

## Redes Bayesianas de Crença

Este exercício está baseado no cap. 14 do livro Inteligência Artificial de Russel & Norvig, 3. edição, pág. 447.

A estrutura de dados chamada Rede de Crença Bayesiana é uma boa ferramenta para representar variáveis e suas interdependências. Elas representam distribuições de probabilidades de maneira completa e concisa. A RCB também é conhecida como *Bayesian Believe Network* ou *BBN*. A teoria básica das RCB é o Teorema de Bayes. Elas são um excelente formalismo para representar conhecimento, ao mesmo tempo que viabilizam um grau importante de aprendizado. São base de importante segmento de sistemas com conhecimento na IA.

O teorema de Bayes associa quatro variáveis sendo 3 conhecidas ou supostas e a quarta desconhecida. Ele se concentra na probabilidade condicionada. Ou seja, na probabilidade de uma hipótese ser verdadeira SE tiver acontecido determinado acontecimento.

Para entender o funcionamento do Teorema de Bayes, vamos dar nomes aos bois:

**A** deve ser o evento para o qual se quer achar a probabilidade. É a pergunta a responder

**B** é o evento secundário, que efetivamente ocorreu.

$P(A)$  É a probabilidade PRÉVIA de A na população em geral. Vamos chamar esta variável de  $x$  e ela é um valor entre 0 e 1.

$P(B|A)$  É a probabilidade de B ocorrer quando A ocorreu. Chamando a esta variável de  $y$ .

$P(B|\sim A)$  A probabilidade de B ocorrer quando A NÃO ocorre. Variável  $z$ .

$P(A|B)$  Finalmente esta é a nossa incógnita. Queremos saber qual a probabilidade de A ocorrer, DADO que B ocorreu. Ela é dada pela fórmula

$$P(A|B) = \frac{x \cdot y}{(x \cdot y) + z \cdot (1 - x)}$$

No primeiro exemplo suponha que você é uma mulher que mora com o seu namorado. Ao voltar de uma longa viagem de trabalho, você vai arrumar sua gaveta de roupas íntimas e leva um susto: está lá uma roupa íntima (por hipótese uma calcinha de rendas) que não é sua. Você nunca compraria tal peça. Imediatamente você é assaltada por uma dúvida. Será que o seu parceiro está tendo um caso com outra mulher? Antes de chutar o pau da barraca, como você conhece tudo sobre probabilidade condicional, e tudo sobre o Reverendo Bayes, vamos esfriar a cabeça e fazer algumas contas. Dando nomes aos bois:

**A** é o evento estar sendo traída pelo namorado.

**B** é o evento aparecer uma calcinha desconhecida na sua gaveta.

Agora vem a parte prévia de achar/estimar probabilidades. A primeira é achar  $P(A)$  ou seja, na população, qual a probabilidade de seu namorado ter outra? A melhor coisa a fazer aqui é buscar esta probabilidade no IBGE ou similar. Se não for possível, você precisa fazer um esforço e desconsiderar a existência do evento B para esta estimativa. No nosso exemplo, há alguns estudos que sugerem que 4 parceiros casados (em 100) traem seus cônjuges ao ano. ([Sil12, pág 252]). Então  $x = 0.04$ .

Continuemos. Agora precisa-se descobrir qual a probabilidade da calcinha aparecer no seu armário dado que seu namorado tem outra. O desgraçado podia ter mais cuidado, pelo menos. Vamos estimar em 50%, ou seja  $y = 0.5$ .

Mais uma probabilidade condicional. Agora, a probabilidade é a calcinha ter aparecido, mas seu namorado está inocente. As hipóteses são várias: a. você comprou a calcinha e esqueceu. b. A pessoa que lava e cuida de suas roupas fez uma confusão qualquer. c. Uma amiga de toda a confiança usou sua lavandaria e ... d. A calcinha é do seu namorado. ... Tudo muito pouco provável, então  $z = 0.05$  ou 5% apenas.

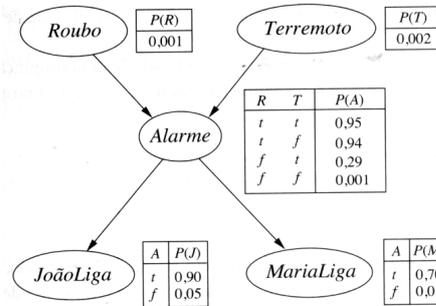
Finalmente, vai-se calcular o que se pretendia, a saber: qual a probabilidade de você estar portando vistosos chifres, dado que a calcinha está lá na gaveta.  $P(A|B) = \frac{0.04 \times 0.5}{(0.04 \times 0.5) + 0.05 \times (1 - 0.04)} =$  que é 0.2941. Logo a probabilidade procurada é 29,4%.

Uma rede bayesiana é um grafo orientado em que cada nó é identificado com informações de probabilidade quantitativa.

- Cada nodo corresponde a uma variável aleatória que pode ser discreta ou contínua.
- Um conjunto de setas conecta pares de nodos. Se  $X \rightarrow Y$  o nodo X será dito pai de Y. O grafo não tem ciclos.
- A regra para as arestas:  $X \rightarrow Y$  quando Y depende de X. Em outras palavras X é causa de Y.
- Cada nodo X tem uma distribuição de probabilidade condicional  $P(X_i|Pais(X_i))$  que quantifica o efeito dos pais sobre o nodo.

A topologia da rede especifica os relacionamentos de independência condicional que são válidos no domínio. O significado de uma seta  $X \rightarrow Y$  é que X tem influência direta sobre Y o que sugere que as causas devem ser pais dos efeitos.

Vamos construir um exemplo: você tem um alarme contra roubos em sua casa. Ele é muito eficiente, mas ocasionalmente também responde a pequenos terremotos. (Por óbvio, na cidade do exemplo, de vez em quando rola um terremoto). Você tem 2 vizinhos, João e Maria que prometeram chamar você no trabalho quando ouvirem o alarme. João quase sempre liga quando ouve o alarme, mas às vezes se confunde com o telefone e chama você indevidamente só porque o telefone tocou. Maria é meio surda e frequentemente falha ao ouvir o alarme. Dada as evidências de quem ligou ou não, quer-se estimar a probabilidade de um roubo.



Note que uma tabela para variável booleana com  $k$  pais contém  $2^k$  probabilidades independentemente especificáveis. Um nodo sem pais tem apenas uma linha representando as probabilidades anteriores a cada valor possível da variável. Veja que a rede não tem nodos correspondentes ao fato de Maria ser surda ou ao fato do telefone tocar confundindo João. Esses fatores são resumidos na incerteza associada aos vínculos de alarme para *JoãoLiga* e *MariaLiga*. Isso mostra ao mesmo tempo a preguiça e a ignorância em operação: seria muito trabalho descobrir se esses fatores seriam mais ou menos prováveis em qualquer caso específico e na verdade não temos nenhum modo razoável de obter as informações relevantes. A equação fundamental de uma rede de crença bayesiana é  $P(x_1, \dots, x_n) = \prod_{i=1}^n P(x_i|pais(X_i))$

Por exemplo, a probabilidade de que o alarme soou, mas não ocorreu nenhum roubo ou terremoto e João e Maria ligaram é:

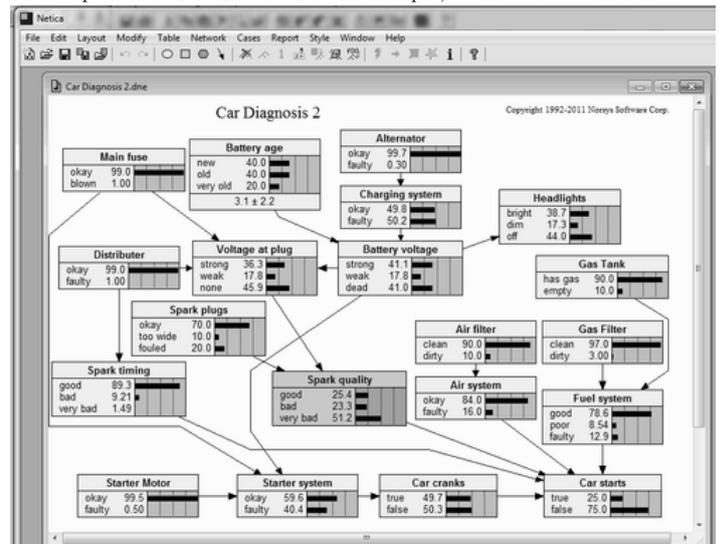
$$P(j, m, a, \sim r, \sim t) = P(j|a)P(m|a)P(a|\sim r \wedge \sim t)P(\sim r)P(\sim t) = 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998 = 0.00628 \text{ ou } 0.62\%$$

É importante salientar que as probabilidades obtidas em uma RCB podem ser conseguidas usando os teoremas fundamentais da probabilidade ( $\sum P = 1$ ,  $P(a \wedge b) = p(a) \cdot P(b)$  entre outros), mas esta estratégia cedo exibe sinais de esgotamento, já que tem comportamento exponencial. A cada nova variável inserida no contexto a TPC (tabela de probabilidades completa) dobra de tamanho. A RCB através de uma análise de causa e efeito diminui este tamanho e mantém a TPC gerenciável, mesmo para um computador.

O algoritmo para construção de uma RCB (obtido de [Rus13]).

- 1: algoritmo RCB
- 2: escolha conjunto de variáveis relevantes  $X_i$
- 3: escolha uma ordem para  $X_i$
- 4: enquanto existem variáveis
- 5: escolha  $X_i$  e adicione um nodo na rede
- 6: escolha os pais de  $X_i$  dentre os nodos que já existem na rede
- 7: conecte os pais (origem) a  $X_i$  (destino)
- 8: determine probabilidade condicional para  $X_i$
- 9: fim{enquanto}
- 10: se  $X_i$  não tem pais, anote sua probabilidade incondicional
- 11: fim{algoritmo}

Existem pacotes de software em condições de automatizar RCB. Uma implementação interessante é a de elaborar APIs para interfacear com as principais linguagens do mercado. Assim é possível criar um software parte convencional e parte RCB. Um desses produtos é o Netica (<http://www.norsys.com/netica.html>) que embora pago, permite seu uso livre para redes com menos do que 60 nodos. É o suficiente para ver uma rede em seu pleno funcionamento. Eis um exemplo, retirado desse ambiente.



### Para você fazer

Baixe e instale o software netica. Estude seus exemplos e ponha para rodar algum deles.

