы построили класс моделей (исходный класс). Чтобы более полно использовать информацию экспертов расширим исходный класс моделей теории Т. Учитывая одновременно информацию нескольких экспертов, можно доуточнить (расширить или сузить) каждую модель и эти новые модели теории Т добавить в исходный класс. Доуточнять модели можно следующим образом. Вместо «знания», заданного экспертом в виде предиката P_i в модели M_j, далее будем рассматривать его доуточнение - предикат $\widetilde{P}_i$.

Под доуточнением $\widetilde{P}_i$ предиката P_i понимаем другую интерпретацию этого предиката в модели M_j одним из способов:

1) оставить отношение без изменений;

2) исключить те элементы из P_i, в истинности которых j -ый эксперт не совсем уверен;

3) добавить в отношение новые элементы и исключить некоторые старые, например, с учетом «знаний» других экспертов;

4) выполнить пункты 2) и 3) одновременно.

Обозначим произвольный расширенный класс моделей теории T через $Mod_n(\Omega)$. Введем расстояние на множестве «знаний» экспертов с помощью построенного класса моделей $Mod_n(\Omega)$ *мощности не более, чем n* . Модели различаются интерпретациями, но на них выполнены все аксиомы теории T.

Определим расстояние между формульными подмножествами, (предикатами) в каждой модели $M_i \in Mod_n(\Omega)$, как меру их симметрической разности или ее модификацию.

Определение3. Расстоянием между различными предикатами $P_k^{M_i}$ и $P_j^{M_i}$, определенными в модели M_i, назовем величину

$$\rho_{M_i}\left(P_k^{M_i}, P_j^{M_i}\right) = \mu\left(\widetilde{P}_k^{M_i} \Delta \widetilde{P}_j^{M_i}\right).$$

Замечание. Это определение корректно, если предикаты одной местности и с одинаковым набором переменных. Если рассматриваемые предикаты имеют разную местность (арность) или разный набор переменных, и эксперт считает отсутствующую в одной из формул переменную x_i несущественной, полагаем, что она принимает любое из возможных значений. В противном случае (если она существенна) доопределяем эту переменную, добавив конъюнктивно к нужной формуле предикат $\widetilde{P}_{x_i}$, уточняющий значения этой переменной. В дальнейшем, с учетом выше сказанного, будем изучать расстояние между формулами от одних и тех же переменных (одной арности).

Можно рассматривать случай вычисления расстояния между различными интерпретациями одной и той же формулы на одном носителе, но это предмет другой работы.

Расстояние между формулами, определенными на множестве моделей $Mod_n(\Omega)$, определим как среднее на множестве расстояний в моделях.

Определение4. Расстоянием между формулами P_k и P_j, определенными на множестве $Mod_n(\Omega)$, назовем величину

$$\rho_1\left(P_k, P_j\right) = \frac{\sum\limits_{M_i \in Mod_n(\Omega)} \rho_{M_i}\left(P_k^{M_i}, P_j^{M_i}\right)}{\left|Mod_n(\Omega)\right|}.$$

Теперь рассмотрим способ определения расстояния между предложениями (замкнутыми формулами). Обозначим через $Mod(\phi)$ множество моделей из $Mod_n(\Omega)$, на которых истинно предложение ϕ, т.е. $Mod(\phi) = \left\{ M_i \in Mod_n(\Omega) \mid \ M_i \models \phi \right\}$.

Очевидно, существуют такие модели, на которых предложение истинно, и такие, на которых оно ложно (если оно не тавтология). Естественно измерять различие информации, содержащейся в предложениях, количеством моделей, на которых предложения принимают разные значения истинности.

Определение5. Расстоянием между предложениями ϕ и ψ назовем величину

$$\rho_2(\phi,\psi)=\frac{\left|Mod\left(\left(\phi\wedge\neg\psi\right)\vee\left(\neg\phi\wedge\psi\right)\right)\right|}{\left|Mod_n(\Omega)\right|}.$$

Рассмотрим еще один способ определения расстояния между формулами. Дополним множество Ω константами из множества носителей моделей. Для этого множества M рассмотрим произвольные кортежи $\overline{a}$ длины местности формул, равной $l(\overline{a})$. При подстановке кортежей в формулы, в предположении, что формулы имеют одинаковую местность (как этого добиться, было указано выше), формулы становятся предложениями.

Определение6. Расстоянием между формулами $\phi(\overline{x})$ и $\psi(\overline{x})$ назовем величину

$$\rho_3\left(\phi(\overline{x}),\psi(\overline{x})\right)=\min_{\overline{a}\in M^{l(\overline{a})}}\rho_2\left(\phi(\overline{a}),\psi(\overline{a})\right).$$

Доказана следующая теорема, из которой следует, что предложенные расстояния действительно являются метриками. В теореме приведены и некоторые дополнительные свойства введенных расстояний. При доказательствах теорем используется теоретико модельные методы цитируемых работ [1-8].

Теорема1. Для любого измеримого расширения (= все интересующие нас формулы-знания имеют измеримую область истинности, т.е. меру) исходного класса моделей теории T для любых формул («знаний» экспертов) ϕ, ψ, χ и для любой функции ρ_i справедливы следующие свойства:

1. $0\le\rho_i(\phi,\psi)\le 1$
2. $\rho_i(\phi,\psi)=\rho_i(\psi,\phi)$ (симметричность расстояния).
3. $\rho_i(\phi,\psi)\le\rho_i(\phi,\chi)+\rho_i(\chi,\psi)$ (неравенство треугольника).
4. $\phi\equiv\psi\Leftrightarrow\rho_i(\phi,\psi)=0$ ($\phi\equiv\psi$ здесь и далее обозначает эквивалентность

формул относительно всех моделей, то есть для любой модели M_i верно $\phi^{M_i}\equiv\psi^{M_i}$)

5. $\phi\equiv\neg\psi\Rightarrow\rho_i(\phi,\psi)=1$
6. $\rho_i(\phi,\psi)=1-\rho_i\left(\phi,\neg\psi\right)=\rho_i\left(\neg\phi,\neg\psi\right)$.

7. $\rho_i(\phi,\psi) = \rho_i(\phi \wedge \psi, \phi \vee \psi)$.

8. $\rho_i(\phi, \neg\phi) = \rho_i(\phi,\psi) + \rho_i(\psi, \neg\phi)$.

Доказательство теоремы следует из определений, свойств вероятностной меры, теоретико-модельных и логических вычислений как в [7].

Мера опровержимости (информативности)высказываний экспертов, вероятности формул и их свойства

С точки зрения важности информации, сообщенной экспертом, естественно считать, что информативность высказывания (непустого предиката) тем выше, чем меньше число элементов, ему удовлетворяющих (более точно, чем меньше мера, определенная на этом подмножестве). Поэтому введем меру опровержимости, которая используется как информативность для выполнимых формул, следующим образом.

Определение7. Пусть $\phi(\overline{x})$ - формула, отражающая «знание» эксперта, тогда мерой опровержимости формулы $\phi(\overline{x})$ назовем величину $I_i(\phi(\overline{x})) = \rho_i(\phi(\overline{x}),1)$, где **1** -- тождественно истинный предикат, например, $\overline{x} = \overline{x}$.

Для введенных расстояний получаем:

$$I_i(\phi) = \begin{cases} \dfrac{\sum\limits_{M_i \in Mod_n(\Omega)} \mu\left(\neg\phi^{M_i}\right)}{|Mod_n(\Omega)|}, \text{если } \rho_1 \\ \dfrac{|Mod(\neg\phi)|}{|Mod_n(\Omega)|}, \text{если } \rho_2 \\ \dfrac{\min\limits_{\overline{a} \in M^{l(\overline{a})}} |Mod(\neg\phi(\overline{a}))|}{|Mod_n(\Omega)|}, \text{если } \rho_3 \end{cases}$$

Для меры опровержимости справедлива следующая теорема.

Теорема2. Для любого измеримого расширения исходного класса моделей теории T для любых формул ϕ, ψ и для любой функции ρ_i справедливы следующие утверждения:

1. $0 \le I_i(\phi) \le 1$.

2. Если $\phi \equiv \psi$, то $I_i(\phi) = I_i(\psi)$.

3. $I_i(1) = 0$ (1 - тождественно истинная формула)

4. $I_i(0) = 1$ (0- тождественно ложная формула).

5. $I_i(\phi) = 1 - I_i(\neg\phi)$.

6. $I_i(\phi) \le I_i(\phi \wedge \psi)$.

7. $I_i(\phi) \ge I_i(\phi \vee \psi)$.

8. $I_i(\phi \wedge \psi) = \rho_i(\phi, \psi) + I_i(\phi \vee \psi)$.

9. Если $\rho_i(\phi, \psi) = 0$, то $I_i(\phi \wedge \psi) = I_i(\phi \vee \psi) = I_i(\phi)$.

10. $I_i(\phi \wedge \psi) = \dfrac{I_i(\phi) + I_i(\psi) + \rho_i(\phi, \psi)}{2}$.

11. $I_i(\phi \vee \psi) = \dfrac{I_i(\phi) + I_i(\psi) - \rho_i(\phi, \psi)}{2}$.

Для доказательства теоремы используются введенные определения, сформулированные выше свойства метрики и теоретико-модельные вычисления.

На практике же чаще эксперт задает высказывание с его «вероятностью». А вопрос состоит в изучении и согласовании таких высказываний, введении некоторой метрики на таких высказываниях. Первоочередной задачей, на наш взгляд, является определение вероятностей для формул с помощью модельного подхода.

Определение 8. Пусть $\phi(\bar{x})$ - формула, выражающая знание эксперта, тогда вероятность высказывания эксперта определим как величину $P_i\left(\phi(\bar{x})\right) = I_i\left(\neg\phi(\bar{x})\right) = \rho_i\left(\phi(\bar{x}), \neg 1\right)$.

Теорема3. Для любого измеримого расширения исходного класса моделей теории T для любых формул ϕ, ψ и для любого ρ_i справедливы следующие утверждения:

1- $0 \le P_i(\phi) \le 1$.

2. Если $\phi \equiv \psi$, то $P_i(\phi) = P_i(\psi)$.

3. $P_i(1) = 1$.

4. $P_i(0) = 0$.

5. $P_i(\phi) = 1 - P_i(\neg\phi)$.

6. $P_i(\phi) \ge P_i(\phi \wedge \psi)$.

7. $P_i(\phi) \le P_i(\phi \vee \psi)$.

8. $P_i(\phi \wedge \psi) = P_i(\phi \vee \psi) - \rho_i(\phi, \psi)$.

9- Если $\rho_i(\phi, \psi) = 0$, то $P_i(\phi \wedge \psi) = P_i(\phi \vee \psi) = P_i(\phi)$.

10. $P_i(\phi \wedge \psi) = \dfrac{P_i(\phi) + P_i(\psi) - \rho_i(\phi, \psi)}{2}$.

11. $P_i(\phi \vee \psi) = \dfrac{P_i(\phi) + P_i(\psi) + \rho_i(\phi, \psi)}{2}$.

Так вычисленные с использованием модельного подхода вероятности позволяют уточнять «вероятности» экспертов с учетом моделей теории T и будут применяться в дальнейшем.

ЗАКЛЮЧЕНИЕ

Полученные результаты остаются справедливыми для произвольных моделей и формульных подмножеств, аппроксимируемых измеримыми и отвечающим знаниям экспертов. Результаты можно использовать для нахождения усредненных расстояний, мер опровержимости и вероятностей высказываний экспертов. Исследование найдет применение к вопросам наилучшего согласования экспертных высказываний, построения решающих функций распознавания образов и разработки экспертных систем.

Список литературы

1. Г.С. Лбов, Н.Г. Старцева. Логические решающие функции и вопросы статистической устойчивости решений. Новосибирск: Издательство Института математики, 1999. С. 85-102.
2. Н.Г. Загоруйко, М.В. Бушуев. Меры расстояния в пространстве знаний // Анализ данных в экспертных системах. Новосибирск,1986. Выпуск 117:Вычислительные системы. С.24-35.
3. А.А. Викентьев, Г.С. Лбов. О метризации булевой алгебры предложений и информативности высказываний экспертов // Доклад РАН 1998.Т.361, №2 С.174-176.
4. A.A. Vikentiev, G.S. Lbov. Setting the metric and informativeness on statements of experts // Pattern Recognition and Image Analysis. 1997 V.7, N2, P.175-183.
5. Г. Кейслер, Ч.Ч. Чэн Теория моделей. Москва:Мир,1977.
6. Ю.Л. Ершов, Е.А. Палютин Математическая логика. Санкт-Петербург, 2004.
7. А.А.Викентьев, Л.Н. Коренева К вопросу о расстояниях между формулами, описывающими структурированные объекты. // Математические методы распознавания образов (ММРО-99). РАН ВЦ, Москва, 1999. С.151-154.
8. H. Gaifman Concerning measures in the first order calculus. // Israel Journal of Mathematics, v. 2 (1), 1964, p. 1-18.

Информация об авторе

Александр А. Викентьев – с.н.с., канд.физ-мат.наук , Институт математики СО РАН, пр. Академика Коптюга, д.4, Лаборатория анализа данных. Доцент, Новосибирский госуд. университет; e-mail: vikent@math.nsc.ru