Tratamento de Incertezas

Raciocínio Probabilístico

Raciocínio Probabilístico

- O Raciocínio Probabilístico é talvez o mais antigo que trata com mecanismos de incerteza. A teoria da probabilidade oferece uma maneira quantitativa de codificar incertezas.
- Apóia-se em informações probabilísticas sobre fatos de um domínio e chega a uma conclusão a respeito de um novo fato, conclusão esta, que fica associada a uma probabilidade.
  - Possui uma semântica clara
  - Probabilidades podem ser obtidas a partir de dados.
  - Permite incorporar novas evidências facilmente.

Raciocínio Probabilístico

- Exemplo:
  - Lógica Clássica:
    - Tempo Chuvoso → Trânsito Engarrafado
  - Nem sempre! E se hoje for feriado?
    - Tempo Chuvoso e ¬Feriado → Trânsito Engarrafado
  - Nem sempre! E se...

  - A idéia é associar "graus de crenças" às proposições (uma maneira de fazer isto é através de probabilidades)
  - Assim, ao invés de:
    - Tempo Chuvoso → Trânsito Engarrafado

    teremos:
    - p(Tempo Chuvoso → Trânsito Engarrafado)=0.85 <->
    - P(Trânsito Engarrafado | Tempo Chuvoso)=0.85

    como conhecimento representado no estado interno do agente
  - As probabilidades resumem a incerteza dada a complexidade do ambiente e limitações (teóricas e práticas) do agente.

Raciocínio Probabilístico

- Frequentistas x Bayesianos
  - Quando se fala de probabilidade neste contexto, não se faz referência a números, e sim, a um tipo de raciocínio.

- Exemplo:
  - "A chance de que um paciente portador da doença D apresente no futuro próximo o sintoma S é p".
    - A verdade desta afirmação não é o valor preciso de p, mas um "valor de crença" do médico.

Revisão de Conceitos Básicos

- **Experimento Aleatório**
  - Um experimento aleatório, E, é um experimento em que o resultado não pode ser predito, mesmo que seja repetido várias vezes e nas mesmas condições.
  - Experimentos:
    - E1: “Arremesso de moeda”.
    - E2: “Arremesso de 2 moedas”.
    - E3: “Arremesso de dado”.
    - E4: “Retirar uma bola de uma urna contendo bolas numeradas de 1 a 100”.
  - Variáveis aleatórias:
    - E1: “Face da moeda para cima”.
    - E2: “Faces de 2 moedas para cima”.
    - E3: “Face do dado para cima”.
    - E4: “Número da bola”

Revisão de Conceitos Básicos

- **Espaço amostral**
  - É o conjunto de todos os valores que a variável aleatória pode assumir.
  - Exemplos de espaços amostrais:
    - E1: {cara, coroa}
    - E2:{(cara,cara), (cara,coroa), (coroa,cara),(coroa,coroa)} ou E2:{(cara,cara),(cara,coroa),(coroa,coroa) }(se a ordem não é relevante)
    - E3: {1, 2, 3, 4, 5, 6}
    - E4: {1, 2, ..., 100}
  - Classificação dos Espaços Amostrais
    - Discreto x Contínuo
    - Finito x Infinito

Revisão de Conceitos Básicos

- **Eventos**
  - Um evento é um subconjunto do espaço amostral.
- Exemplos:
  - E1: “Jogar uma moeda três vezes e observar quantas vezes aparece cara e quantas vezes aparece coroa”.
  - Evento: “Aparecer 2 caras (k) e uma coroa (c) em qualquer ordem”.

    Subconjunto: {(kkc), (kck), (ckk)}
  - E2: “Jogar um dado e observar o número na face de cima”.
  - Evento: “Número é par”.

    Subconjunto: {2,4,6}

Revisão de Conceitos Básicos

- **Variável Aleatória**
  - É aquela que assume valores num espaço amostral e para a qual está determinada a probabilidade de ocorrência de cada um dos elementos do espaço amostral.
  - Se o espaço amostral for contínuo, é necessário conhecer a função distribuição de probabilidade para caracterizar a variável aleatória.
  - Se o espaço amostral for discreto, é necessário conhecer a função massa de probabilidade para caracterizar a variável aleatória.

Revisão de Conceitos Básicos

- **Probabilidade**
  - $P(A)$ = É a medida da **probabilidade** de um evento A ocorrer. Número real no intervalo [0, 1].
  - Se o espaço amostral consiste de N elementos igualmente prováveis e o evento A corresponde a um subconjunto de r elementos do espaço amostral, então a probabilidade de ocorrer A é dada por: $P(A) = r / N$
  - Probabilidade como Frequência Relativa:
    - Sequência de repetições do experimento sob idênticas condições
    - fn = número de ocorrências de A nas primeiras n repetições.
    - fn/n = fração dos experimentos em que A ocorre nas primeiras n repetições. Quando n cresce, esta fração tende a se estabilizar.

Revisão de Conceitos Básicos

- **Probabilidade**
  - Exemplo: Uma urna contém 10 bolas numeradas de 0 a 9. Num experimento é preciso selecionar uma bola da urna e anotar seu número. Deseja-se encontrar a probabilidade dos eventos:

    A = "número da bola é 5"

    B = "número da bola é ímpar"

    C = "número da bola é múltiplo de 3"
  - O espaço amostral é S={0,1,...,9} e os resultados correspondentes aos eventos acima são:

    A={5} B={1,3,5,7,9} C={3,6,9}
  - Se for suposto que os resultados são equiprováveis, então:

    P(A)= P(5) = 1/10

    P(B)=P(1)+P(3)+P(5)+P(7)+P(9) = 1/10 + 1/10 + 1/10 + 1/10 + 1/10 = 5/10.

    P(C)=P(3)+P(6)+P(9) = 1/10 + 1/10 + 1/10 = 3/10.

Revisão de Conceitos Básicos

- **Probabilidade A Priori**
  - A probabilidade a priori, também chamada de probabilidade incondicional, de um evento é a probabilidade atribuída a um evento na ausência de conhecimento que suporte a sua ocorrência ou ausência, isto é, a probabilidade do evento anterior a qualquer evidência.
  - A probabilidade a priori é simbolizada por P(evento)
- **Probabilidade Conjunta (Joint Distribution)**
  - Fornece a probabilidade de dois ou mais eventos ocorrerem simultaneamente.
- **Distribuição de Probabilidade**
  - Tabulação das probabilidades de todos os valores possíveis de uma ou mais variáveis aleatórias.

Revisão de Conceitos Básicos

- **Probabilidade Conjunta (Joint Distribution)**
  - Ao se jogar uma moeda duas vezes, temos os eventos:
  - A={Obter cara (k) na primeira jogada} -> P(A) = 1/2
  - B={Obter cara (k) na segunda jogada também} -> P(B)=1/2
  - Logo,
  - $P(A \wedge B) = P(A)P(B) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$
- Atenção!!!

$$P(A \wedge B)=P(A)P(B)$$

  - Só é válida se os eventos forem independentes, como no exemplo anterior.
- Dois eventos são independentes se a ocorrência de um não altera a probabilidade de ocorrência do outro. Caso contrário, eles são chamados de eventos **<u>não independentes</u>** ou eventos **<u>dependentes</u>**.

Revisão de Conceitos Básicos

- **Probabilidade a Posteriori ou Probabilidade Condicional**
  - A probabilidade a posteriori (após o fato), também chamada **probabilidade condicional**, de um evento é a probabilidade de um evento dada alguma evidência.
  - A probabilidade posterior é simbolizada por

    P(evento | evidência).
  - Probabilidade Condicional de A dado B (i.e., Probabilidade Condicional de ocorrer A, dado que sabemos que B ocorre).
    - $P(A|B)=P(A \wedge B)/P(B)$, com $P(B) \neq 0$
  - P(A|B) é a probabilidade de A e B relativamente a probabilidade de B.

Revisão de Conceitos Básicos

- **Probabilidade Condicional**
  - A probabilidade condicional responde questões do tipo:
    - "Dado que ocorreu o evento A, qual a probabilidade de ocorrer o evento B?"
  - Por exemplo:
    - Qual a probabilidade de um paciente estar com meningite, dado que ele está com dor de cabeça?
    - Qual a probabilidade de chover, dado que o céu está completamente nublado?
    - Qual a probabilidade de ter meningite dado que o paciente sente náuseas e está com dor?

Revisão de Conceitos Básicos

- **Probabilidade Condicional**
  - Um experimento consiste em jogar um dado uma vez. Sejam os eventos:

    A: "Número na face é 6"

    B: "Número na face é maior que 4"
  - O dado foi jogado. Qual a probabilidade de ocorrer o evento A (sair 6), sabendo que o evento B ocorreu?
  - $P(A|B) = P(A \wedge B) / P(B)$

    $= (1/6) / (1/3)$

    $= 1/2$

Revisão de Conceitos Básicos

- **Probabilidade Condicional**

	Não Fuma	Fuma	Total
Sem Câncer	40	10	50
Câncer	7	3	10
Total	47	13	60

- Qual a probabilidade de uma pessoa ter câncer, dado que ela fuma?

$$P(C=\text{“}S\text{”} \mid F=\text{“}S\text{”}) = P(C=\text{“}S\text{”} \wedge F=\text{“}S\text{”}) / P(F=\text{“}S\text{”})$$

$$= (3/60)/ (13/60)$$

$$= 0{,}23 \text{ ou } 23\%$$

Revisão de Conceitos Básicos

– Outro Exemplo:
  - Inteligência do Aluno (I): i^0(baixa),i^1(alta)
  - Dificuldade da Disciplina (D): d^0(fácil), d^1(difícil)
  - Conceito do Aluno (C):c^1(A), c^2(B), c^3(C)

2X2X3 eventos independentes

I	D	C	Prob.
i^0	d^0	c^1	0.126
I^0	d^0	c^2	0.168
i^0	d^0	c^3	0.126
i^0	d^1	c^1	0.009
i^0	d^1	c^2	0.045
I^0	d^1	c^3	0.126
i^1	d^0	c^1	0.252
I^1	d^0	c^2	0.0224
I^1	d^0	c^3	0.0056
i^1	d^1	c^1	0.06
I^1	d^1	c^2	0.036
I^1	d^1	c^3	0.024

soma=1

$P(i^0,d^1,c^3)=0.126$

Revisão de Conceitos Básicos

– Exemplo:

- Condicionamento

  Condicionamento em c^1: (elimina toas as combinações onde $c^i \neq c^1$

I	D	C	Prob.
i^0	d^0	c^1	0.126
~~i^0~~	~~d^0~~	~~c^2~~	~~0.168~~
~~i^0~~	~~d^0~~	~~c^3~~	~~0.126~~
i^0	d^1	c^1	0.009
~~i^0~~	~~d^1~~	~~c^2~~	~~0.045~~
~~i^0~~	~~d^1~~	~~c^3~~	~~0.126~~
i^1	d^0	c^1	0.252
~~i^1~~	~~d^0~~	~~c^2~~	~~0.0224~~
~~i^1~~	~~d^0~~	~~c^3~~	~~0.0056~~
i^1	d^1	c^1	0.06
~~i^1~~	~~d^1~~	~~c^2~~	~~0.036~~
~~i^1~~	~~d^1~~	~~c^3~~	~~0.024~~

Revisão de Conceitos Básicos

- Exemplo:
  - Condicionamento: Redução

I	D	C	Prob.
i^0	d^0	c^1	0.126
i^0	d^1	c^1	0.009
i^1	d^0	c^1	0.252
i^1	d^1	c^1	0.06

Revisão de Conceitos Básicos

- Exemplo:
  - Condicionamento: Renormalização

I	D	C	Prob.
i^0	d^0	c^1	0.126
i^0	d^1	c^1	0.009
i^1	d^0	c^1	0.252
i^1	d^1	c^1	0.06

medida não normalizada $P(I,D,c^1)$ = 0.447

/0.447

I	D	Prob.
i^0	d^0	0.282
i^0	d^1	0.02
i^1	d^0	0.564
i^1	d^1	0.134

P(I ^ D | c1)

Probabilidade Condicional

Revisão de Conceitos Básicos

- **Distribuição Marginal**
  - Oferece as probabilidades de vários valores das variáveis no subconjunto sem referenciar aos valores das outras variáveis.
  - O termo variável marginal é usado para referir às variáveis no subconjunto de variáveis sendo retidas. Estes termos são denominados de "marginal" porque eles costumam ser encontrados através da soma de valores em uma tabela ao longo de linhas ou colunas, e a escrita dessa soma é dada nas margens da tabela.

Revisão de Conceitos Básicos

– Exemplo:
  - Marginalize I

P(I,D)

I	D	Prob.
i^0	d^0	0.282
i^0	d^1	0.02
i^1	d^0	0.564
i^1	d^1	0.134

D	Prob.
d^0	0.846
d^1	0.154

Revisão de Conceitos Básicos

- **Inferência Probabilística**
  - 1º. Método: Usar a Distribuição Conjunta Total como "base de conhecimento"
  - Exemplo: Qual a probabilidade do aluno ter baixa inteligência ($I=i^0$) sabendo que o aluno tirou conceito A ($C=c^1$)?

I	D	C	Prob.
i^0	d^0	c^1	0.126
i^0	d^1	c^1	0.009
i^1	d^0	c^1	0.252
i^1	d^1	c^1	0.06

$$P(I=i^0 \mid C=c^1) = P(I=i^0 \text{ e } C=c^1) / P(C=c^1) = \frac{0.126+0.009}{0.126+0.009+0.252+0.06} = 0.302$$

Revisão de Conceitos Básicos

- **Inferência Probabilística**
  - 1º. Método: Usar a Distribuição Conjunta Total como "base de conhecimento"
  - Problema:
    - Para n variáveis booleanas a distribuição conjunta total possui 2^n probabilidades!
    - Na modelagem de problemas realistas pode haver centenas (ou até milhares) de variáveis aleatórias
    - Assim, a distribuição conjunta total não é uma ferramenta prática para inferência probabilística
    - Serve apenas como fundamento teórico sobre o qual podem ser elaboradas abordagens mais efetivas

Inferência Probabilística

- Teorema de Bayes:
  - Seja:
    - P(A|B) a probabilidade de que a hipótese A seja verdadeira dada a evidência B.
    - P(B|A) a probabilidade que a evidência B será observada se a hipótese A for verdadeira.
    - P(A) a probabilidade "a priori" que a hipótese A é verdadeira na ausência de qualquer evidência específica.
    - P(B) a probabilidade "a priori" que a evidência B é verdadeira na ausência de qualquer evidência específica.

    - $P(A|B)=P(A \wedge B)/P(B)$
    - $P(B|A)=P(B \wedge A)/P(A)$

    Como $P(A \wedge B)=P(B \wedge A)$

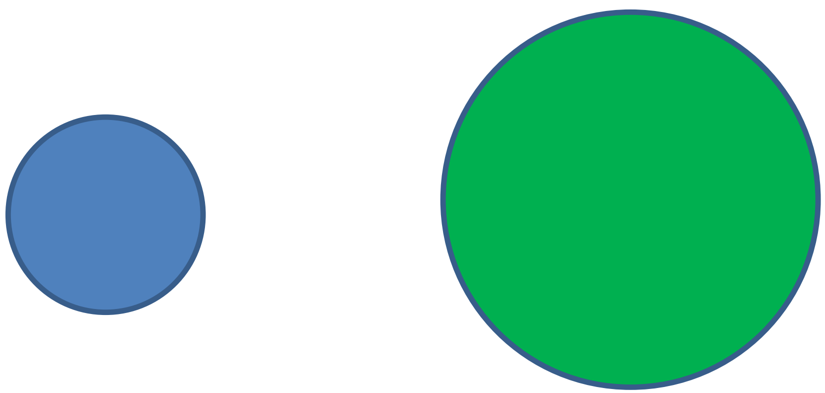
    - $P(A|B)=P(B|A) * P(A)/P(B)$

Inferência Probabilística

- Teorema de Bayes:
  - Na prática a regra de Bayes é útil porque existem muitos casos onde temos boas estimativas para $P(B|A)$, $P(B)$ e $P(A)$ e precisamos calcular $P(A|B)$.

Inferência Probabilística

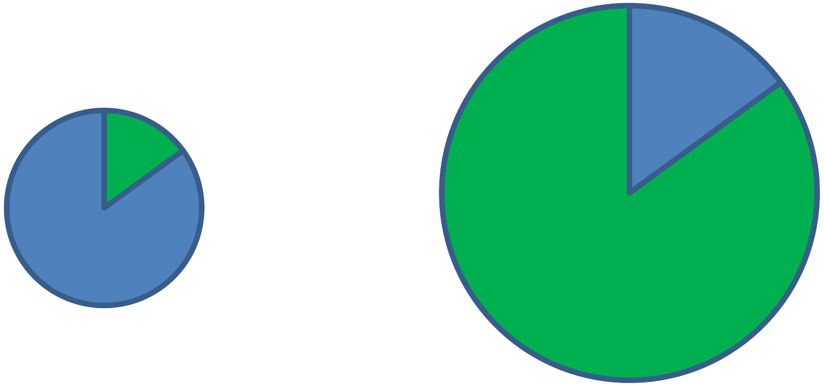
- Exemplo Introdutório:
  - Em uma cidade existem duas companhias de taxis, os azuis e os verdes. 15% das taxis são azuis e 85% dos taxis são verdes.
  - Em um acidente noturno, uma testemunha assegurou ter visto um taxi azul envolvido no acidente.
  - Experimentos conduzidos pela polícia atestaram que esta testemunha era capaz de acertar a cor de um taxi sob as mesmas condições de iluminação em 80% das vezes.
  - De que cor era o taxi envolvido no acidente?



De cada 100 taxis, 15 são azuis e 85 são verdes.

Inferência Probabilística

- Exemplo Introdutório:
  - Em uma cidade existem duas companhias de taxis, os azuis e os verdes. 15% das taxis são azuis e 85% dos taxis são verdes.
  - Em um acidente noturno, uma testemunha assegurou ter visto um taxi azul envolvido no acidente.
  - Experimentos conduzidos pela polícia atestaram que esta testemunha era capaz de acertar a cor de um taxi sob as mesmas condições de iluminação em 80% das vezes.
  - De que cor era o taxi envolvido no acidente?



De cada 100 taxis, 15 – (20%) são azuis e vistos azuis
e 85 – (20%) são verdes e vistos como verdes. Mas 20% dos 85 são verdes mas vistos como azuis e 20% dos 15 são azuis mas vistos como verdes.

Inferência Probabilística

- Exemplo Introdutório:
  - Em uma cidade existem duas companhias de taxis, os azuis e os verdes. 15% das taxis são azuis e 85% dos taxis são verdes.
  - Em um acidente noturno, uma testemunha assegurou ter visto um taxi azul envolvido no acidente.
  - Experimentos conduzidos pela polícia atestaram que esta testemunha era capaz de acertar a cor de um taxi sob as mesmas condições de iluminação em 80% das vezes.
  - De que cor era o taxi envolvido no acidente?

CorTaxi	Testemunha	Prob
Azul	ViuAzul	0.15*0.8=0.12
Azul	ViuVerde	0.15*0.2=0.03
Verde	ViuAzul	0.85*0.2=0.17
Verde	ViuVerde	0.85*0.8=0.68

Inferência Probabilística

- Exemplo Introdutório
  - P(Taxi=Azul | Testemunha=ViuAzul) =

$$\frac{P(Taxi=Azul,Testemunha=ViuAzul)}{P(Testemunha=ViuAzul)}$$

CorTaxi	Testemunha	Prob
Azul	ViuAzul	0.15*0.8=0.12
Azul	ViuVerde	0.15*0.2=0.03
Verde	ViuAzul	0.85*0.2=0.17
Verde	ViuVerde	0.85*0.8=0.68

P(Taxi=Azul | Testemunha=ViuAzul) =
0.12 / (0.12+0.17) = 0.12/0.29 = 0.41

Inferência Probabilística

- Aplicando o Teorema de Bayes no Problema dos Taxis:
  - Taxi=Azul: hipótese = taxi envolvido no acidente ser REALMENTE azul;
  - Testemunha=ViuAzul: evidência = a testemunha ver um taxi de cor azul.
  - P(B|A) = percentagem das vezes em que a testemunha viu um taxi azul e o taxi era REALMENTE azul = 80%
  - P(A) = probabilidade a priori. Percentagem das vezes em que um taxi é azul = 15%.
  - P(B) = probabilidade a priori. Percentagem das vezes que a testemunha vê um taxi azul = 12 + 17 = 29%.

- P(Taxi=Azul | Testemunha=ViuAzul) =
  0.8 * 0.15 / 0.29 = 0.413 ~ 41%

Inferência Probabilística

- Resumo do que vimos até agora:
  - Se dispomos da distribuição de probabilidade conjunta: $P(h|e)=P(h,e)/P(e)$

I	D	Prob.
i^0	d^0	0.282
I^0	d^1	0.02
I^1	d^0	0.564
I^1	d^1	0.134

Calcule $P(i^0|d^1)$

Inferência Probabilística

- Resumo do que vimos até agora:
  - Se dispomos da distribuição de probabilidade conjunta:
    $P(h|e)=P(h,e)/P(e)$

I	D	Prob.
i^0	d^0	0.282
I^0	d^1	0.02
I^1	d^0	0.564
I^1	d^1	0.134

Calcule $P(i^0|d^1)$

$P(i^0,d^1)/P(d^1)=0.02/(0.02+0.134)=0.02/0.136=0.147$

Inferência Probabilística

- Resumo do que vimos até agora:
  - Se dispomos da distribuição de probabilidade condicional:

$P(h|e)=P(e|h).P(h) / P(e)$

  - Sabendo que 85% dos taxis são verdes e 15% são azuis e que dado que a probabilidade de uma testemunha acertar a cor do taxi é 80% ($P(t=x|c=x)=0.8$), calcule qual a probabilidade de um taxi ser azul dado que a testemunha diz que viu um taxi azul.

Inferência Probabilística

- Resumo do que vimos até agora:
  - Se dispomos da distribuição de probabilidade condicional:

$P(h|e)=P(e|h).P(h) / P(e)$

  - Sabendo que 85% dos taxis são verdes e 15% são azuis e que dado que a probabilidade de uma testemunha acertar a cor do taxi é 80% ($P(t=x|c=x)=0.8$), calcule qual a probabilidade de um taxi ser azul dado que a testemunha diz que viu um taxi azul.
  - $P(e|h).P(h)/P(e) = 0.8*0.15/(0.8*0.15+0.2*0.85)=$

$= 0.12/(0.12+0.17)=0.12/0.29=0.413$

Inferência Probabilística

- As distribuições são intercambiáveis, dada uma podemos descobrir a outra!
- Exemplo:
  - Um médico sabe que meningite causa dor no pescoço em 50% dos casos. Ele sabe que a probabilidade a priori de um paciente ter meningite (m=s) é 1/50000 e a probabilidade a priori de qualquer paciente ter uma dor no pescoço (d=s) é 1/20.
  - Tem-se que:
    - $P(d=s|m=s) = 0.5$
    - $P(m=s) = 1/50000$
    - $P(d=s) = 1/20$
  - Um paciente chega ao consultório com dor no pescoço. Qual a probabilidade dele estar com meningite – $P(m=s|d=s)$?
    - $P(m=s|d=s) = P(d=s|m=s)*P(m) / P(d=s) = 0.5 * 1/50000 / 1/20 = 0.0002$
  - Você seria capaz de reconstruir a Tabela de Distribuição de Probabilidade Conjunta?

Inferência Probabilística

- Teorema de Bayes – Combinação de Evidências
  - Continuando neste problema, temos então que a certeza que do diagnóstico médico de meningite para um paciente que chega ao consultório se queixando de dores de cabeça é de apenas 0.0002, ou seja 0.02%.
  - Mas além da dor de cabeça, o paciente se queixa de náuseas. O médico também sabe que 95% dos pacientes com meningite acabam apresentando náuseas. Ele também sabe que se o paciente não tem meningite, então a probabilidade dele apresentar náuseas (por alguma outra razão) é de 20%
  - Como esta nova evidência afeta a certeza do médico em relação ao diagnóstico de meningite?

Inferência Probabilística

- Teorema de Bayes – Combinação de Evidências

  – Imaginemos agora que queremos a distribuição:

  P(meningite=s | dor=s ^ náuseas=s)

  $$P(meningite=s \mid dor=s \wedge náuseas=s) = \frac{P(dor=s \wedge náuseas=s \mid meningite=s) * P(meningite=s)}{P(dor=s \wedge náuseas=s)}$$

  – Problema
    - Precisamos conhecer as probabilidades condicionais $P(dor=s \wedge náuseas=s|meningite=s)$
    - Problema similar a conhecer distribuição conjunta

Inferência Probabilística

- Teorema de Bayes – Combinação de Evidências

  – Solução: INDEPENDÊNCIA CONDICIONAL
    - dor e náuseas não são absolutamente independentes
      – Se o paciente tem náuseas, isso significa que provavelmente há uma meningite que por sua vez deve provocar dor;

    - No entanto, são independentes dada a presença ou ausência de meningite
      – dor=s e náuseas=s são diretamente causadas por meningite=s
      – Mas nenhuma tem efeito direto sobre a outra

Inferência Probabilística

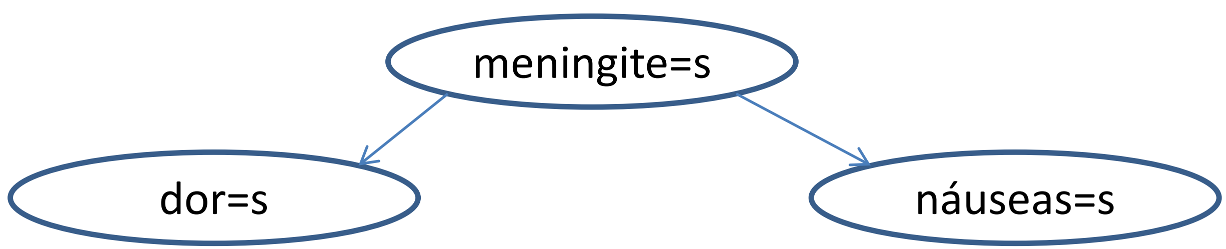
- Teorema de Bayes – Combinação de Evidências

  – Solução: INDEPENDÊNCIA CONDICIONAL

$$P(\text{meningite}=s \mid \text{dor}=s \wedge \text{náuseas}=s) = \frac{P(\text{dor}=s \wedge \text{náuseas}=s \mid \text{meningite}=s) * P(\text{meningite}=s)}{P(\text{dor}=s \wedge \text{náuseas}=s)}$$

$$P(\text{dor}=s \wedge \text{náuseas}=s \mid \text{meningite}=s) = P(\text{dor}=s \mid \text{meningite}=s) * P(\text{náuseas}=s \mid \text{meningite}=s)$$

- Conceitualmente meningite separa dor e náuseas, porque é uma causa direta de ambas



Inferência Probabilística

- Teorema de Bayes – Combinação de Evidências
  - Continuando neste problema, temos então que a certeza que do diagnóstico médico de meningite para um paciente que chega ao consultório se queixando de dores de cabeça é de apenas 0.0002, ou seja 0.02%.
  - Mas além da dor de cabeça, o paciente se queixa de náuseas. O médico também sabe que 95% dos pacientes com meningite acabam apresentando náuseas. Ele também sabe que se o paciente não tem meningite, então a probabilidade dele apresentar náuseas (por alguma outra razão) é de 20%
  - Como esta nova evidência afeta a certeza do médico em relação ao diagnóstico de meningite?

Inferência Probabilística

- Teorema de Bayes – Combinação de Evidências
  - Vamos ao cálculos:

O que já temos: p(d=s)=0.05
p(m=s)=0.00002
p(n=s|m=s)=0.95 -> p(n=n|m=s)=0.05
p(n=s|m=n)=0.20 -> p(n=s|m=n)=0.80

p(m=s|d=s,n=s)=p(d=s,n=s|m=s)*p(m=s)/p(d=s,n=s)
=p(d=s|m=s)*p(n=s|m=s)*p(m=s)/p(d=s,n=s)
=0.5*0.95*0.00002/p(d=s,n=s)

p(d=s,n=s) = p(d=s)*p(n=s)
= 0.05*(p(n=s|m=s)*p(m=s)+p(n=s|m=n)*p(m=n))
= 0.05*(0.95*0.00002+0.2*0.99998)
= 0.05*0.2
= 0.01

p(m=s|d=s,n=s)=0.0000095/0.01=0.00095~0.095%

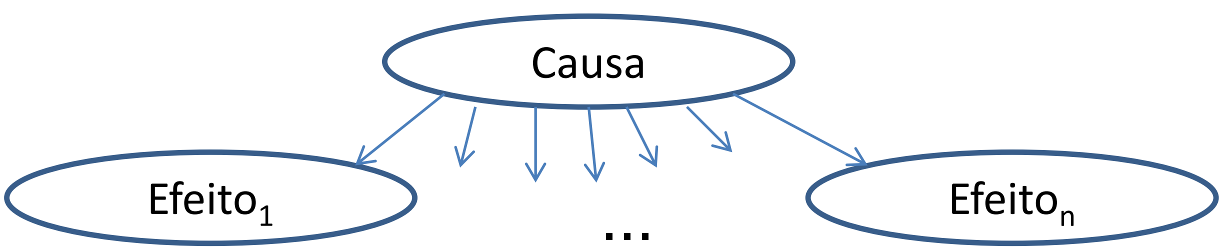
Inferência Probabilística

- Teorema de Bayes – Combinação de Evidências
  - Consequências:

$$P(dor=s, náuseas=s, meningite=s) =$$
$$P(dor=s \wedge náuseas=s \mid meningite=s) * P(meningite=s) =$$
$$P(dor=s \mid meningite=s) * P(náuseas=s \mid meningite=s) * P(meningite=s)$$

  - Generalizando (modelo Bayesiano Ingênuo):

$$P(Causa, Efeito_1, Efeito_2, ..., Efeito_n)$$
$$= P(Causa) * P(Efeito_1 \mid Causa) * P(Efeito_2 \mid Causa) *....* P(Efeito_n \mid Causa)$$



Redes Bayesianas

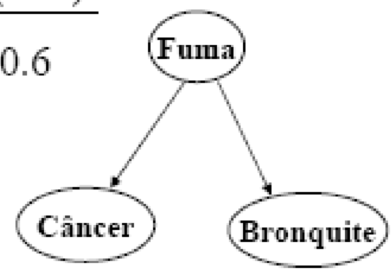
- *As redes bayesianas consistem em uma notação gráfica para representar independência condicional e assim especificar de forma concisa distribuições conjuntas totais*
- Notação
  - Uma rede bayesiana é um *grafo acíclico orientado*
  - Existe um nó para cada variável aleatória
  - Um arco orientado (seta) de um nó Y a um nó X é lido como Y é pai de X;
    - Pais(X) = {conjunto de todos nós pais de X}
  - Cada nó X_i tem uma distribuição de probabilidade condicional associada
    - P(Xi | Pais(X_i)).
- Esta distribuição quantifica o efeito das variáveis aleatórias representadas nos nós pais sobre a variável aleatória representada no nó X_i.

Redes Bayesianas

- A topologia da rede especifica independências condicionais válidas no domínio; "conhecimento causal"
- Uma seta de X para Y intuitivamente significa que X tem influência direta sobre Y
- Havendo a topologia, resta especificar uma distribuição condicional para cada variável dados os seus pais
- Ao final, topologia + distribuições, implicitamente especificam a distribuição conjunta total

Redes Bayesianas

P(F=s)	P(F=n)
0.4	0.6



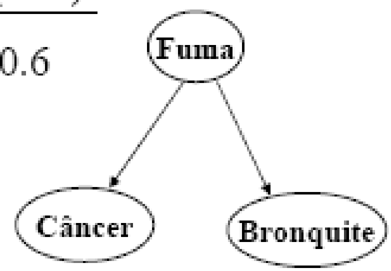
F	P(C=s)	P(C=n)
s	0.7	0.3
n	0.1	0.9

F	P(B=s)	P(B=n)
s	0.8	0.2
n	0.3	0.7

- Qual a probabilidade de alguém ter câncer, sabendo que ela fuma? $P(C|F) = 0.7$
- Qual a probabilidade de alguém ter bronquite, dado que ela fuma? $P(B|F) = 0.8$
- Qual a porcentagem de pessoas que fumam? $P(F) = 0.4$

Redes Bayesianas

P(F=s)	P(F=n)
0.4	0.6



F	P(C=s)	P(C=n)
s	0.7	0.3
n	0.1	0.9

F	P(B=s)	P(B=n)
s	0.8	0.2
n	0.3	0.7

- P(F,C,B)=P(F)*P(C|F)*P(B|F)
- Qual a probabilidade de uma pessoa fumar, ter câncer e ter bronquite?

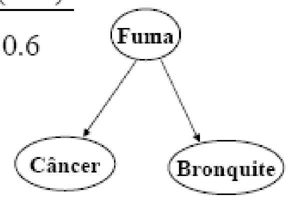
  $P(F_s,C_n,B_s)=P(F_s)*P(C_n|F_s)*P(B_s|F_s)$

  $P(F_s,C_n,B_s)=0,4*0,3*0,8$

  $P(F_s,C_n,B_s)=0,096$

Redes Bayesianas

P(F=s)	P(F=n)
0.4	0.6



F	P(C=s)	P(C=n)
s	0.7	0.3
n	0.1	0.9

F	P(B=s)	P(B=n)
s	0.8	0.2
n	0.3	0.7

- Sabendo que uma pessoa está com câncer, qual o probabilidade desta pessoa ser fumante?

$P(F_s|C_s) = P(F_s,C_s,B)/P(C_s)$

$P(F_s,C_s,B)=P(F_s)*P(C_s|F_s)*P(B_s|F_s) + P(F_s)*P(C_s|F_s)*P(B_n|F_s)$

$= 0,4*0,7*0,8+0,4*0,7*0,2=0,4*0,7*1=0,28$

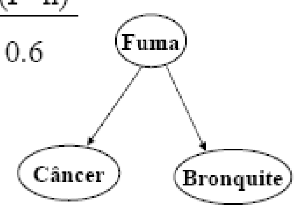
$P(C_s)=P(F_s)*P(C_s|F_s)+P(F_n)*P(C_s|F_n)$

$= 0,4*0,7 + 0,6*0,1 = 0,34$

$P(F_s|C_s) = 0,28/0,34 = 0,823$ -> ~82% chance de ser fumante

Redes Bayesianas

P(F=s)	P(F=n)
0.4	0.6



F	P(C=s)	P(C=n)
s	0.7	0.3
n	0.1	0.9

F	P(B=s)	P(B=n)
s	0.8	0.2
n	0.3	0.7

- Sabendo que uma pessoa está com câncer, qual o probabilidade desta pessoa ser fumante?

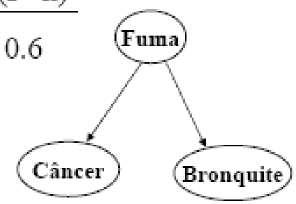
  $P(F_s|C_s)=P(C_s|F_s).P(F_s) / P(C_s)=0.7*0.4 / P(C_s)$

  $P(C_s)=P(C_s|F_s)*P(F_s)+P(C_s|F_n)*P(F_n)=0.7*0.4+0.1*0.6=0.34$

  $P(F_s|C_s)=0.7*0.4/0.34=0.28/0.34=0.823$ -> 82% de ser fumante

Redes Bayesianas

P(F=s)	P(F=n)
0.4	0.6



F	P(C=s)	P(C=n)
s	0.7	0.3
n	0.1	0.9

F	P(B=s)	P(B=n)
s	0.8	0.2
n	0.3	0.7

- Sabendo que uma pessoa está com bronquite, qual o probabilidade desta pessoa ter câncer?

P(F,Cs,Bs) = P(Fs,Cs,Bs)+P(Fn,Cs,Bs)/P(Bs)

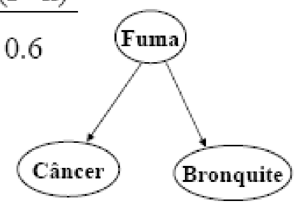
P(Fs,Cs,Bs)=P(Fs)*P(Cs|Fs)*P(Bs|Fs)=0,4*0,7*0,8=0,224

P(Fn,Cs,Bs)=P(Fn)*P(Cs|Fn)*P(Bs|Fn)=0,6*0,1*0,3=0,018

P(Bs)=P(Bs|Fs)*P(Fs)+P(Bs|Fn)*P(Fn)

Redes Bayesianas

P(F=s)	P(F=n)
0.4	0.6



F	P(C=s)	P(C=n)
s	0.7	0.3
n	0.1	0.9

F	P(B=s)	P(B=n)
s	0.8	0.2
n	0.3	0.7

- Sabendo que uma pessoa está com bronquite, qual o probabilidade desta pessoa ter câncer?

P(F,Cs,Bs) = P(Fs,Cs,Bs)+P(Fn,Cs,Bs)/P(Bs)

P(Bs)=P(Bs|Fs)*P(Fs)+P(Bs|Fn)*P(Fn)

P(Fs)*P(Bs|Fs)=0,4*0,8=0,32

P(Fn)*P(Bs|Fn)=0,6*0,3=0,18

Redes Bayesianas

- Sabendo que uma pessoa está com bronquite, qual o probabilidade desta pessoa ter câncer?

P(F,Cs,Bs) = P(Fs,Cs,Bs)+P(Fn,Cs,Bs)/P(Bs)

P(Cs|Bs)=P(Bs,Cs|Fs)+P(Bs,Cs|Fn)/P(Bs)=

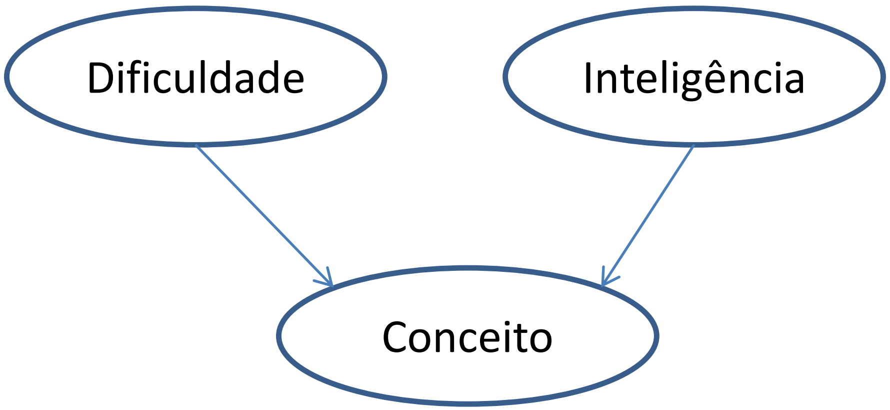
$$\frac{0.224+0.018}{0.32+0.18} = 0.484$$

Redes Bayesianas

- Exemplo:
  - Um aluno quer obter uma carta de referência do professor Mauro para concorrer a uma bolsa no Ciências sem Fronteiras.
  - Para dar esta carta de recomendação, o professor se baseia em informações sobre a inteligência do aluno, o conceito que o aluno tirou na sua disciplina, no grau de dificuldade da disciplina ministrada e se o aluno passou ou não no TOEFL.
  - O conceito (A, B ou C) que o aluno tirou na disciplina depende do grau de dificuldade da disciplina (Difícil ou Fácil) e da inteligência do aluno (Alta ou Baixa)
  - A inteligência do aluno também influencia na sua chance de passar ou não no TOEFL.
  - Quanto maior o conceito do aluno na disciplina, maior a chance do professor dar a carta de recomendação.
  - Quanto maior a inteligência do aluno, maior a chance dele tirar um conceito alto na disciplina e maior a chance dele passar no TOEFL.
  - Quanto mais difícil a disciplina, menor a chance do aluno tirar um conceito alto na disciplina.

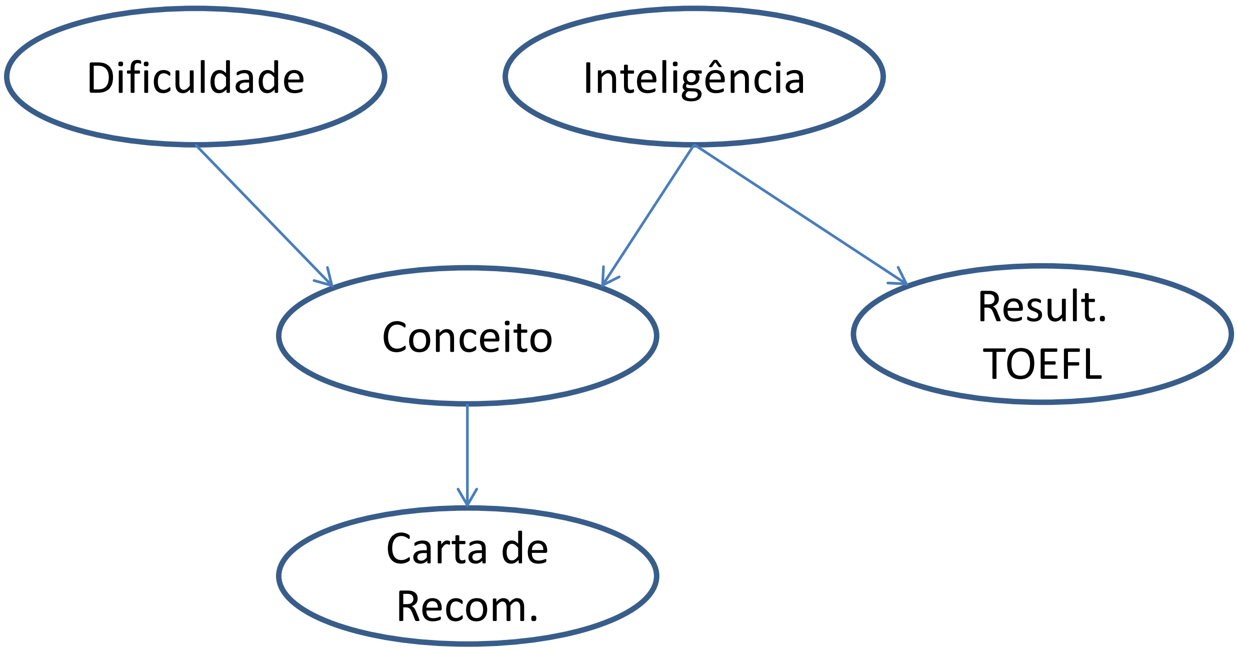
Redes Bayesianas

- Exemplo:
  - Vamos montar a Rede Bayesiana que modela o raciocínio do professor Mauro:



Redes Bayesianas

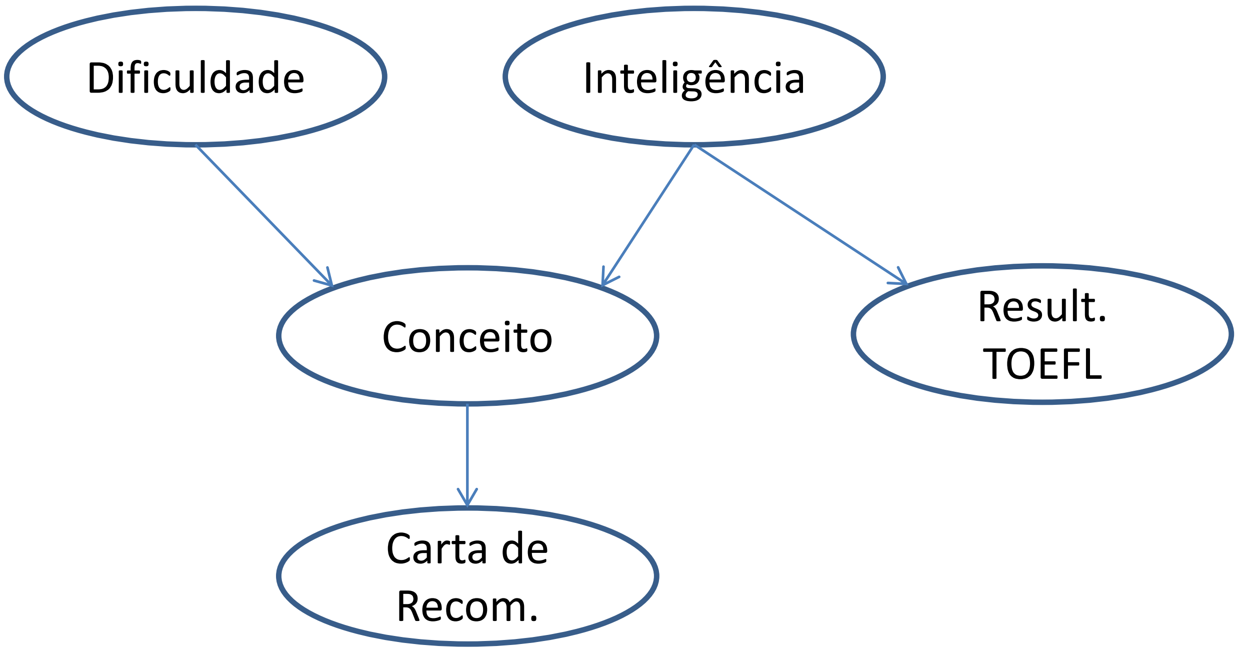
- Exemplo:
  - Vamos montar a Rede Bayesiana que modela o raciocínio do professor Mauro:



Redes Bayesianas

- Exemplo:
  - Vamos preencher as Tabelas de Probabilidades Condicionais:

Fácil-d_f	Difícil-d_d
0.6	0.4

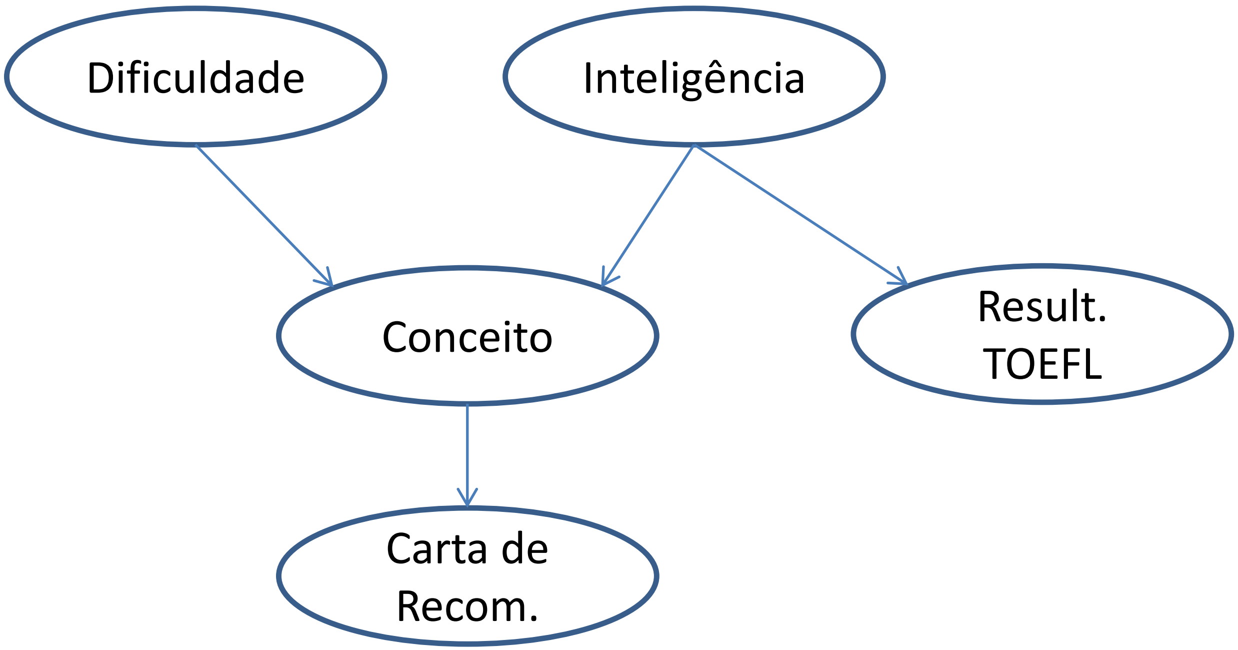


Redes Bayesianas

- Exemplo:
  – Vamos preencher as Tabelas de Probabilidades Condicionais:

Fácil-d_f	Difícil-d_d
0.6	0.4

Baixa-i_b	Alta-i_a
0.7	0.3

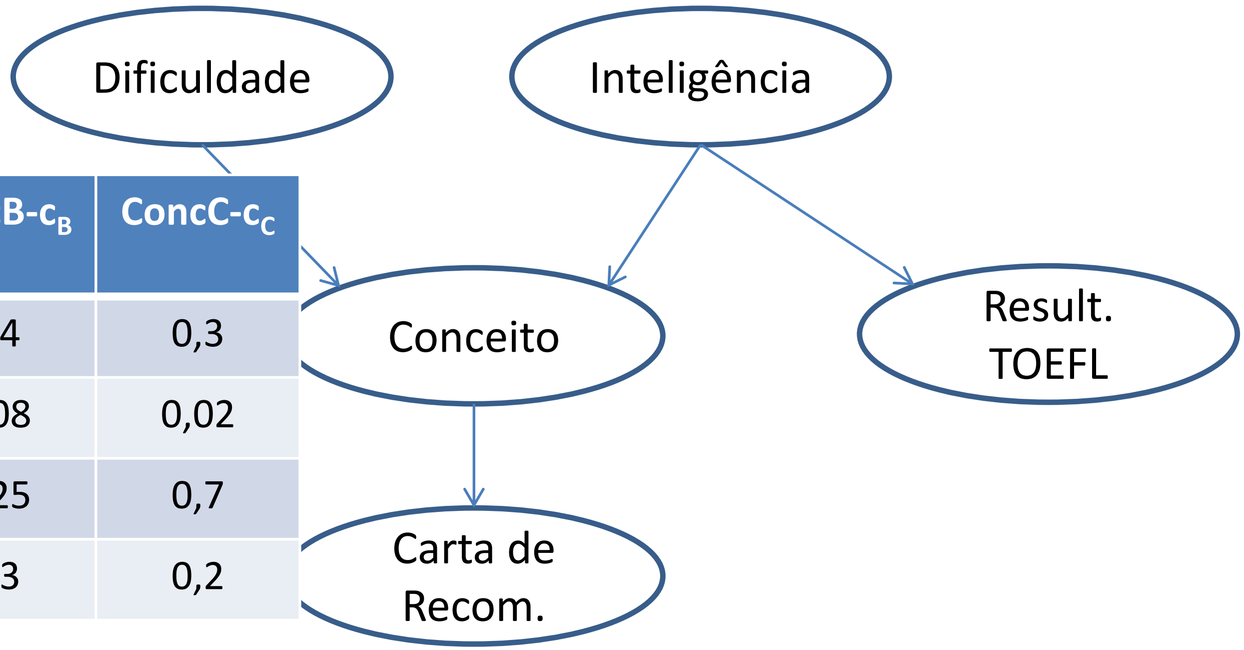


Redes Bayesianas

- Exemplo:
  - Vamos preencher as Tabelas de Probabilidades Condicionais:

Fácil-d_f	Difícil-d_d
0.6	0.4

Baixa-i_b	Alta-i_a
0.7	0.3



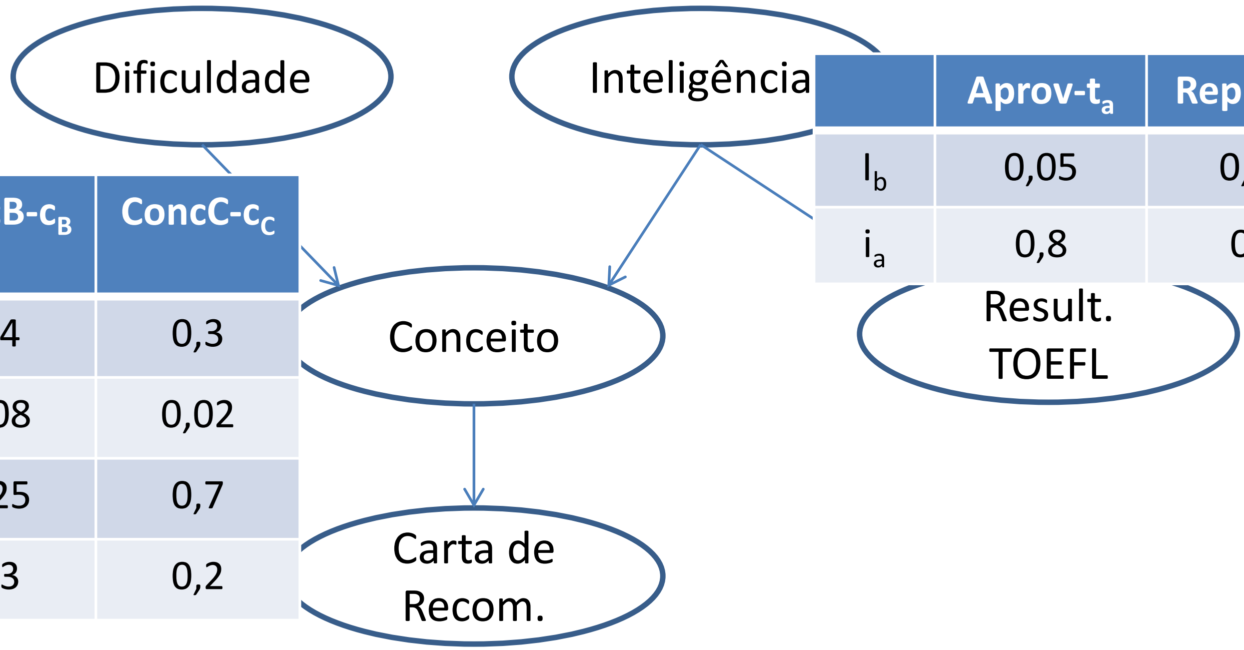
	ConcA-c_A	ConcB-c_B	ConcC-c_C
d_f,i_b	0,3	0,4	0,3
d_f,i_a	0,9	0,08	0,02
d_d,i_b	0,05	0,25	0,7
d_d,i_a	0,5	0,3	0,2

Redes Bayesianas

- Exemplo:
  – Vamos preencher as Tabelas de Probabilidades Condicionais:

Fácil-d_f	Difícil-d_d
0.6	0.4

Baixa-i_b	Alta-i_a
0.7	0.3



	ConcA-c_A	ConcB-c_B	ConcC-c_C
d_f,i_b	0,3	0,4	0,3
d_f,i_a	0,9	0,08	0,02
d_d,i_b	0,05	0,25	0,7
d_d,i_a	0,5	0,3	0,2

	Aprov-t_a	Reprov-t_r
I_b	0,05	0,95
i_a	0,8	0,2

Redes Bayesianas

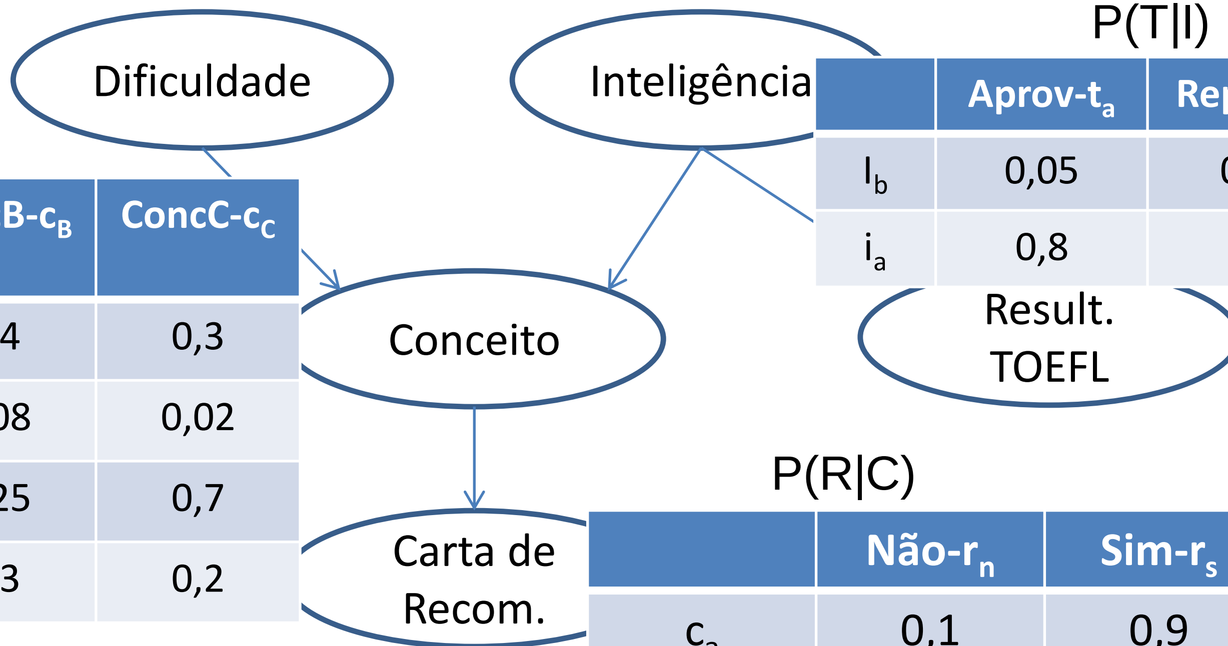
- Exemplo:
  - Vamos preencher as Tabelas de Probabilidades Condicionais:

P(D)

Fácil-d_f	Difícil-d_d
0.6	0.4

P(I)

Baixa-i_b	Alta-i_a
0.7	0.3



P(T|I)

	Aprov-t_a	Reprov-t_r
I_b	0,05	0,95
i_a	0,8	0,2

P(C|D,I)

	ConcA-c_A	ConcB-c_B	ConcC-c_C
d_f,i_b	0,3	0,4	0,3
d_f,i_a	0,9	0,08	0,02
d_d,i_b	0,05	0,25	0,7
d_d,i_a	0,5	0,3	0,2

P(R|C)

	Não-r_n	Sim-r_s
c_a	0,1	0,9
c_b	0,4	0,6
c_c	0,99	0,01

Redes Bayesianas

- Exemplo:
  - P(D,I,C,T,R)=?
    - P(D)*P(I)*P(C|D,I)*P(T|I)*P(R|C)
  - Qual a probabilidade de, em uma população, eu encontrar um aluno inteligente, que cursou uma disciplina difícil, obteve um conceito B e passou no TEOFL e para o qual o professor deu uma carta de recomendação?

Redes Bayesianas

- Exemplo:
  - P(D,I,C,T,R)=?
    - P(D)*P(I)*P(C|D,I)*P(T|I)*P(R|C)

  - Qual a probabilidade de, em uma população, eu encontrar um aluno inteligente, que cursou uma disciplina difícil, obteve um conceito B e passou no TEOFL e para o qual o professor deu uma carta de recomendação?

$P(d_d,i_a,c_b,t_a,r_s)=P(d_d)*P(i_a)*P(c_b|d_d,i_a)*P(t_a|i_a)*P(r_s|c_b)$

$$\left.\begin{array}{l} P(d_d)=0,4 \\ P(i_a)=0,3 \\ P(c_b|d_d,i_a)=0,3 \\ P(t_a|i_a)=0,8 \\ P(r_s|c_b)=0,6 \end{array}\right\} 0,4*0,3*0,3*0,8*0,6=0,01728 \sim 1,72\%$$

Redes Bayesianas

- Exemplo:
  - P(D,I,C,T,R)=?
    - P(D)*P(I)*P(C|D,I)*P(T|I)*P(R|C)

  - Qual a probabilidade de um professor dar uma carta de recomendação para um aluno inteligente, que cursou uma disciplina difícil, obteve um conceito B e passou no TEOFL?

$$P(r_s|d_d,i_a,c_b,t_a)=P(d_d,i_a,c_b,t_a|r_s)* P(r_s) / P(d_d,i_a,c_b,t_a)$$

Redes Bayesianas

- Exemplo:
  - P(D,I,C,T,R)=?
    - P(D)*P(I)*P(C|D,I)*P(T|I)*P(R|C)
  - Qual a probabilidade de um professor dar uma carta de recomendação para um aluno inteligente, que cursou uma disciplina difícil, obteve um conceito B e passou no TEOFL?

$P(r_s|d_d,i_a,c_b,t_a)=P(d_d,i_a,c_b,t_a|r_s)*P(r_s) / P(d_d,i_a,c_b,t_a)$

$P(d_d,i_a,c_b,t_a,r_s)=P(d_d)*P(i_a)*P(c_b|d_d,i_a)*P(t_a|i_a)*P(r_s|c_b)$

$$\left.\begin{array}{l} P(d_d)=0,4 \\ P(i_a)=0,3 \\ P(c_b|d_d,i_a)=0,3 \\ P(t_a|i_a)=0,8 \\ P(r_s|c_b)=0,6 \end{array}\right\} 0,4*0,3*0,3*0,8*0,6=0,01728 \sim 1,72\%$$

Redes Bayesianas

- Exemplo:
  - $P(D,I,C,T,R)=?$
    - $P(D)*P(I)*P(C|D,I)*P(T|I)*P(R|C)$
  - Qual a probabilidade de um professor dar uma carta de recomendação para um aluno inteligente, que cursou uma disciplina difícil, obteve um conceito B e passou no TEOFL?

$P(r_s|d_d,i_a,c_b,t_a)=P(d_d,i_a,c_b,t_a|r_s)* P(r_s) / P(d_d,i_a,c_b,t_a)$

$P(d_d,i_a,c_b,t_a)=P(d_d)*P(i_a)*P(c_b|d_d,i_a)*P(t_a|i_a)$

$$\left.\begin{array}{l} P(d_d)=0,4 \\ P(i_a)=0,3 \\ P(c_b|d_d,i_a)=0,3 \\ P(t_a|i_a)=0,8 \end{array}\right\} \quad 0,4*0,3*0,3*0,8=0,0288 \sim 2,88\%$$

Redes Bayesianas

- Exemplo:
  - P(D,I,C,T,R)=?
    - P(D)*P(I)*P(C|D,I)*P(T|I)*P(R|C)
  - Qual a probabilidade de um professor dar uma carta de recomendação para um aluno inteligente, que cursou uma disciplina difícil, obteve um conceito B e passou no TEOFL?

$$P(r_s|d_d,i_a,c_b,t_a)=P(d_d,i_a,c_b,t_a|r_s)*\ P(r_s)\ /\ P(d_d,i_a,c_b,t_a)$$

$$\frac{P(d_d,i_a,c_b,t_a,r_s)}{P(d_d,i_a,c_b,t_a)}=\frac{P(\cancel{d_d})*P(\cancel{i_a})*P(c_b|\cancel{d_d},i_a)*P(\cancel{t_a}|i_a)*P(r_s|c_b)}{P(\cancel{d_d})*P(\cancel{i_a})*P(c_b|\cancel{d_d},i_a)*P(\cancel{t_a}|i_a)}$$

$$\frac{0{,}0172}{0{,}0288} = 0{,}5972 \sim 60\% = P(r_s|c_b)$$

Redes Bayesianas

- Exemplo:
  - Qual a probabilidade de um aluno para o qual o professor não deu a carta de recomendação ser um aluno inteligente?

$P(i_a|r_n)=P(r_n|i_a)^* P(i_a) / P(r_n)$

$P(r_n|i_a)^*P(i_a) =$

$P(r_n,c_a,d_f,i_a)+P(r_n,c_a,d_d,i_a)+P(r_n,c_b,d_f,i_a)+P(r_n,c_b,d_d,i_a)+\ldots$

$+P(r_n,c_c,d_f,i_a)+P(r_n,c_c,d_d,i_a)$

$= P(r_n|c_a)^*P(c_a|d_f,i_a)^*P(d_f)^*P(i_a) + \ldots$

$+P(r_n|c_a)^*P(c_a|d_d,i_a)^*P(d_d)^*P(i_a) + \ldots$

$+P(r_n|c_b)^*P(c_b|d_f,i_a)^*P(d_f)^*P(i_a) + \ldots$

$+P(r_n|c_b)^*P(c_b|d_d,i_a)^*P(d_d)^*P(i_a) + \ldots$

$+P(rn|c_c)^*P(c_c|d_f,i_a)^*P(d_f)^*P(i_a) + \ldots$

$+P(rn|c_c)^*P(c_a|d_d,i_a)^*P(d_d)^*P(i_a)$

Redes Bayesianas

- Exemplo:
  - Qual a probabilidade de um aluno para o qual o professor não deu a carta de recomendação ser um aluno inteligente?

$P(i_a|r_n)=P(r_n|i_a)^* \; P(i_a) \;/\; P(r_n)$

$P(r_n|i_a)^*P(i_a) =$

$P(r_n,c_a,d_f,i_a)+P(r_n,c_a,d_d,i_a)+P(r_n,c_b,d_f,i_a)+P(r_n,c_b,d_d,i_a)+\ldots$

$+P(r_n,c_c,d_f,i_a)+P(r_n,c_c,d_d,i_a)$

$= P(r_n|c_a)^*P(c_a|d_f,i_a)^*P(d_f)^*P(i_a) + \quad \ldots=0{,}1^*0{,}9^*0{,}6^*0{,}3$

$+P(r_n|c_a)^*P(c_a|d_d,i_a)^*P(d_d)^*P(i_a) \; + \ldots=0{,}1^*0{,}5^*0{,}4^*0{,}3$

$+P(r_n|c_b)^*P(c_b|d_f,i_a)^*P(d_f)^*P(i_a) \; + \ldots=0{,}4^*0{,}08^*0{,}6^*0{,}3$

$+P(r_n|c_b)^*P(c_b|d_d,i_a)^*P(d_d)^*P(i_a) \; + \ldots=0{,}4^*0{,}3^*0{,}4^*0{,}3$

$+P(rn|c_c)^*P(c_c|d_f,i_a)^*P(d_f)^*P(i_a) \; + \ldots=0{,}99^*0{,}02^*0{,}6^*0{,}3$

$+P(rn|c_c)^*P(c_a|d_d,i_a)^*P(d_d)^*P(i_a) =0{,}99^*0{,}2^*0{,}4^*0{,}3$

$= 0{,}0162+0{,}006+0{,}00576+0{,}0144+0{,}003564+0{,}02376= 0{,}069684$

Redes Bayesianas

- Exemplo:
  – Qual a probabilidade de um aluno para o qual o professor não deu a carta de recomendação ser um aluno inteligente?

$P(i_a|r_n)=P(r_n|i_a)* P(i_a) / P(r_n)$

$P(r_n) = P(r_n,c_a,d_f,i_a)+P(r_n,c_a,d_d,i_a)+P(r_n,c_b,d_f,i_a)+P(r_n,c_b,d_d,i_a)+...$

$+P(r_n,c_c,d_f,i_a)+P(r_n,c_c,d_d,i_a) + ... = 0,069684$

$+ P(r_n,c_a,d_f,i_b)+P(r_n,c_a,d_d,i_b)+P(r_n,c_b,d_f,i_b)+P(r_n,c_b,d_d,i_b)+...$

$+P(r_n,c_c,d_f,i_b)+P(r_n,c_c,d_d,i_b) =$

$= P(r_n|c_a)*P(c_a|d_f,i_b)*P(d_f)*P(i_b) + ...=0,1*0,3*0,6*0,7$

$+P(r_n|c_a)*P(c_a|d_d,i_b)*P(d_d)*P(i_b) + ...=0,1*0,05*0,4*0,7$

$+P(r_n|c_b)*P(c_b|d_f,i_b)*P(d_f)*P(i_b) + ...=0,4*0,4*0,6*0,7$

$+P(r_n|c_b)*P(c_b|d_d,i_b)*P(d_d)*P(i_b) + ...=0,4*0,25*0,4*0,7$

$+P(r_n|c_c)*P(c_c|d_f,i_b)*P(d_f)*P(i_b) + ...=0,99*0,3*0,6*0,7$

$+P(r_n|c_c)*P(c_a|d_d,i_b)*P(d_d)*P(i_b) =0,99*0,7*0,4*0,7$

$= 0,0126+0,0014+0,0672+0,028+0,12474+0,19404 = 0,42798$

Redes Bayesianas

- Exemplo:
  - Qual a probabilidade de um aluno para o qual o professor não deu a carta de recomendação ser um aluno inteligente?

$$P(i_a|r_n)=P(r_n|i_a)*P(i_a) / P(r_n)$$

$$\frac{P(r_n|i_a)*P(i_a)}{P(r_n)} = \frac{0{,}069684}{0{,}069684 + 0{,}42798} = 0{,}14002 \sim 14\%$$

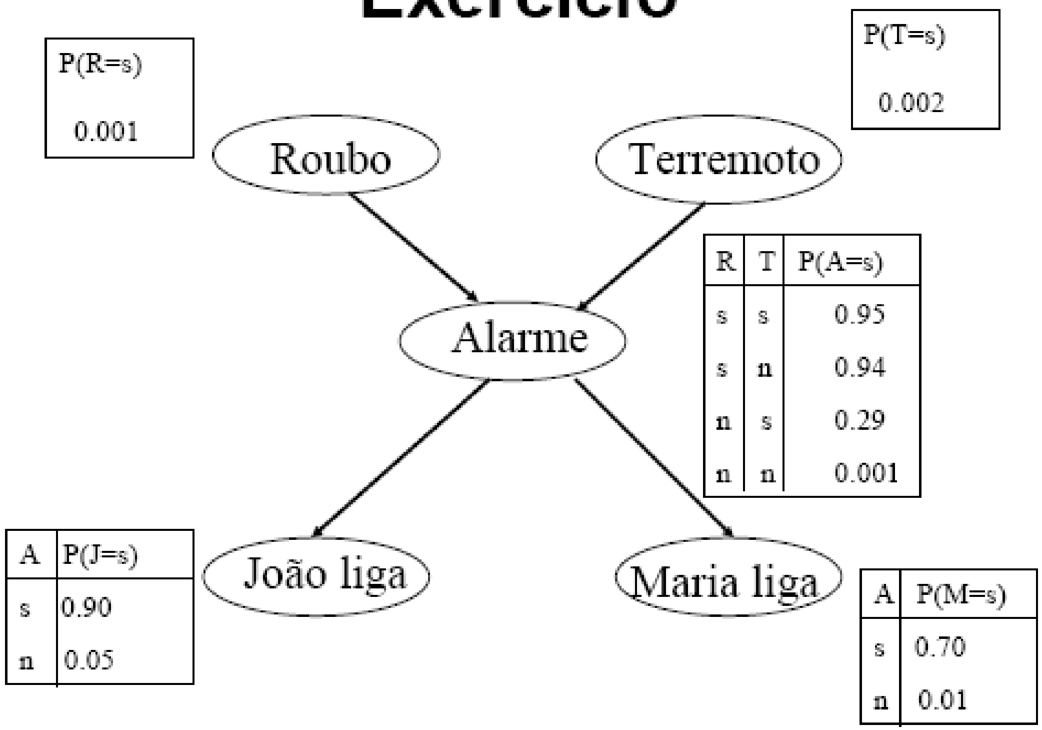
Redes Bayesianas

- Outro Exemplo:
  - Você instalou um alarme contra roubos na sua casa, que dispara em caso de invasão.
  - Infelizmente, o alarme é sensível a terremotos
  - Quando o alarme disparar, seus 2 vizinhos, João e Maria, disseram que vão te ligar.
  - João as vezes confunde o alarme com a sirene do carro de bombeiros.
  - Maria ouve música num volume alto e nem sempre escuta o alarme
  - João te ligou.
  - Maria te ligou.
  - Qual a probabilidade de estarem roubando sua casa???

Redes Bayesianas

- Outro Exemplo:

Exercício



P(R=s)
0.001

P(T=s)
0.002

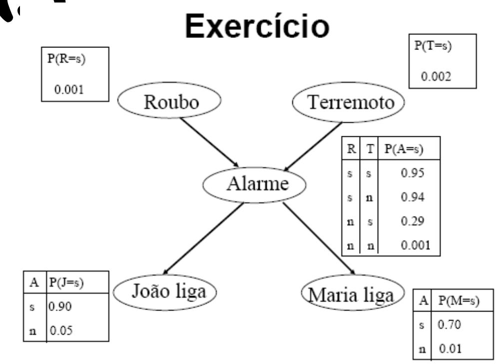
R	T	P(A=s)
s	s	0.95
s	n	0.94
n	s	0.29
n	n	0.001

A	P(J=s)
s	0.90
n	0.05

A	P(M=s)
s	0.70
n	0.01

• Qual a probabilidade de estar havendo um roubo, sabendo que a Maria e o João ligaram?

Redes Bayesianas



- Outro Exemplo:
  - $P(R|J,M) = \dfrac{P(J,M|R).P(R)}{P(J,M)} =$

$$\frac{P(J,M,A,T,R)+P(J,M,\neg A,T,R)+P(J,M,A,\neg T,R)+P(J,M,\neg A,\neg T,R)}{P(J,M)}$$

$P(J,M,A,T,R)=P(J|A).P(M|A).P(A|T,R).P(T).P(R)=0{,}9.0{,}7.0{,}95.0{,}002.0{,}001=$
$= 0{,}000001197$

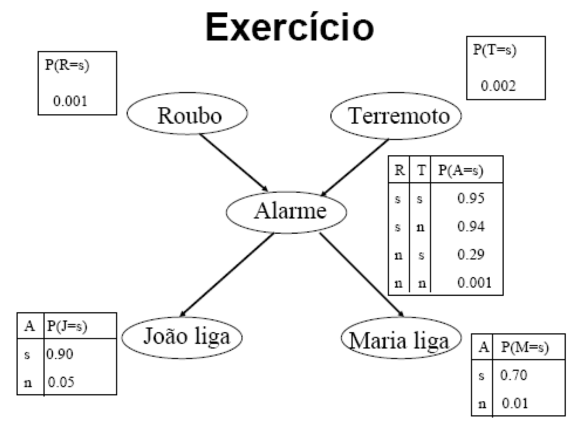
$P(J,M,\neg A,T,R)=P(J|\neg A).P(M|\neg A).P(\neg A|T,R).P(T).P(R)=$
$= 0{,}05.0{,}01.0{,}05.0{,}002.0{,}001= 0{,}00000000005$

$P(J,M,A,\neg T,R)=P(J|A).P(M|A).P(A|\neg T,R).P(\neg T).P(R)=0{,}9.0{,}7.0{,}94.0{,}998.0{,}001$
$= 0{,}0005910156$

$P(J,M,\neg A,\neg T,R)= P(J|\neg A).P(M|\neg A).P(\neg A|\neg T,R).P(\neg T).P(R)=$
$= 0{,}05.0{,}01.0{,}06.0{,}998.0{,}001= 0{,}0000002994$

$P(J,M|R).P(R)=0{,}00059224259$

Redes Bayesianas



- Outro Exemplo:
  - $P(R|J,M)=\dfrac{P(J,M|R).P(R)}{P(J,M)} =$

$P(J,M)=P(J,M|R).P(R)+P(J,M|\neg R).P(\neg R)$

$P(J,M,A,T,\neg R)=P(J|A).P(M|A).P(A|T,\neg R).P(T).P(\neg R)=0{,}9.0{,}7.0{,}29.0{,}002.0{,}999= 0{,}0003650346$

$P(J,M,\neg A,T,\neg R)=P(J|\neg A).P(M|\neg A).P(\neg A|T,\neg R).P(T).P(\neg R)=$

$= 0{,}05.0{,}01.0{,}71.0{,}002.0{,}999= 0{,}00000070929$

$P(J,M,A,\neg T,\neg R)=P(J|A).P(M|A).P(A|\neg T,\neg R).P(\neg T).P(\neg R)=$

$= 0{,}9.0{,}7.0{,}001.0{,}998.0{,}999= 0{,}00062811126$

$P(J,M,\neg A,\neg T,\neg R)= P(J|\neg A).P(M|\neg A).P(\neg A|\neg T,\neg R).P(\neg T).P(\neg R)=$

$= 0{,}05.0{,}01.0{,}999.0{,}998.0{,}999= 0{,}000498002499$

$P(J,M)=0{,}002084100239$ $\quad P(R|J,M)=\dfrac{P(J,M|R).P(R)}{P(J,M)}=\dfrac{0{,}00059224259}{0{,}002084100239}=0{,}284$

Redes Bayesianas

- Considerações Finais
  - As redes bayesianas fornecem um modo conciso de representar :
    - Relacionamentos de independência condicional
    - Uma distribuição conjunta total; cada entrada é definida como o produto das entradas correspondentes nas distribuições condicionais locais. Uma RB, com frequência, é exponencialmente menor que a distribuição conjunta total.
  - A inferência para redes bayesianas unicamente conectadas demora um tempo linear no tamanho da rede. No caso geral, o problema é intratável (NP-hard).

Raciocínio Probabilístico

- Exercícios:
  1. Suppose there are two full bowls of cookies. Bowl #1 has 10 chocolate chip and 30 plain cookies, while bowl #2 has 20 of each. Our friend Fred picks a bowl at random, and then picks a cookie at random. We may assume there is no reason to believe Fred treats one bowl differently from another, likewise for the cookies. The cookie turns out to be a plain one. How probable is it that Fred picked it out of Bowl #1?

  2. (desafio) The blue M&M was introduced in 1995. Before then, the color mix in a bag of plain M&Ms was (30% Brown, 20% Yellow, 20% Red, 10% Green, 10% Orange, 10% Tan). Afterward it was (24% Blue , 20% Green, 16% Orange, 14% Yellow, 13% Red, 13% Brown).

     A friend of mine has two bags of M&Ms, and he tells me that one is from 1994 and one from 1996. He won't tell me which is which, but he gives me one M&M from each bag. One is yellow and one is green. What is the probability that the yellow M&M came from the 1994 bag?