知識工学　第14回

二宮　崇

教科書と資料

- 教科書
  - Artificial Intelligence: A Modern Approach (3rd Edition): Stuart Russell, Peter Norvig (著), Prentice Hall, 2009

- この講義のウェブサイト

http://aiweb.cs.ehime-u.ac.jp/~ninomiya/ke/

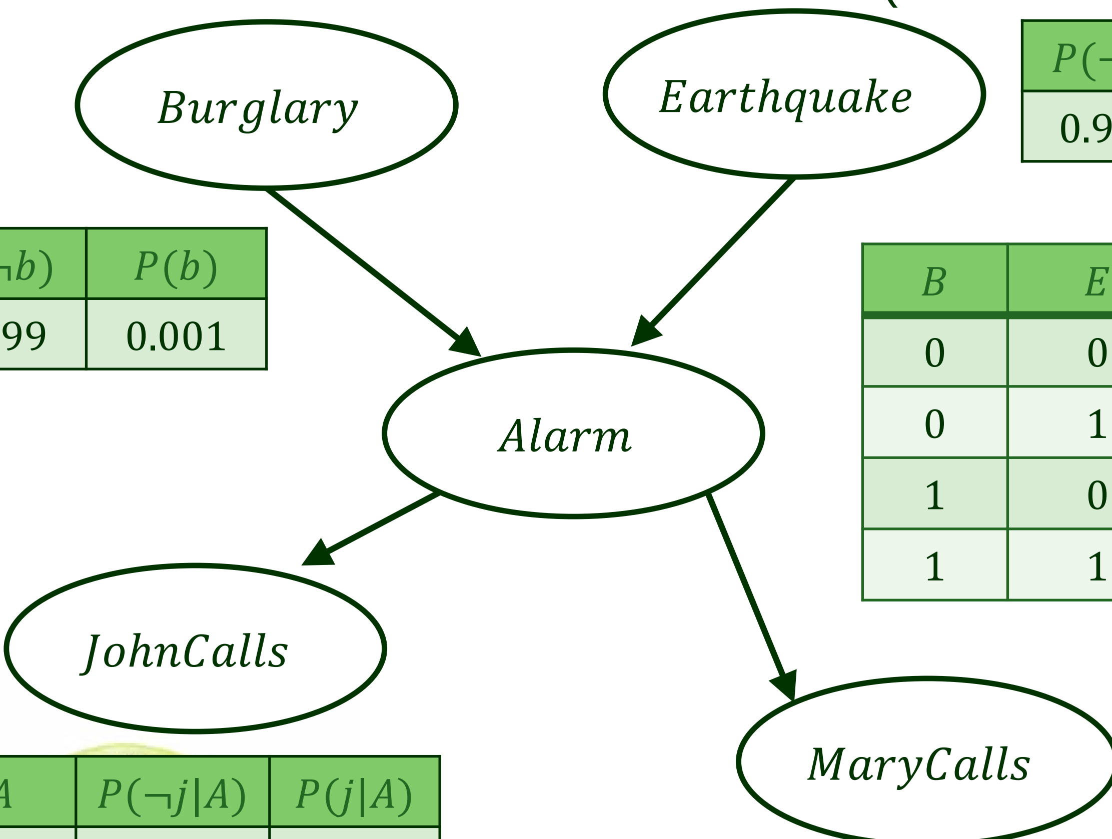
本日の講義内容

- ベイジアンネットの厳密推論 (§14.4)
  - 変数消去アルゴリズム (§14.4.2)
  - 厳密推論の計算量 (§14.4.3)
- ベイジアンネットの近似推論 (§14.5)
  - 棄却サンプリング
  - 尤度重み付け法

ベイジアンネットの復習

- 防犯アラームの例

条件付き確率表
(Conditional Probability Table, CPT)



$P(\neg e)$	$P(e)$
0.998	0.002

$P(\neg b)$	$P(b)$
0.999	0.001

B	E	$P(\neg a \mid B,E)$	$P(a \mid B,E)$
0	0	0.999	0.001
0	1	0.71	0.29
1	0	0.06	0.94
1	1	0.05	0.95

A	$P(\neg j \mid A)$	$P(j \mid A)$
0	0.95	0.05
1	0.10	0.90

A	$P(\neg m \mid A)$	$P(m \mid A)$
0	0.99	0.01
1	0.30	0.70

ベイジアンネットの復習

- 完全結合分布の表現 (§14.2.1)
  - $P(X_1 = x_1 \land \cdots \land X_n = x_n)$を$P(x_1, \cdots, x_n)$と略記する
  - $P(x_1, \cdots, x_n) = \prod_{i=1}^{n} P(x_i | parents(x_i))$
  - ただし、$parents(x_i)$は$Parents(X_i)$に含まれる変数の具体的な値の組

例: アラームが鳴り(a)、しかし、泥棒(b)も入らず、地震(e)も起きずに、ジョン(j)とメアリー(m)が電話をする確率

$$P(j \land m \land a \land \neg b \land \neg e)$$
$$= P(j|a)P(m|a)P(a|\neg b \land \neg e)P(\neg b)P(\neg e)$$
$$= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998 = 0.000628$$

ベイジアンネットの復習

- 列挙による推論 (§14.4.1)

$$P(X|e_1,\cdots,e_n) = \alpha P(X,e_1,\cdots,e_n) = \alpha \sum_{Y_1,\cdots,Y_m} P(X,e_1,\cdots,e_n,Y_1,\cdots,Y_m)$$

  - ベイジアンネットを使って$P(X,e_1,\cdots,e_n,Y_1,\cdots,Y_m)$を求めて$P(X|e_1,\cdots,e_n)$を計を計算すれば良い。

- 例

$$P(Burglary|JohnCalls = 1, MaryCalls = 1)$$

$$P(B|j,m) = \alpha P(B,j,m) = \alpha \sum_E \sum_A P(B,j,m,E,A)$$

- 時間計算量 $O(n2^n)$

ベイジアンネットの厳密推論(§14.4)

- 変数消去アルゴリズム(§14.4.2)
  - 同じ計算を二度しないように工夫する動的計画法の一種

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} \sum_E \underbrace{P(E)}_{f_2(E)} \sum_A \underbrace{P(A|B,E)}_{f_3(A,B,E)} \underbrace{P(j|A)}_{f_4(A)} \underbrace{P(m|A)}_{f_5(A)}$$

- 式の注釈付けしたパートを因子(factor)と呼ぶ。

ベイジアンネットの厳密推論(§14.4)

- 変数消去アルゴリズム(続き)

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} \sum_E \underbrace{P(E)}_{f_2(E)} \sum_A \underbrace{P(A|B,E)}_{f_3(A,B,E)} \underbrace{P(j|A)}_{f_4(A)} \underbrace{P(m|A)}_{f_5(A)}$$

A	$f_4(A)\ (= P(j|A))$
0	0.05
1	0.90

A	$f_5(A)\ (= P(m|A))$
0	0.01
1	0.70

ベイジアンネットの厳密推論(§14.4)

- 変数消去アルゴリズム (続き)

$$P(B|j,m) = \alpha f_1(B) \times \sum_E f_2(E) \times \sum_A f_3(A,B,E) \times f_4(A) \times f_5(A)$$

$$= \alpha f_1(B) \times \sum_E f_2(E) \times f_6(B,E)$$

$$f_6(B,E)$$
$$= \sum_A f_3(A,B,E) \times f_4(A) \times f_5(A)$$
$$= f_3(a,B,E) \times f_4(a) \times f_5(a) + f_3(\neg a,B,E) \times f_4(\neg a) \times f_5(\neg a)$$

ベイジアンネットの厳密推論(§14.4)

- 変数消去アルゴリズム(続き)

$$
\begin{aligned}
P(B|j,m) &= \alpha f_1(B) \times \sum_E f_2(E) \times \sum_A f_3(A,B,E) \times f_4(A) \times f_5(A) \\
&= \alpha f_1(B) \times \sum_E f_2(E) \times f_6(B,E) \\
&= \alpha f_1(B) \times f_7(B)
\end{aligned}
$$

$$
\begin{aligned}
&f_7(B) \\
&\quad = \textstyle\sum_E f_2(E) \times f_6(B,E) \\
&\quad = f_2(e) \times f_6(B,e) + f_2(\neg e) \times f_6(B,\neg e)
\end{aligned}
$$

ベイジアンネットの厳密推論(§14.4)

- 変数消去アルゴリズム (続き)
  - 因子に対する操作(掛け算)
  - $f\left(X_1 \cdots X_j, Y_1 \cdots Y_k, Z_1 \cdots Z_l\right) = f_1\left(X_1 \cdots X_j, Y_1 \cdots Y_k\right) \times f_2(Y_1 \cdots Y_k, Z_1 \cdots Z_l)$

A	B	$f_1(A,B)$	B	C	$f_2(B,C)$	A	B	C	$f_3(A,B,C)$
0	0	0.1	0	0	0.4	0	0	0	$0.1 \times 0.4 = 0.04$
0	1	0.9	0	1	0.6	0	0	1	$0.1 \times 0.6 = 0.06$
1	0	0.7	1	0	0.8	0	1	0	$0.9 \times 0.8 = 0.72$
1	1	0.3	1	1	0.2	0	1	1	$0.9 \times 0.2 = 0.18$
						1	0	0	$0.7 \times 0.4 = 0.28$
						1	0	1	$0.7 \times 0.6 = 0.42$
						1	1	0	$0.3 \times 0.8 = 0.24$
						1	1	1	$0.3 \times 0.2 = 0.06$

ベイジアンネットの厳密推論(§14.4)

- 変数消去アルゴリズム(続き)
- 因子に対する操作 (足し算)

$$f(B,C) = \sum_{A} f_3(A,B,C) = f_3(\neg a,B,C) + f_3(a,B,C)$$

B	C	$f_3(\neg a,B,C)$	$f_3(a,B,C)$	$f(B,C)$
0	0	0.04	0.28	0.32
0	1	0.06	0.42	0.48
1	0	0.72	0.24	0.96
1	1	0.18	0.06	0.24

ベイジアンネットの厳密推論(§14.4)

- 防犯アラームでの具体的な計算例

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} \sum_E \underbrace{P(E)}_{f_2(E)} \sum_A \underbrace{P(A|B,E)}_{f_3(A,B,E)} \underbrace{P(j|A)}_{f_4(A)} \underbrace{P(m|A)}_{f_5(A)}$$

B	$f_1(B)\ (= P(B))$
0	0.999
1	0.001

E	$f_2(E)\ (= P(E))$
0	0.998
1	0.002

A	B	E	$f_3(A,B,E)\ (= P(A\|B,E))$
0	0	0	0.999
0	0	1	0.71
0	1	0	0.06
0	1	1	0.05
1	0	0	0.001
1	0	1	0.29
1	1	0	0.94
1	1	1	0.95

ベイジアンネットの厳密推論(§14.4)

- 防犯アラームでの具体的な計算例

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} \sum_{E} \underbrace{P(E)}_{f_2(E)} \sum_{A} \underbrace{P(A|B,E)}_{f_3(A,B,E)} \underbrace{P(j|A)}_{f_4(A)} \underbrace{P(m|A)}_{f_5(A)}$$

A	$f_4(A)\ (= P(j\|A))$
0	0.05
1	0.90

A	$f_5(A)\ (= P(m\|A))$
0	0.01
1	0.70

ベイジアンネットの厳密推論(§14.4)

- 防犯アラームでの具体的な計算例

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} \sum_E \underbrace{P(E)}_{f_2(E)} \sum_A \underbrace{P(A|B,E)}_{f_3(A,B,E)} \underbrace{P(j|A)}_{f_4(A)} \underbrace{P(m|A)}_{f_5(A)}$$

A	B	E	$f_3(A,B,E)$	$f_4(A)$	$f_5(A)$	$f_3(A,B,E) \times f_4(A) \times f_5(A)$
0	0	0	0.999	0.05	0.01	0.0004995
0	0	1	0.71	0.05	0.01	0.000355
0	1	0	0.06	0.05	0.01	0.00003
0	1	1	0.05	0.05	0.01	0.000025
1	0	0	0.001	0.90	0.70	0.00063
1	0	1	0.29	0.90	0.70	0.1827
1	1	0	0.94	0.90	0.70	0.5922
1	1	1	0.95	0.90	0.70	0.5985

ベイジアンネットの厳密推論(§14.4)

- 防犯アラームでの具体的な計算例

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} \sum_{E} \underbrace{P(E)}_{f_2(E)} f_6(B,E)$$

$$f_6(B,E) = \sum_{A} f_3(A,B,E) \times f_4(A) \times f_5(A)$$
$$= f_3(a,B,E) \times f_4(a) \times f_5(a) + f_3(\neg a,B,E) \times f_4(\neg a) \times f_5(\neg a)$$

B	E	$f_6(B,E)$
0	0	$0.0004995 + 0.00063 = 0.0011295$
0	1	$0.000355 + 0.1827 = 0.183055$
1	0	$0.00003 + 0.5922 = 0.59223$
1	1	$0.000025 + 0.5985 = 0.598525$

ベイジアンネットの厳密推論(§14.4)

- 防犯アラームでの具体的な計算例

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} \sum_{E} \underbrace{P(E)}_{f_2(E)} f_6(B,E)$$

B	E	$f_2(E)$	$f_6(B,E)$	$f_2(E) \times f_6(B,E)$
0	0	0.998	0.0011295	0.001127241
0	1	0.002	0.183055	0.00036611
1	0	0.998	0.59223	0.59104554
1	1	0.002	0.598525	0.00119705

ベイジアンネットの厳密推論(§14.4)

- 防犯アラームでの具体的な計算例

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} f_7(B)$$

$$\begin{aligned} f_7(B) &= \sum_E f_2(E) \times f_6(B,E) \\ &= f_2(e) \times f_6(B,e) + f_2(\neg e) \times f_6(B,\neg e) \end{aligned}$$

B	$f_7(B)$
0	$0.001127241 + 0.00036611 = 0.001493351$
1	$0.59104554 + 0.00119705 = 0.59224259$

ベイジアンネットの厳密推論(§14.4)

- 防犯アラームでの具体的な計算例

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} f_7(B)$$

B	$f_1(B)$	$f_7(B)$	$f_1(B) \times f_7(B)$	$P(B|j,m)$
0	0.999	0.001493351	0.001491857649	0.715828
1	0.001	0.59224259	0.00059224259	0.284172

ベイジアンネットの厳密推論(§14.4)

- 厳密推論の計算量 (§14.4.3)
  - 多重木(polytree)
    - 任意の二つのノード間に多くても一つの無向パスしか存在しないネットワーク。
    - 単結合ネットワーク。
    - 多重木の計算量
      - CPTの項目総数に対し線形
      - 親の数が定数で制限→ノード数に対し線形

ベイジアンネットの厳密推論(§14.4)

- 厳密推論の計算量 (§14.4.3)
  - 複結合ネットワーク
    - 向きを考慮しない場合にループが存在するネットワーク
    - 複結合ネットワークの計算量
      - 指数的な時間・空間的計算量
      - #P困難

ベイジアンネットの近似推論(§14.5)

- ダイレクトサンプリング(§14.5.1)
  - 厳密推論は大規模な複結合ネットワークに対し指数的な計算時間
  - 近似的に推論を行う方法を考える
    - 棄却サンプリング
    - 尤度重み付け法

ベイジアンネットの近似推論(§14.5)

- 棄却サンプリング
  - 0〜1の間の乱数を発生
  - 乱数を用いて、ベイジアンネットの確率分布に従ったサンプルを生成
    - サンプル=全確率変数に対する具体的な値
  - 推論式

$$P(X|e_1,\cdots,e_n) \approx \frac{N(X,e_1,\cdots,e_n)}{N(e_1,\cdots,e_n)}$$

  - ただし、
    - $N(e_1,\cdots,e_n)$は$E_1,\cdots,E_n = e_1,\cdots,e_n$となったサンプルの数
    - $N(X,e_1,\cdots,e_n)$は$E_1,\cdots,E_n = e_1,\cdots,e_n$かつ$X$となったサンプルの数

ベイジアンネットの近似推論(§14.5)

- 棄却サンプリング
  - サンプルの生成方法
    - ベイジアンネットの根ノードからスタート
    - 各ノード毎に乱数を発生させる
    - CPTに従って、各ノードの値を決定する
    - 以上を子ノードに対して繰り返す

ベイジアンネットの近似推論(§14.5)

- 棄却サンプリング
  - きちんとした順序で各ノードの値を決定していけば、全ての親ノードが決定した後に子ノードの値を決定できる
  - CPTでの決定方法
    - 親の値: $B = 0, E = 1$
    - 乱数: r $(0 \leq r \leq 1)$
    - $r < 0.71$ならば$A = 0$とする
    - $r \geq 0.71$ならば$A = 1$とする

B	E	$P(\neg a)$	$P(a)$
0	0	0.999	0.001
0	1	0.71	0.29
1	0	0.06	0.94
1	1	0.05	0.95

ベイジアンネットの近似推論(§14.5)

- 棄却サンプリングの問題点
  - $E_1, \cdots, E_n = e_1, \cdots, e_n$とならなかったサンプルは計算にまったく使われない→無駄の多いサンプリング
    - 特に$e_1, \cdots, e_n$が非常に稀な現象であるとき、$E_1, \cdots, E_n = e_1, \cdots, e_n$となるサンプルがなかなか得られない。
- 尤度重み付け法
  - $E_1, \cdots, E_n = e_1, \cdots, e_n$となるように強制してサンプリング
  - 本来の分布に従ったサンプルとなるように補正(重み付け)をかける。

ベイジアンネットの近似推論(§14.5)

- 尤度重み付け法
  - 1サンプルの生成と重みの計算方法
    - w := 1.0
    - 棄却サンプリングと同様の方法で各ノードの値を順次決定
    - 証拠eに対応する確率変数Eのノードを決定するときは、乱数を用いず、$E = e$とする
    - CPTにより得られる$P(E = e|parents(E))$に対し、$w := w \times P(E = e|parents(E))$
    - この手順を繰り返し、サンプル一つ生成したところで得られるwがそのサンプルに対する重みとなる。
- 棄却サンプリングではサンプル一つを1個と数えたが、これをw個と数える

ベイジアンネットの近似推論(§14.5)

- よく用いられているサンプリング手法
  - マルコフ連鎖モンテカルロ (Markov chain Monte Carlo: MCMC)
    - ギブスサンプラー (Gibbs sampler)
    - メトロポリス-ヘイスティングサンプラー (Metropolis-Hastings sampler)