www.datascienceacademy.com.br

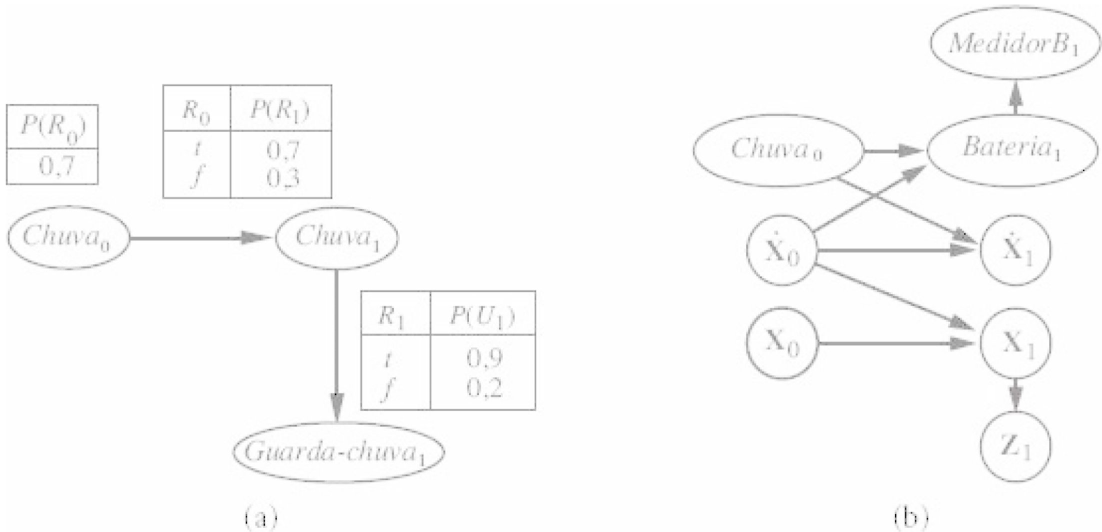
Introdução à Inteligência Artificial

Redes Bayesianas Dinâmicas

Uma rede bayesiana dinâmica, ou DBN, é uma rede bayesiana que representa um modelo de probabilidade temporal. Já vimos exemplo de DBNs: a rede de guarda-chuva. Em geral, cada fatia de uma DBN pode ter qualquer número de variáveis de estados Xt e variáveis de evidência Et. Por simplicidade, vamos supor que as variáveis e seus vínculos são reproduzidos exatamente de uma fatia para outra e que a DBN representa um processo de Markov de primeira ordem, de forma que cada variável possa ter pais somente em sua própria fatia ou na fatia imediatamente precedente.

Deve ficar claro que todo modelo oculto de Markov pode ser representado como uma DBN com uma única variável de estado e uma única variável de evidência. Também ocorre que toda DBN de variáveis discretas pode ser representada como um MOM (Modelo Oculto de Markov); podemos combinar todas as variáveis de estados na DBN em uma única variável de estado cujos valores são todas as tuplas de valores possíveis das variáveis de estados individuais. Agora, se todo MOM for uma DBN e toda DBN puder ser convertida em um MOM, qual será a diferença? A diferença é que, decompondo-se o estado de um sistema complexo em suas variáveis constituintes, a DBN será capaz de tirar proveito da escassez no modelo de probabilidade temporal. Por exemplo, suponha que uma DBN tenha 20 variáveis de estados booleanas, cada uma das quais tem três pais na fatia precedente. Então, o modelo de transição de DBN tem 20 × 23 = 160 probabilidades, enquanto o MOM correspondente tem 220 estados e, portanto, 240, ou, aproximadamente, um trilhão de probabilidades na matriz de transição. Isso é ruim por pelo menos três razões: primeiro, o próprio MOM exige muito mais espaço; em segundo lugar, a enorme matriz de transição torna a inferência de MOM muito mais dispendiosa; em terceiro lugar, o problema de aprender um número tão enorme de parâmetros torna o modelo MOM puro inadequado para problemas grandes. O relacionamento entre DBNs e MOMs é aproximadamente análogo ao relacionamento entre redes bayesianas comuns e distribuições conjuntas totalmente tabuladas.

Para construir uma DBN, devemos especificar três tipos de informações: a distribuição anterior sobre as variáveis de estados, P(X0); o modelo de transição P(Xt+1|Xt); e o modelo de sensores P(Et|Xt). Para especificar os modelos de transição e de sensores, também devemos especificar a topologia das conexões entre fatias sucessivas e entre as variáveis de estados e de evidência. Como os modelos de transição e de sensores são supostamente estacionários — os mesmos para todo t —, é mais conveniente simplesmente especificá-los para a primeira fatia. Por exemplo, a especificação de DBN completa para o mundo do guarda-chuva é dada pela rede de três nós da figura (a) abaixo. A partir dessa especificação, é possível construir a DBN completa com um número infinito de fatias de tempo, conforme seja necessário, copiando a primeira fatia.



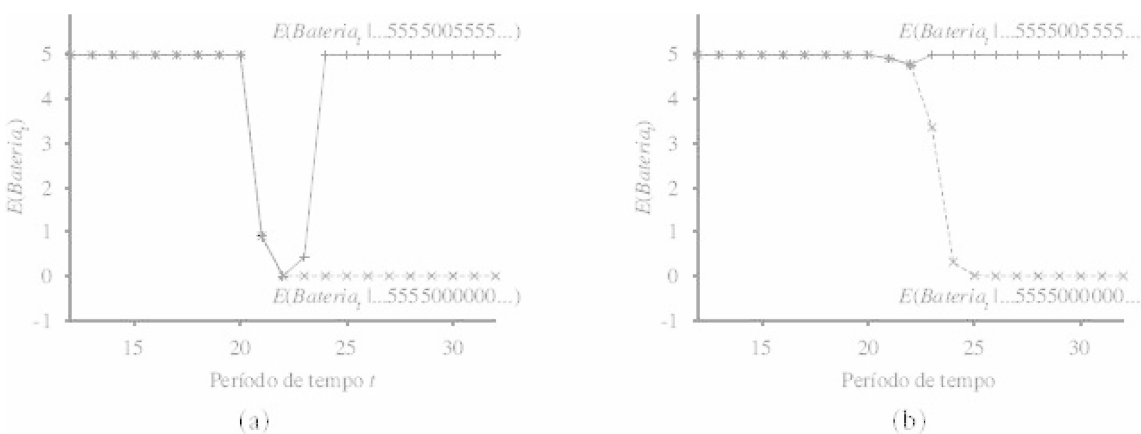
Agora, vamos considerar um exemplo mais interessante: o monitoramento de um robô alimentado por bateria que se move no plano X−Y. Primeiro, precisamos de variáveis de estados, que incluirão tanto Xt = (Xt, Yt) para representar a posição quanto para representar a velocidade.

Presumiremos algum método para medir a posição — talvez uma câmera fixa ou um GPS (Global Positioning System) a bordo — produzindo medições de Zt. A posição no período de tempo seguinte dependerá da posição atual e da velocidade. A velocidade no período seguinte dependerá da velocidade atual e do estado da bateria. Acrescentamos Bateriat para representar o nível de carga real da bateria, que tem como pais o nível da bateria anterior e a velocidade, e também adicionamos MedidorBt, que mede o nível de carga da bateria. Isso nos dá o modelo básico mostrado na figura (b).

Vale a pena examinarmos com maior profundidade a natureza do modelo de sensores correspondente a MedidorBt. Vamos supor, por simplicidade, que tanto Bateriat quanto MedidorBt possam assumir valores discretos de 0 até 5, de modo muito semelhante ao medidor do nível da bateria em um laptop típico. Se o medidor for sempre preciso, a TPC (tabela de probabilidade condicional) P(MedidorBt | Bateriat) deve ter probabilidades 1,0 "ao longo da diagonal" e probabilidades 0,0 nos outros lugares. Na realidade, o ruído sempre interfere nas medições. No caso de medições contínuas, poderia ser usada em vez disso uma distribuição gaussiana com pequena variância. No caso de nossas variáveis discretas, podemos fazer a aproximação de uma gaussiana utilizando uma distribuição na qual a probabilidade de erro caia de maneira apropriada, de modo que a probabilidade de um erro grande seja muito pequena. Utilizaremos a expressão modelo de erro gaussiano para cobrir tanto a versão contínua quanto a versão discreta.

Qualquer pessoa com experiência prática em robótica, controle de processos computadorizados ou outras formas de detecção automática admitirá prontamente o fato de que pequenas quantidades de ruído de medição são muitas vezes o menor dos problemas. Os sensores reais falham. Quando um sensor falha, ele não envia necessariamente um sinal afirmando: “Oh, a propósito, os dados que estou prestes a lhe enviar não têm o menor sentido.“ Em vez disso, ele simplesmente envia os dados sem sentido. A espécie mais simples de falha é chamada falha transiente, na qual o sensor decide ocasionalmente enviar alguns dados sem sentido. Por exemplo, o sensor de nível da bateria pode ter o hábito de enviar um zero quando alguém se choca com o robô, mesmo que a bateria esteja completamente carregada.

Vamos ver o que acontece quando ocorre uma falha transiente com um modelo de erro gaussiano que não admite tais falhas. Por exemplo, suponha que o robô esteja calmamente sentado e observe 20 leituras consecutivas da bateria indicando o nível de carga 5. Em seguida, o medidor da bateria tem um ataque temporário e a leitura seguinte é MedidorB21 = 0. Qual será o modelo de erro gaussiano simples que nos levará a acreditar em Bateria21? De acordo com a regra de Bayes, a resposta depende tanto do modelo de sensores P(MedidorB21 = 0 | Bateria21) quanto da previsão P(Bateria21 | MedidorB1:20). Se a probabilidade de um grande erro de sensor for significativamente menos provável que a probabilidade de uma transição para Bateria21 = 0, mesmo que esta última seja muito improvável, então a distribuição posterior atribuirá alta probabilidade ao fato de a bateria estar sem carga. Uma segunda leitura igual a zero em t = 22 tornará essa conclusão quase certa. Se a falha transiente desaparecer em seguida e a leitura retornar a 5 a partir de t = 23 em diante, a estimativa do nível da bateria voltará rapidamente a 5, como se fosse por mágica. Esse curso de eventos é ilustrado na curva superior da figura (a) abaixo, que mostra o valor esperado de Bateriat ao longo do tempo, usando um modelo de erro gaussiano discreto.



Apesar da recuperação, existe um tempo (t = 22) em que o robô se convence de que sua bateria está sem carga; então, presume-se que ele deve enviar um sinal de socorro e se desligar. Infelizmente, seu modelo de sensores super simplificado fez com que ele se perdesse. Como isso pode ser corrigido? Considere um exemplo familiar extraído da atividade humana diária de dirigir: em curvas bruscas ou em colinas íngremes, o “tanque de combustível se esvazia” e às vezes faz a luz de advertência acender. Em vez de procurar pelo telefone de emergência, lembramos que o medidor de nível de combustível simplesmente comete um erro muito grande quando o combustível está jogando de um lado para outro dentro do tanque. Moral da história é: para o sistema manipular falhas de sensores de maneira apropriada, o modelo de sensores deve incluir a possibilidade de falha. A espécie mais simples de modelo de falha para um sensor dá certa margem de probabilidade de que o sensor retornará algum valor completamente incorreto, independentemente do estado verdadeiro do mundo. Por exemplo, se o medidor de bateria falhar retornando 0, poderemos dizer que:

$$P(MedidorB_t = 0 \mid Bateria_t = 5) = 0{,}03$$

O que presumivelmente é muito maior que a probabilidade atribuída pelo modelo de erro gaussiano simples. Vamos chamá-lo modelo de falha transiente. De que maneira ele nos ajuda quando estamos diante de uma leitura igual a 0? Considerando-se que a probabilidade prevista de uma bateria sem carga, de acordo com as leituras feitas até o momento, é muito menor que 0,03, a melhor explicação da observação MedidorB21 = 0 é que o sensor falhou temporariamente. Por intuição, podemos imaginar que a crença sobre o nível da bateria tem certa quantidade de “inércia” que ajuda a superar oscilações temporárias na leitura do medidor. A curva superior da figura (b) anterior mostra que o modelo de falha transiente pode manipular falhas transientes sem mudança catastrófica nas crenças.

Essa é a situação no caso de oscilações temporárias. E, se houver falha persistente de um sensor? Infelizmente, falhas desse tipo são bastante comuns. Se o sensor retornar 20 leituras iguais a 5 seguidas por 20 leituras iguais a 0, então o modelo de falha de sensor transiente descrito no parágrafo anterior terá como resultado o fato de o robô ser levado gradualmente a acreditar que sua bateria está sem carga quando de fato talvez o medidor tenha falhado. A curva inferior da figura (b) mostra a “trajetória” de crença para esse caso. Depois de t = 25 — cinco leituras iguais a 0 —, o robô se convence de que sua bateria está sem carga. É óbvio que preferiríamos que o robô acreditasse que o medidor de sua bateria está quebrado, se esse de fato fosse o evento mais provável.

Até agora, apenas arranhamos a superfície do problema de representação de processos complexos. A variedade de modelos de transição é enorme, englobando tópicos tão discrepantes quanto a modelagem do sistema endócrino humano e a modelagem de vários veículos trafegando em uma autoestrada. A modelagem de sensores também é por si só um vasto subcampo, mas, mesmo fenômenos sutis, como flutuação de sensores, descalibração repentina e os efeitos de condições exógenas (como as condições do clima) sobre leituras de sensores, podem ser manipuladas por representação explícita dentro de redes bayesianas dinâmicas. E por isso, Inteligência Artificial e Internet das Coisas, formam uma dupla capaz de revolucionar o mundo!

Referências:

Livro: Inteligência Artificial
Autor: Peter Norvig