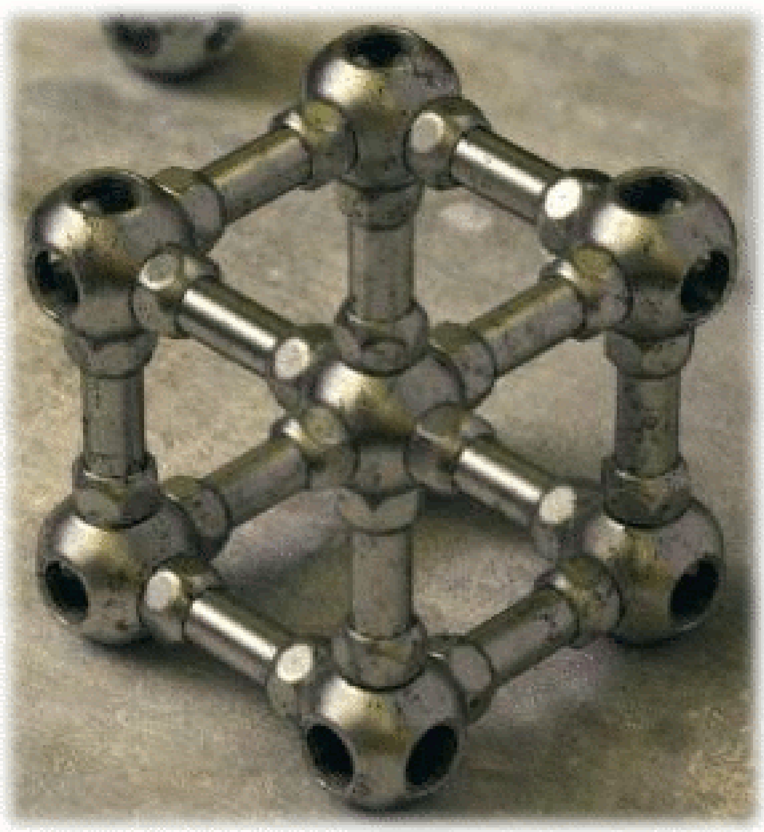

Caracterização da Independência Condicional em Lógica Modal [1]

Author(s):

Alexandre Matos Arruda

Marcelo Finger

[1]This work was supported by Fapesp Project LogProb, grant 2008/03995-5, São Paulo, Brazil.

Caracterização da Independência Condicional em Lógica Modal*

Alexandre Matos Arruda Marcelo Finger

1 Introdução

A independência condicional (IC) é um dos conceitos básicos em probabilidade. Sua importância na estatística foi acentuada por Dawid [14] em 1979, onde algumas propriedades formais da IC foram mencionadas. Desde então muitos artigos foram publicados sobre essas propriedades (exemplos em [4], [15]).

O grande interesse em axiomatizar a independência condicional é por seus benefícios na teoria da probabilidade para sistemas especialistas. A noção de IC pode ser interpretada como um certo relacionamento entre sintomas (que podem ser descritos por variáveis aleatórias) e portanto promete a possibilidade de determinar uma estrutura adequada de um sistema especialista perguntando diretamente a especialistas (veja [12], [2]).

A importância da IC para sistemas especialistas foi explicitamente mostrada e focada especialmente por Pearl que em [10] formulou uma conjectura concreta para a caracterização de RIC (relação de independência condicional) correpondente a vetores aleatórios cuja distribuição são medidas positivas. Essa conjectura foi refutada em [12] ao se encontrar uma nova propriedade independente dos RICs. Outra questão motivada pelo mesmo trabalho foi resolvida em [9].

Entretando, alguns resultados positivos foram alcançados a esse respeito nos quais certas subclasses de afirmações de IC foram caracterizadas. Em [8] e independentemente em [6] uma completa caracterização para a classe "marginal" de afirmações de IC (isto é, afirmações $I(A; B|C)$ com C fixo) foram encontradas. Em [3] e também em [11] (usando diferentes descrições formais) a classe dos "contextos-fixos" de afirmações de IC (isto é, $I(A; B|C)$ onde $A \cup B \cup C$ é fixo) foram caracterizadas. Em conexão com estes resultados gostaríamos de mencionar atenção ao artigo [7] onde classes especiais de RIC's ("monotônica" na condição) são caracterizadas.

Em [13] Studeny demonstra que a IC não é caracterizável. A prova é similar a prova de que as dependências multivalorados embutidas em teoria de banco de dados não são axiomatizáveis [5]. Essa prova porém é limitada a uma linguagem proposicional. Aqui iremos propor uma axiomatização para a IC utilizando uma lógica de maior de poder de expressão que a lógica proposicional clássica: A Lógica Modal.

*This work was supported by Fapesp Project 2008/03995-5 (LogProb)

Na seção 2 iremos apresentar nossa linguagem para a lógica modal e alguns conceitos importantes como a definição de independência condicional.

Uma parte da axiomatização será proposta na seção 3.

Conclusões serão apresentadas na seção 4.

2 Conceitos Preliminares

A lógica modal trabalha no contexto de uma linguagem com necessidade e possibilidade. As sentenças das linguagem são da seguinte forma:

$$\mathbb{P}_0, \mathbb{P}_1, \mathbb{P}_2$$

$$\bot, \top, \neg A, A \wedge B, A \vee B, A \rightarrow B, A \leftrightarrow B, \Box A, \Diamond A$$

Sentenças da forma $\mathbb{P}_n$ (para $n = 0, 1, 2, \ldots$) são atômicas. $\top$ é uma constante para verdade; $\bot$ é uma constante para falsidade. $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ são símbolos de negação, conjunção, disjunção, condicionalidade e bicondicionalidade, respectivamente. $\Box$ é o símbolo da necessidade e $\Diamond$ é o símbolo da possibilidade. Uma abordagem mais detalhada pode ser encontrada em [1].

2.1 A Verdade e Mundos Possíveis

De acordo com a idéia de Leibniz, necessidade é o que é verdadeiro em todo mundo possível e possibilidade é o que é verdadeiro em algum deles. Linguisticamente, uma sentença da forma $\Box A$ - *necessariamente* A - é verdade se e somente se A é verdade em todo mundo possível e uma sentença da forma $\Diamond A$ - *possivelmente* A - é verdade apenas no caso que A é verdadeiro em algum mundo possível. Temos também uma sequência infinita de conjuntos de mundos possíveis,

$$P_0, P_1, P_2, \ldots$$

A intuição atrás desta modelagem é o fato de que para cada número natural n, o conjunto P são apenas aqueles mundos possíveis nos quais as correspondente sentença atômica $\mathbb{P}_n$ é verdade. Em outras palavras, a sequência $P_0, P_1, P_2, \ldots$ interpreta as sentenças atômicas estipulando a que cada mundo possíveis eles são verdadeiros (e, por omissão, quais eles são falsos): $\mathbb{P}_n$ é verdadeiro em um mundo possível α se e somente se α está no conjunto P_n.

Um modelo é um par

$$< W, P >$$

em que W é um conjunto de mundos possíveis e P abrevia uma sequência infinita $P_0, P_1, P_2, \ldots$ de subconjuntos de W. Note que W pode conter mundos possíveis que não estão em nenhum dos conjuntos P_n; de fato, qualquer ou todos esses conjuntos podem ser vazios. Também não é necessário que o mundo atual apareça em todo modelo.

Em termos de um mundo possível em um modelo nos estabelecemos as condições de verdade para as sentenças de acordo com suas formas. Usamos o simbolismo

$$\models^{\mathfrak{A}}_{\alpha} A$$

onde A é uma sentença e α é um mundo possível no modelo $\mathfrak{A} =< W, P >$ ou podemos também escrever

$$A \text{ é verdade em } \alpha \text{ em } \mathfrak{A}$$

As condições de verdade são estabelecidas:

1. $\models^{\mathfrak{A}}_{\alpha} \mathbb{P}_n$ sse $\alpha \in P_n$, para $n = 0, 1, 2, \ldots$;
2. $\models^{\mathfrak{A}}_{\alpha} \top$;
3. não $\models^{\mathfrak{A}}_{\alpha} \bot$;
4. $\models^{\mathfrak{A}}_{\alpha} \neg A$ sse não $\models^{\mathfrak{A}}_{\alpha}$;
5. $\models^{\mathfrak{A}}_{\alpha} A \wedge B$ sse ambos $\models^{\mathfrak{A}}_{\alpha} A$ e $\models^{\mathfrak{A}}_{\alpha} B$;
6. $\models^{\mathfrak{A}}_{\alpha} A \vee B$ sse ou $\models^{\mathfrak{A}}_{\alpha} A$ ou $\models^{\mathfrak{A}}_{\alpha} B$, ou ambos;
7. $\models^{\mathfrak{A}}_{\alpha} A \rightarrow B$ se $\models^{\mathfrak{A}}_{\alpha} A$ então $\models^{\mathfrak{A}}_{\alpha} B$;
8. $\models^{\mathfrak{A}}_{\alpha} \Box A$ sse para todo β em $\mathfrak{A}$, $\models^{\mathfrak{A}}_{\beta} A$;
9. $\models^{\mathfrak{A}}_{\alpha} \Diamond A$ sse para algum β em $\mathfrak{A}$, $\models^{\mathfrak{A}}_{\beta} A$.

2.2 O Problema da Caracterização da Independência Condicional

Seja $[\xi_i]_{i \in N}$ um vetor randômico ($2 \leq card\ N < \infty$) e suponha por simplicidade que seus componentes são variáveis aleatórias finitamente-valoradas. Então definimos uma relação ternária disjunta I sobre $exp\ N$ (disjunta significa que seu domínio é o conjunto de triplas de subconjuntos de N par a par disjuntos):

$$I(A; B|C) \text{ vale sse } [\xi_i]_{i \in A} \text{ é condicionalmente independente de } [\xi_i]_{i \in B} \text{ dado } [\xi_i]_{i \in C}.$$

Chamaremos essa relação de *relação de independência condicional* (RIC) correpondente a $[\xi_i]_{i \in N}$ como ela descreve todos os relacionamentos condicionais independentes ao longo de seus subvetores.

Nossa questão é saber se é possivel caracterizar RICs como relações ternárias disjuntas satisfazendo um conjunto de propriedades do seguinte tipo:

$$[I(A_1, B_1|C_1) \wedge I(A_2, B_2|C_2) \wedge \ldots \wedge I(A_r, B_r|C_r)] \rightarrow I(A_{r+1}, Br + 1|C_{r+1})$$

Essa questão é chamada de *problema da caracterização* da RIC em [12]

3 Axiomatição

De acordo com Studeny [13] tal caracterização não é possível, porém sua prova limita-se somente a lógica proposicional clássica. Nossa idéia é caracterizar a independência condicional usando uma linguagem modal onde ganhamos um poder maior de expressão do que o visto em uma linguagem proposicional clássica.

Considere uma linguagem que além dos conectivos booleanos clássicos, também possui uma modalidade binária "|".

Como axiomas referentes à esta nova modalidade, temos:

(**A0**)	Todos as tautologias clássicas
(**A1**)	$(B \mid A) \to (A \to B)$
(**A2**)	$(A \wedge B) \to (B \mid A)$
(**A3**)	$(A \mid B) \mid C \leftrightarrow (A \mid B \wedge C)$
(**A4**)	$A \to (A \mid (B \mid A))$
(**A5**)	$\neg C \to ((A \mid C) \leftrightarrow (B \mid C))$
(**A6**)	$(A \to B) \to ((C \mid B) \to (C \mid A))$

Mais as regras usuais de Modus Ponens e Substituição Uniforme. Note que no caso da modalidade "|" não temos nenhuma regra de inferência semelhante à regra da necessitação. A noção de dedução é a tradicional, ou seja, uma sequência de fórmulas, onde cada uma é a substituição de um axioma, ou é obtida das fórmulas anteriores na sequência usando Modus Ponens. Desta forma, obtemos os seguintes resultados, que podem ser vistos como lemas auxiliares:

(**L1**)	$(A \wedge B) \leftrightarrow ((A \mid B) \wedge B)$
(**L2**)	$A \leftrightarrow (A \mid (B \mid A))$
(**L3**)	$(A \to B) \to ((A \mid C) \to (B \mid C))$
(**L4**)	$(A \leftrightarrow B) \to ((A \mid C) \leftrightarrow (B \mid C))$
(**L5**)	$(A \leftrightarrow B) \to ((C \mid B) \leftrightarrow (C \mid A))$
(**L6**)	$((A \mid B) \leftrightarrow A) \to ((B \mid A) \leftrightarrow B)$

Em particular, este último resultado vale a pena ser demonstrado. Pelo lema **(L5)**, obtemos $((A \mid B) \leftrightarrow A) \to (B \mid (A \mid B) \leftrightarrow B \mid A)$. Por subsituição em **(L2)**, obtemos $B \mid (A \mid B) \leftrightarrow B$. Combinando estes dois últimos resultados via lógica clássica, ou seja, por **(A0)** obtemos $((A \mid B) \leftrightarrow A) \to ((B \mid A) \leftrightarrow B)$ como desejado.

Basedo neste lema, podemos definir o símbolo $\perp\!\!\!\perp$ como

$$A \perp\!\!\!\perp B =_{\text{def}} (A \mid B) \leftrightarrow A$$

representando a independência estatística das variáveis binárias A e B. Neste contexto o lema **(L6)** representa a comutatividade da independência estatística e pode ser reescrito como:

$$A \perp\!\!\!\perp B \to B \perp\!\!\!\perp A.$$

4 Conclusões

A independência condicional (IC) é um dos conceitos básicos em probabilidade. Sua importância na estatística foi acentuada por Dawid [14] em 1979, onde algumas propriedades formais da IC foram mencionadas. O grande interesse em axiomatizar a independência condicional é por seus benefícios na teoria da probabilidade para sistemas especialistas (mais detalhes em [12], [2]).

Studény em [13] demonstrou a impossibilidade da caracterização da independência condicional. Porém sua prova restringe-se à utilização de uma linguagem proposicional clássica.

Pretendemos contornar essa restrição buscando uma caracterização para esse problema utilizando uma linguagem modal, que possui maior poder de expressão que a linguagem proposicional clássica.

Referências

[1] Brian F. C. *Modal logic: an introduction*. Cambridge University Press, 1980.

[2] Perez A. e Jirousek R. Constructing and intensional expert system (ines). *Medical Decision Making: diagnostic strategies and expert systems*, pages 307–315, 1985.

[3] Geiger D. e Pearl J. Logical and algorithmic properties of conditional independence. *Proc. 2nd International Workshop on AI and Statistics*, 1989.

[4] Mouchart M. e Rolin J. M.: A note on conditional independence with statistical applications. Conditional independence in statistical theory. *Statistica*, XLIV(4):557–584, 1984.

[5] Sagiv Y. e Walecka S. F. Subset dependencies and a completeness result for a subclass of embedded multivalued dependencies. *Journal of the Association for Computing Machinery*, 29(1):103–117, 1982.

[6] Matús F. Independence and radon projetcion on compact groups. *PhD. Thesis*, 1988.

[7] Matús F. Ascending and descending conditional independence relations. *in Transactions of the 11-th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes vol. B*, pages 189–200, 1992.

[8] Paz A. e Pearl J. Geiger D. Axioms and algorithms for inference involving probabilistic independence. *Technical report CSD 890031*, 1990.

[9] Kramosil I. A note on nonaxiomatizability of independence relations generated by certain probabilistic structure. *Kybernetika*, 24(6):439–446, 1988.

[10] Pearl J. Markov and bayes networks: a comparison of two grafical representations of probabilistic knowledge. *Technical report CSD 860024*, 1986.

[11] Malvestuto F. M. A unique formal system for binary decompositions of database relations, probability distributions and graphs. *Information Science*, 1992.

[12] Studený M. Multiinformation and the problem of characterization of conditional-independence relations. *Problems of Control and Information Theory*, 1:3–26, 1989.

[13] Studený M. Conditional independence relations have no finite complete characterization. In *Information Theory, Statistical Decision Functions and Random Processes, Transactions of the 11th Prague Conference*, pages 377–396. Kluwer, Dordrecht, 1992.

[14] Dawid A. P. Conditional independence in statistical theory. *Journal of the Royal Statistical Society B*, 41:1–31, 1979.

[15] van Putter C. e van Shuppen J. H. Invariance properties of conditional independence relation. *The Annals of Probability*, 13(3):934–945, 1985.