探索的モデリング (第７回)

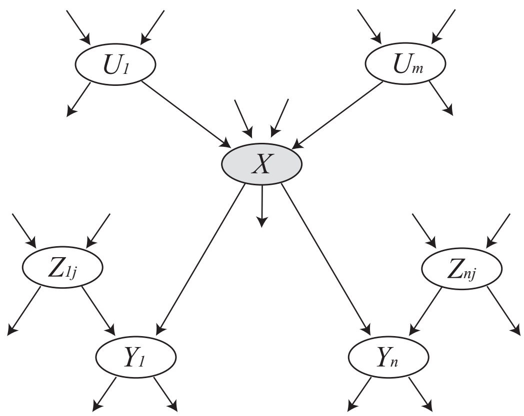
井庭崇
iba@sfc.keio.ac.jp

1　単結合ベイジアンネット：多重木における厳密推論

ここでは、ネットワークが「多重木」(polytree) と呼ばれる単純な構造をしているときの推論の方法を取り上げる。多重木のネットワークとは、ノードのどの組の間にもせいぜい１本の経路があるだけのネットワーク、つまり、単結合ネットワークのことである。

質問ノードとして X、証拠ノードとしていくつかの集合 $\mathbf{E}$ （X は含まれない）があるとしよう。ここで行いたいのは、$P(X|\mathbf{E})$ を計算して、$Bel(X)$ を更新することである。

下図は、ノード X と、親ノード U_i と子ノード Y_j、そして子ノードの親ノード Z_{ij}、がリンクで結ばれている一般的な多重木の構造である。



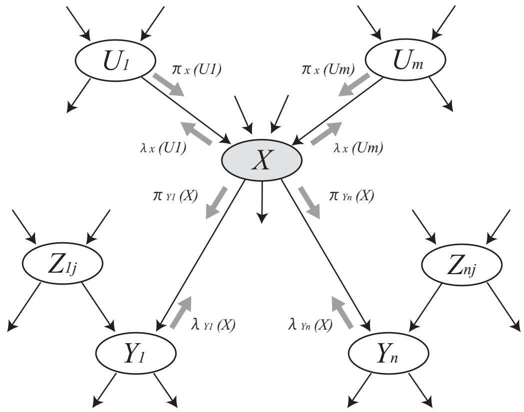
X のための局所的な信念更新のためには、ネットワークの他のすべての部分からの証拠を組み込まなければならない。この図から、私たちは、証拠を次のように分けることができることがわかる。

- ノード $U_1, ..., U_m$ を通じて X につながっている証拠ノードからの X に対する**予測的な裏付け**
- ノード $Y_1, ..., Y_n$ を通じて X につながっている証拠ノードからの X に対する**診断的な裏づけ**

1.1　Kim と Pearl によるメッセージ・パッシング・アルゴリズム

基本的なアイデアは、アルゴリズムのそれぞれイテレーション（繰り返し）で、$\lambda(X)$、$\pi(X)$、CPT (式 3.1) の３種類のパラメータを用いて、$Bel(X)$ を局所的に更新するということである。

$\lambda(X)$ と $\pi(X)$ は、X の親と子のそれぞれから受け取った π と λ のメッセージを用いて計算される。なお、π と λ のメッセージは、X から近傍へも送られるので、その近傍の更新を行うことができる。



このアルゴリズムには、次の３種類のパラメータが必要となる。

- 入ってくるリンク $U_i \to X$ がもたらす予測的裏づけ π の現在の強さ

$$\pi_X(U_i) = P(U_i|E_{U_i \backslash X})$$

　ここで、$E_{U_i \backslash X}$ は、X を除く、U_i につながっているすべての証拠のこと。

- 出ていくリンク $X \to Y_j$ がもたらす診断的裏づけ λ の現在の強さ

$$\lambda_{Y_j}(X) = P(E_{Y_j \backslash X}|X)$$

　ここで、$E_{Y_j \backslash X}$ は、X を除く、親から Y_j につながっているすべての証拠のこと。

- 固定の CPT $P(X|U_i, ..., U_n)$ （X の直接的な親に関して）

これらのパラメータは、以下に述べる３つのステップにおいて、局所的な信念更新を行うのに用いられる。

1. 信念更新 (belief updating)
2. ボトムアップ伝搬 (Bottom-up propagation)
3. トップダウン伝搬 (Top-down propagation)

※以下では、x_i はノード X の i 番目の状態を意味し、、u_i は、$u_1, ..., u_n$ というノードを表すために用いられる点に注意が必要である。

1.1.1　信念更新 (belief updating)

ノード X についての信念更新は、すべての子ノードと親ノードから来たメッセージ（それらの信念パラメータの変化を示している）によって活性化される。ノード X が活性化されたときは、親ノードからのメッセージ $\pi_X(U_i)$ と、子ノードからのメッセージ $\lambda_{Y_j}(X)$ を調べる。

$$Bel(x_i) = \alpha\lambda(x_i)\pi(x_i) \tag{1}$$

ここで、$\lambda(x_i)$ パラメータの計算方法は、次のようになる。

$$\lambda(x_i) = \begin{cases} 1 & (\text{証拠が } X = x_i \text{ の場合}) \\ 0 & (\text{証拠がほかの } x_j \text{ の場合}) \\ \prod_j \lambda_{Y_j}(x_i) & (\text{それ以外}) \end{cases} \tag{2}$$

この $\lambda(x_i)$ パラメータを通じて証拠が入力される。もし x_i が証拠値ならば 1 になり、もし証拠がほかの値 x_j であれば 0 になる。もし X に入力された証拠がなければ、子ノードから受け取った λ メッセージのすべてを掛け合わせたものになる。

$$\pi(x_i) = \sum_{u_1,...,u_n} P(x_i|u_1,...,u_n) \prod_i \pi_X(u_i) \tag{3}$$

$\pi(x_i)$ パラメータは、親ノードからの π メッセージと CPT の積である。
α は、$\sum_{x_i} Bel(X = x_i) = 1$ となるように正規化するための定数である。

1.1.2　ボトムアップ伝搬 (Bottom-up propagation)

ノード X は、親ノードに送るための新しい λ メッセージを計算する。

$$\lambda_X(U_i) = \sum_{x_i} \lambda(x_i) \sum_{u_k:k\neq i} P(x_i|u_1,...,u_n) \prod_{k\neq i} \pi_x(u_k) \tag{4}$$

ひとつの親ノードへの λ メッセージは、次の情報を統合している。

1. λ メッセージを通じて子ノードから来た情報で、$\lambda(X)$ パラメータに集約される情報
2. CPT の値
3. 親ノードから受け取った π メッセージ

1.1.3　トップダウン伝搬 (Top-down propagation)

ノード X は、子ノードに送るための新しい π メッセージを計算する。

$$\pi_{Y_j}(x_i) = \begin{cases} 1 & (\text{証拠値 } x_i \text{ が入力された場合}) \\ 0 & (\text{証拠がほかの } x_j \text{ の値の場合}) \\ \alpha[\prod_{k\neq j} \lambda_{Y_k}(x_i)] \sum_{u_1,...,u_n} P(X_i|u_1,...,u_n) \prod_i \pi_X(U_i) = \frac{\alpha Bel(x_i)}{\lambda_{Y_j}(X_i)} & (\text{それ以外}) \end{cases} \tag{5}$$

子ノード Y_j へ向かう $\pi_{Y_j}(x_i)$ メッセージは、x_i が証拠値の場合に 1、証拠がほかの値 x_j の場合に 0 になる。X に証拠が入力されなかった場合には、そのメッセージは、次の情報を統合したものになる。

1. Y_j 以外の子ノードからの情報
2. CPT
3. 親ノードから受け取った π メッセージ

1.1.4　初期化と表記上の注意

このアルゴリズムは、証拠を入力する前に、次の初期化が必要である。

- λ 値、λ メッセージ、π メッセージを 1 に設定する。
- 根ノード：もしノード W が親ノードをもっていなければ、$\pi(W)$ を事前確率 $P(W)$ に設定する。

このメッセージ・パッシング・アルゴリズムは、証拠が得られる前にも、ネットワーク上のすべてのノードの信念の計算に用いることができる。

具体的な証拠 $W = w_i$ が得られたときには、ノード W は $w_1, w_2, ..., w_n$ の値が与えられる。

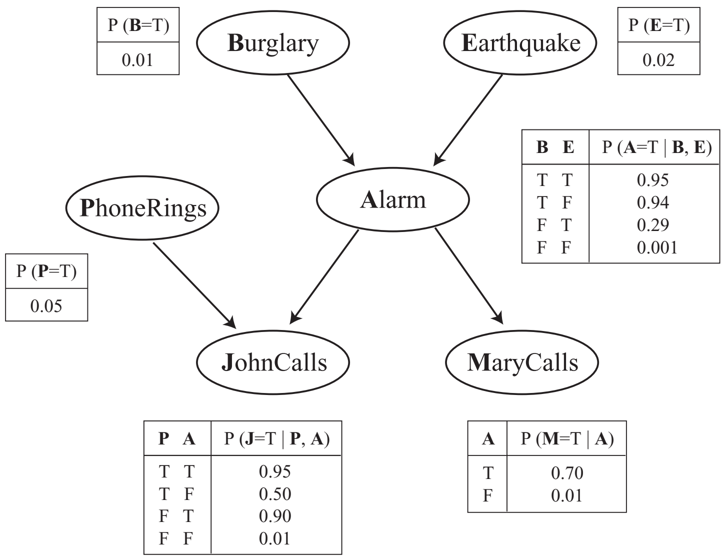
- $\lambda(W) = (0, 0, ..., 0, 1, 0, ..., 0)$ のように、i 番目の位置を 1 にする。

メッセージを表すこの π/λ 表記は、Kim と Pearl によって導入されたものだが、当初は混乱を生んだ。メッセージの両者の種類における書き方は、$\pi_{Child}(Parent)$ と $\lambda_{Child}(Parent)$ である。なので、

- π メッセージは、親ノードから子ノードへ、弧の方向に送られる。表記は、$\pi_{Receiver}(Sender)$。
- λ メッセージは、子ノードから親ノードへ、弧と逆の方向に送られる。表記は、$\lambda_{Sender}(Receiver)$。

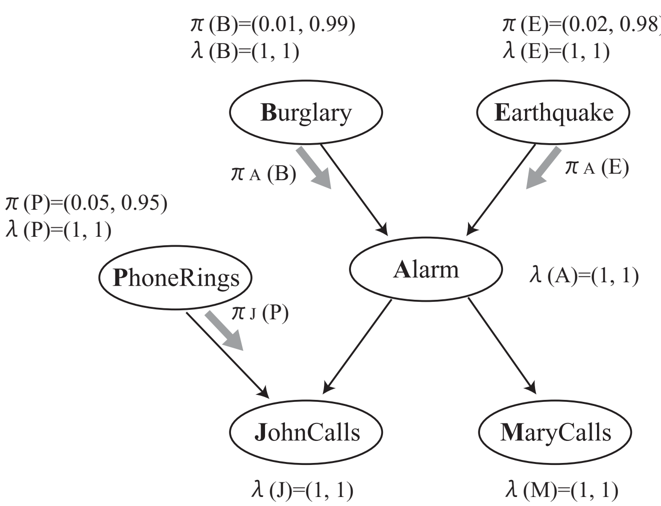
π は、事前確率の役割も担っており、λ はベイズの定理における尤度を担っている。

1.2 例：拡張された地震モデル



この拡張された地震モデルにおいて、最初は証拠なしで計算し、その後ノード M に証拠を入力してメッセージパッシングによる更新を行っていく。

まず最初に、証拠が入力される前に、いくつかのパラメータを初期化する必要がある。事前確率から $\pi(B)$、$\pi(E)$、$\pi(P)$ を初期化し、証拠なしの葉ノードとして、$\lambda(J)$、$\lambda(M)$ を (1, 1) に初期化する。

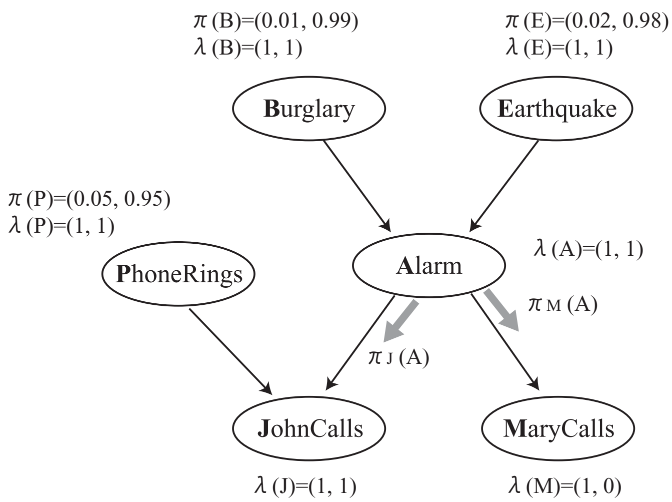


証拠入手前の伝搬のあいだ、診断メッセージ (λ メッセージ) はすべて、単位ベクトルになる。更新関数からわかるように、これらのメッセージを掛けるだけで、つまり単位ベクトルを掛けても、他のパラメータやメッセージは変化しない。

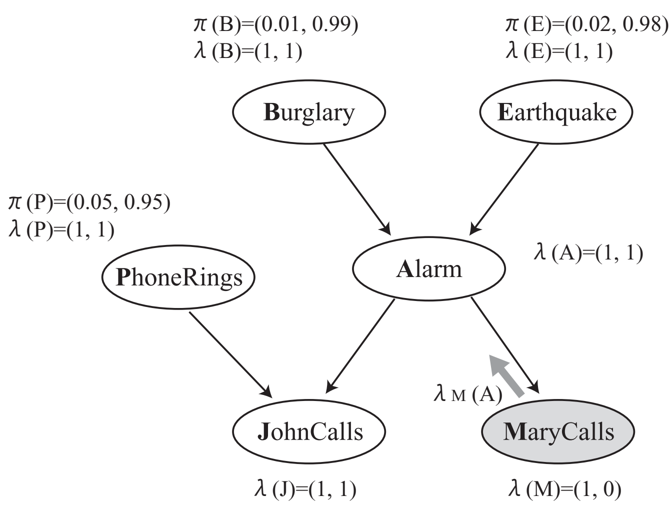
新しいメッセージ $\pi_A(B)$、$\pi_A(E)$、$\pi_J(P)$、が、トップダウン伝搬の式 (5) を用いて計算され、送られると同時に、信念更新関数 (1) を用いて $Bel(B)$、$Bel(E)$、$Bel(P)$ が計算される。

ノード A は、すべてのメッセージを親ノードから受け取っているので、$Bel(A)$ を更新できる。そして、子ノード J と M に渡すための π メッセージを計算する。あとは、P から π メッセージを受け取れば、この段階で、$Bel(J)$ を更新することができるが、すでに A から π メッセージを受け取っているので、このサイクルでは、それは行わない。

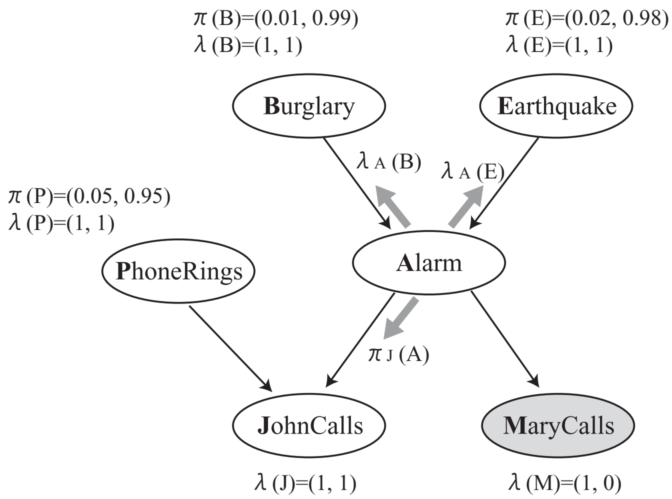
次のメッセージ伝搬フェーズでは、すべての π メッセージはそれぞれの親からもらったので、$Bel(J)$ と $Bel(M)$ を計算することができる。



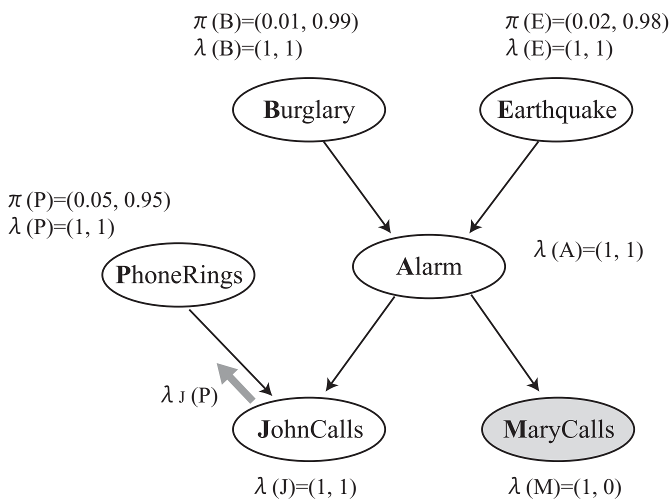
次に、証拠 $M = T$ を、$\lambda(M) = (1, 0)$ として、ネットワーク上で伝搬させる。まず、メッセージ $\lambda_M(A)$ が計算され、A に送られる。



$\lambda(A)$ と、その結果として $Bel(A)$ が再計算され、新しいメッセージが、A の親ノードへ、$\lambda_A(B)$ と $\lambda_A(E)$ として送られ、子ノード J へ、$\pi_J(A)$ として送られる。



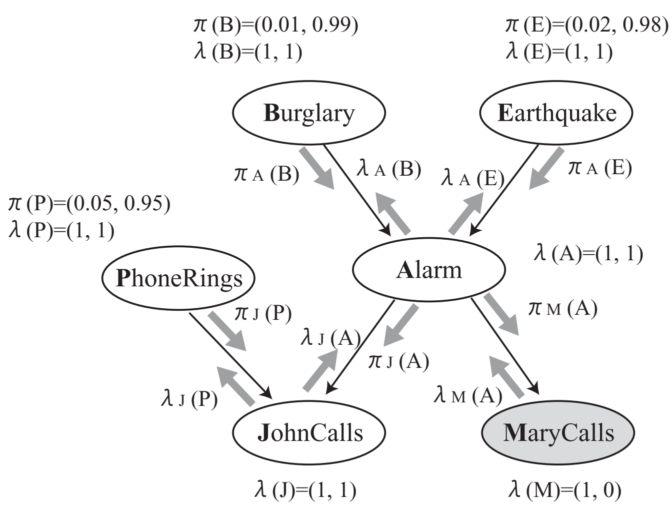
この新しいメッセージを受け取った結果、$Bel(B)$、$Bel(E)$、$Bel(J)$ が再計算され、最後のメッセージ $\lambda_J(B)$ を計算して、J から P へ送信する。



最後の計算は、$Bel(P)$ の更新である。

伝搬回数の最少回数は、3 ステップである。これは、ノード P の証拠ノード M からの距離を表している。

結果として、次のように上下にメッセージパッシングが行われたことになる。

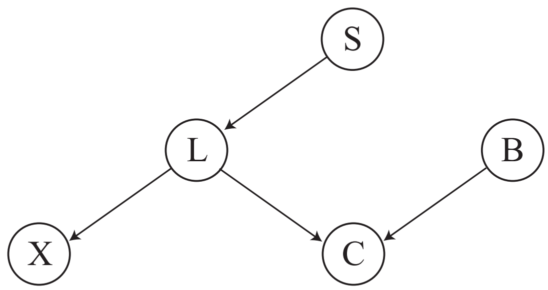


1.3 練習問題

医療診断のためのベイジアンネットワークを考える。

- B = Bronchitis (気管支炎)
- S = Smoker (喫煙家)
- C = Cough (咳払い)
- X = Positive X-ray (X 線検査陽性)
- L = Lung cancer (肺ガン)

すべてのノードはブーリアン (True or False)
患者が喫煙家である事前確率は 0.25、患者が気管支炎である事前確率は 0.05。

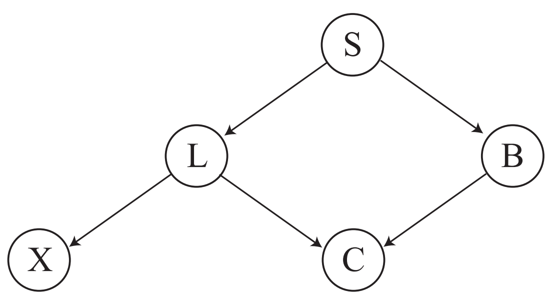


いま証拠として、患者が X 線検査が陽性であることがわかったとすると、

(1) ノード S、B、C、X について、π と λ を書きなさい。

(2) π と λ によるメッセージパッシング・アルゴリズムにおける、データ伝搬の 3 段階を示しなさい。

(3) もし医者が、喫煙が肺ガンだけでなく気管支炎も引き起こすと考えるとした場合について考えてみよう。ネットワークは次のようになる。なぜ、多重木のメッセージパッシングアルゴリズムは、このようなネットワークに対して、信念の更新を行うのに不適当なのかを説明してください。



※今回の資料で取り上げた解説・例は、文献 [1] にもとづいている。

参考文献

[1] Kevin B. Korb and Ann E. Nicholson. *Bayesian Artificial Intelligence.* Chapman & Hall / CRC, 2004.