CERIES 東北大学 電気情報系 グローバルCOEプログラム
Center of Education and Research for Information Electronics Systems

DEX-SMI

Statistical-Mechanical Approach to Probabilistic Information Processing

統計的モデリングとベイジアンネットの原理

東北大学 大学院情報科学研究科
応用情報科学専攻
田中 和之(Kazuyuki Tanaka)
kazu@smapip.is.tohoku.ac.jp
http://www.smapip.is.tohoku.ac.jp/~kazu/

本講演の参考図書

- 田中和之著: 確率モデルによる画像処理技術入門, 森北出版, 2006.
- 田中和之編著: 臨時別冊・数理科学SGCライブラリ「確率的情報処理と統計力学 ---様々なアプローチとそのチュートリアル」, サイエンス社, 2006.
- 田中和之: 大規模確率場における予測と推論, 電子情報通信学会誌, Vol.88, No.9, pp.698-702, September 2005.

不確実性を伴う情報処理の数理モデル

不確実性の数学的表現→確率・統計

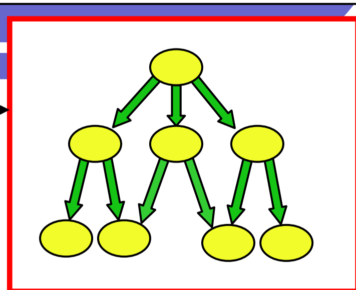
確率推論

モデル化

ネットワーク構造をもつ
数理モデル（ベイジアンネットワーク）

医療診断
故障診断
危険予知

ノードは事象，矢印は
条件付き確率に対応



閉路のあるグラフ

不確実性を伴うデータに耐えうる推論システム

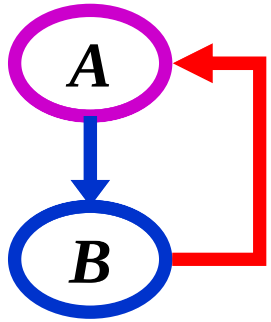
単純な機能を持つたくさんの要素が関連し合い，互いに協力して複雑・高度な機能を生み出す．

重要な概念のひとつ

ベイズの公式の導出

$$\Pr\{A,B\}=\Pr\{A|B\}\Pr\{B\}$$

結合確率
(Joint Probability)



$$\Pr\{A,B\}=\Pr\{B|A\}\Pr\{A\}$$

$$\boxed{\Pr\{A|B\}}=\frac{\Pr\{A,B\}}{\Pr\{B\}}=\frac{\boxed{\Pr\{B|A\}}\boxed{\Pr\{A\}}}{\Pr\{B\}}$$

事後確率
(Posterior Probability)

事前確率
(Prior Probability)

$$\Pr\{B\}=\sum_{A}\Pr\{A,B\}$$

周辺確率
(Marginal Probability)

ベイズの公式による確率的推論の例(1)

A 教授はたいへん謹厳でこわい人で，機嫌の悪いときが 3/4 を占め，機嫌のよい期間はわずかの 1/4 にすぎない．
教授には美人の秘書がいるが，よく観察してみると，教授の機嫌のよいときは，8 回のうち 7 回までは彼女も機嫌がよく，悪いのは 8 回中 1 回にすぎない．
教授の機嫌の悪いときで，彼女の機嫌のよいときは 4 回に 1 回である．
秘書の機嫌からベイズの公式を使って教授の機嫌を確率的に推論することができる．

甘利俊一：情報理論 (ダイヤモンド社，1970) より

ベイズの公式による確率的推論の例（2）

データからヒストグラムを計算し，確率表を作成

教授は機嫌の悪いときが 3/4 を占め，機嫌のよい期間はわずかの 1/4 にすぎない．

$$\Pr\{教授機嫌良い\}=\frac{1}{4} \qquad \Pr\{教授機嫌悪い\}=\frac{3}{4}$$

$\Pr\{教授機嫌\}$

教授機嫌良い	1/4
教授機嫌悪い	3/4

教授の機嫌のよいときは，8 回のうち 7 回までは彼女も機嫌がよく，悪いのは 8 回中 1 回にすぎない．

$$\Pr\{秘書機嫌良い|教授機嫌良い\}=\frac{7}{8}$$

教授の機嫌の悪いときで，彼女の機嫌のよいときは 4 回に 1 回である．

$$\Pr\{秘書機嫌良い|教授機嫌悪い\}=\frac{1}{4}$$

$\Pr\{秘書機嫌|教授機嫌\}$

	秘書機嫌良い	秘書機嫌悪い
教授機嫌良い	7/8	1/8
教授機嫌悪い	1/4	3/4

ベイズの公式による確率的推論の例(3)

$$\Pr\{B\}=\sum_{A}\Pr\{A,B\}=\sum_{A}\Pr\{B\mid A\}\Pr\{A\}$$

$$\Pr\{\text{秘書機嫌良し}\}$$
$$=\Pr\{\text{秘書機嫌良し}\mid\text{教授機嫌良し}\}\Pr\{\text{教授機嫌良し}\}$$
$$+\Pr\{\text{秘書機嫌良し}\mid\text{教授機嫌悪い}\}\Pr\{\text{教授機嫌悪い}\}$$
$$=\frac{7}{8}\times\frac{1}{4}+\frac{1}{4}\times\frac{3}{4}=\frac{13}{32}$$

教授は機嫌の悪いときが 3/4 を占め，機嫌のよい期間はわずかの 1/4 にすぎない．

$$\Pr\{\text{教授機嫌良い}\}=\frac{1}{4}$$

$$\Pr\{\text{教授機嫌悪い}\}=\frac{3}{4}$$

$$\Pr\{\text{秘書機嫌良い}\mid\text{教授機嫌悪い}\}=\frac{1}{4}$$

$$\Pr\{\text{秘書機嫌良い}\mid\text{教授機嫌良い}\}=\frac{7}{8}$$

教授の機嫌のよいときは，8 回のうち 7 回までは彼女も機嫌がよく，悪いのは 8 回中 1 回にすぎない．

ベイズの公式による確率的推論の例(4)

$$\Pr\{A,B\}=\frac{\Pr\{B|A\}\Pr\{A\}}{\Pr\{B\}}$$

$$\Pr\{教授機嫌良し|秘書機嫌良し\}$$

$$=\frac{\Pr\{秘書機嫌良し|教授機嫌良し\}\Pr\{教授機嫌良し\}}{\Pr\{秘書機嫌良し\}}=\frac{\frac{7}{8}\times\frac{1}{4}}{\frac{13}{32}}=\frac{7}{13}$$

$$\Pr\{秘書機嫌良い|教授機嫌良い\}=\frac{7}{8}$$

$$\Pr\{秘書機嫌良い\}=\frac{13}{32}$$

$$\Pr\{教授機嫌良い\}=\frac{1}{4}$$

教授の機嫌のよいときは，8回のうち7回までは彼女も機嫌がよく，悪いのは8回中1回にすぎない．

教授は機嫌の悪いときが3/4を占め，機嫌のよい期間はわずかの1/4にすぎない．

ベイズの公式による確率的推論の例(5)

Prior Probability Table

Professor = Good Mood	**1/4**
Professor = Bad Mood	**3/4**

$\Pr\{\text{Professor}\}$

Conditional Probability Table

	Secretary =Good Mood	**Secretary =Bad Mood**
Professor =Good Mood	**7/8**	**1/8**
Professor =Bad Mood	**1/4**	**3/4**

$\Pr\{\text{Secretary}\,|\,\text{Professor}\}$

$$\Pr\{\text{Professor}\,|\,\text{Secretary}\} = \frac{\Pr\{\text{Secretary}\,|\,\text{Professor}\}\Pr\{\text{Professor}\}}{\Pr\{\text{Secretary}\}}$$

	Professor =Good Mood	**Professor =Bad Mood**
Secretary =Good Mood	**7/13**	**6/13**
Secretary =Bad Mood	**1/19**	**18/19**

ベイジアンネットによる確率推論の例(1)

Cloudy
A=T or F

$\Pr\{A\}$

A=T	0.5
A=F	0.5

Sprinkler
C=T or F

$\Pr\{C \mid A\}$

	C=T	C=F
A=T	0.5	0.5
A=F	0.1	0.9

Rain
B=T or F

$\Pr\{B \mid A\}$

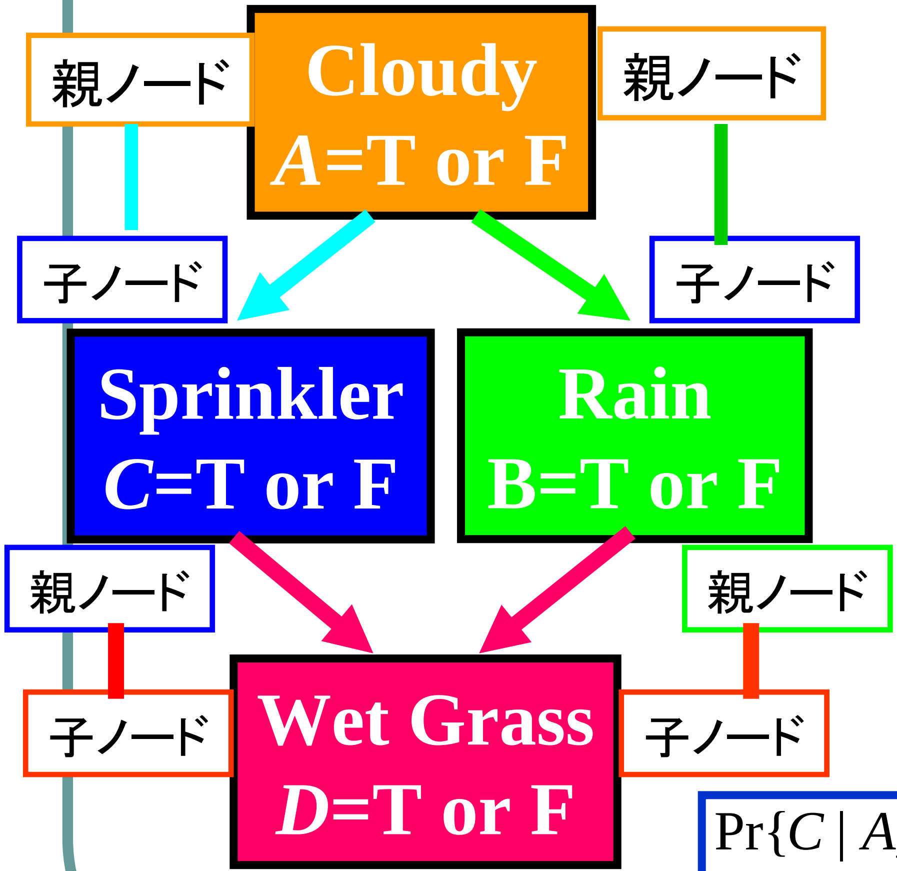
	B=T	B=F
A=T	0.2	0.8
A=F	0.8	0.2

Wet Grass
D=T or F

$\Pr\{D \mid B, C\}$

		A=T	A=F
B=T	C=T	0.99	0.01
B=T	C=F	0.9	0.1
B=F	C=T	0.9	0.1
B=F	C=F	0	1

ベイジアンネットによる確率推論の例(2)



$$
\begin{aligned}
&\Pr\{A,B,C,D\}\\
&=\Pr\{D|A,B,C\}\Pr\{A,B,C\}\\
&=\Pr\{D|A,B,C\}\Pr\{C|A,B\}\\
&\qquad\times\Pr\{A,B\}\\
&=\Pr\{D|A,B,C\}\Pr\{C|A,B\}\\
&\qquad\times\Pr\{B|A\}\Pr\{A\}\\
&=\Pr\{D|B,C\}\Pr\{C|A\}\\
&\qquad\times\Pr\{B\mid A\}\Pr\{A\}
\end{aligned}
$$

$$\Pr\{C\mid A,B\}=\Pr\{C\mid B\}$$

$$\Pr\{D\mid A,B,C\}=\Pr\{D\mid B,C\}$$

因果独立

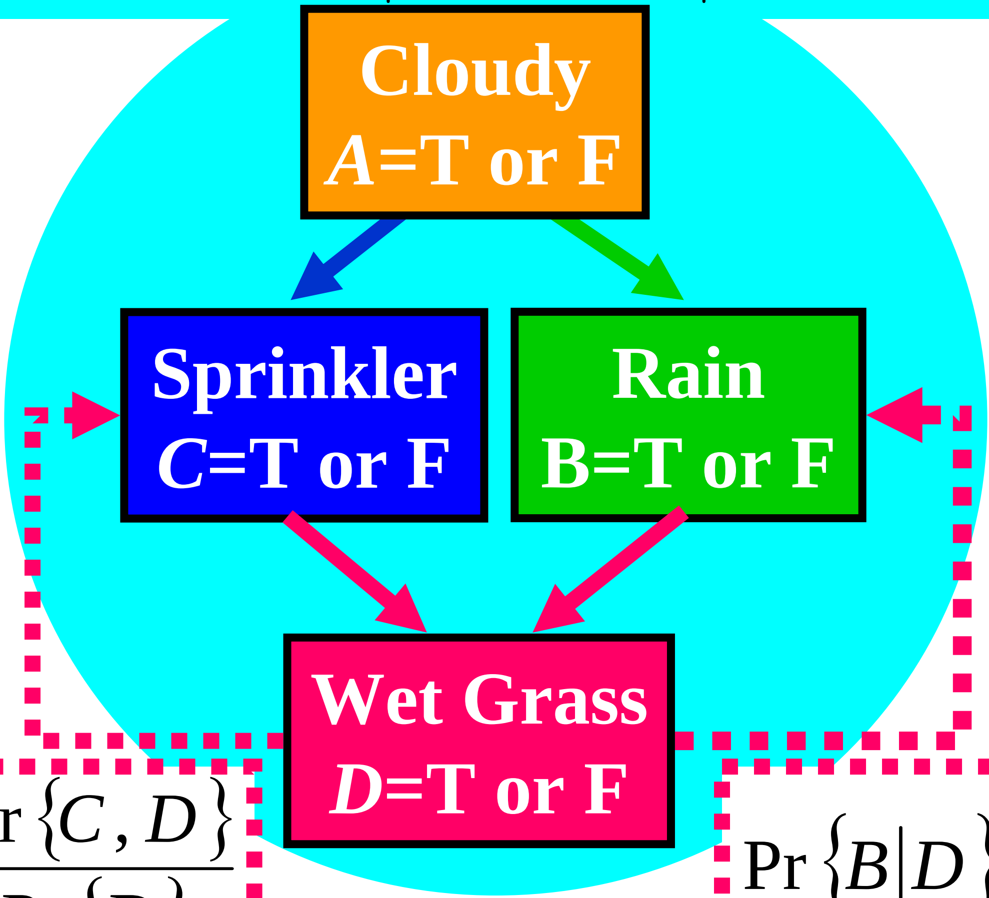
ベイジアンネットによる確率推論の例(3)

$$\Pr\{A,B,C,D\}=\Pr\{D|B,C\}\Pr\{C|A\}\Pr\{B\mid A\}\Pr\{A\}$$

周辺確率

$$\Pr\{C,D\}=\sum_{A}\sum_{B}\Pr\{A,B,C,D\}$$

$$\Pr\{D\}=\sum_{A}\sum_{B}\sum_{C}\Pr\{A,B,C,D\}$$



周辺確率

$$\Pr\{B,D\}=\sum_{A}\sum_{C}\Pr\{A,B,C,D\}$$

$$\Pr\{D\}=\sum_{A}\sum_{B}\sum_{C}\Pr\{A,B,C,D\}$$

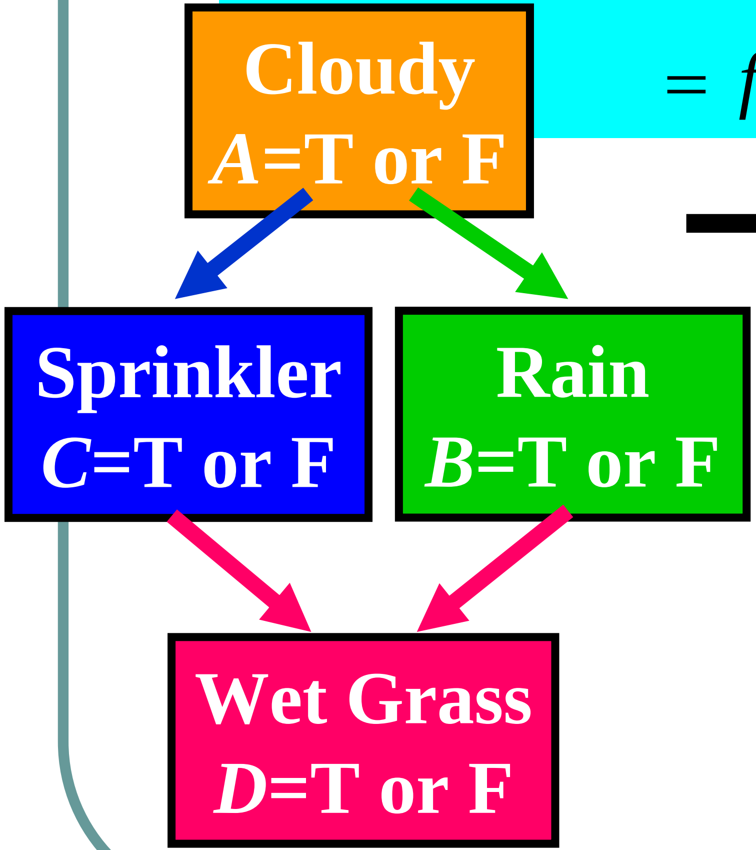
$$\Pr\{C|D\}=\frac{\Pr\{C,D\}}{\Pr\{D\}}$$

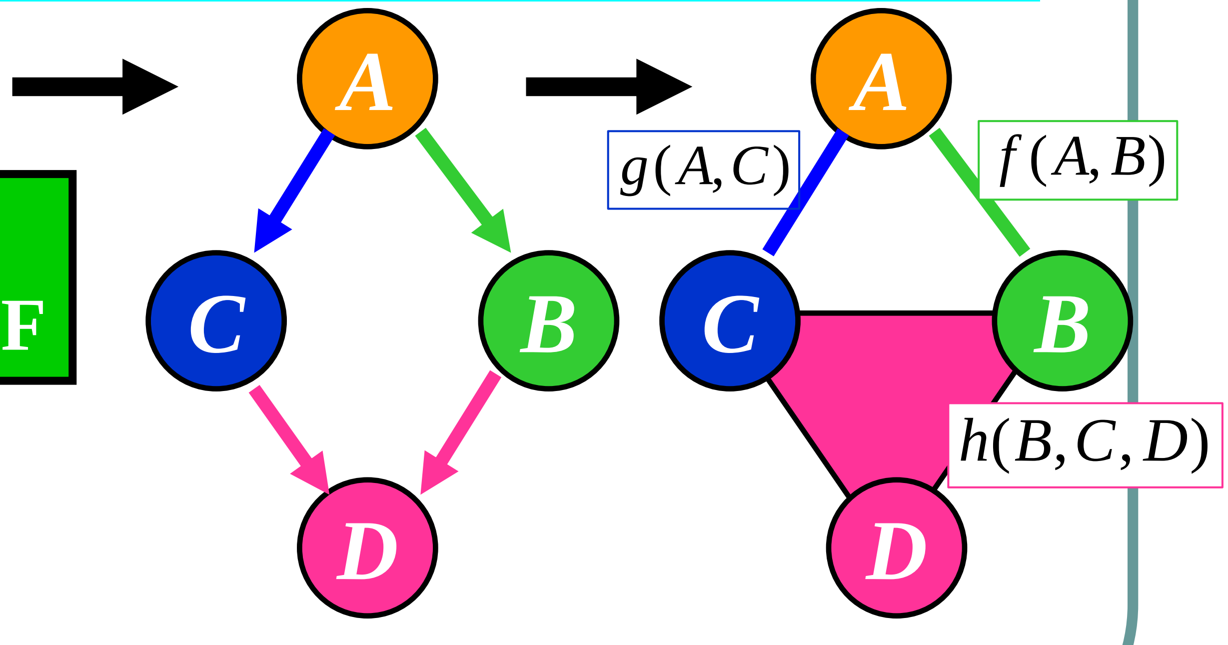
$$\Pr\{B|D\}=\frac{\Pr\{B,D\}}{\Pr\{D\}}$$

ベイジアンネットによる確率推論の例（4）

$$\Pr\{A,B,C,D\}=\underbrace{\Pr\{D|B,C\}}_{h(B,C,D)}\underbrace{\Pr\{C|A\}}_{g(A,C)}\underbrace{\Pr\{B\,|\,A\}\Pr\{A\}}_{f(A,B)}$$

$$=f(A,B)g(A,C)h(B,C,D)$$





有向グラフ　　　無向グラフ

より複雑なベイジアンネット

$$\Pr\{A_2\} \equiv \sum_{A_1}\sum_{A_3}\cdots\sum_{A_N}\Pr\{A_1, A_2, A_3, \cdots, A_N\}$$

定義に基づいて厳密に計算するプログラム

$$A_i = \text{True or False}$$

ノード数とともに指数関数的に計算量が増加 $O(e^N)$

$$2^{N-1} = \exp((N-1)\ln 2)$$

```
Pr{ A_2 } ← 0;
for( A_1 = T or F){
 for( A_3 = T or F){
  for( A_4 = T or F){
   ⋮
    for( A_N = T or F){
     Pr{ A_2 } ← Pr{ A_2 } + Pr{ A_1,A_2 ,⋯ , A_N };
    }
   ⋮
  }
 }
}
```

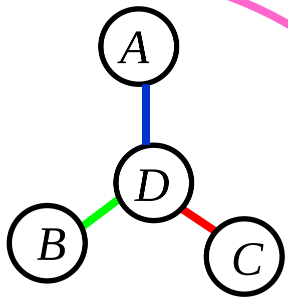
N-1 重の for 文

このプログラムでは
L=10個のノードで1秒かかるとしたら
L=20個で約17分, L=30個で約12日,
L=40個で約34年かかる.

扱いやすい確率モデルのグラフ表現

扱いやすい確率モデルの数理構造

$$\sum_{A=\mathrm{T,F}}\sum_{B=\mathrm{T,F}}\sum_{C=\mathrm{T,F}} f(A,D)\,g(B,D)\,h(C,D) \longrightarrow \sum_{A=\mathrm{T,F}}\sum_{B=\mathrm{T,F}}\sum_{C=\mathrm{T,F}}$$



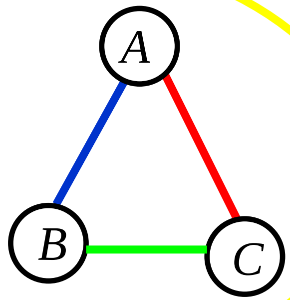
$$=\left(\sum_{A=\mathrm{T,F}} f(A,D)\right)\left(\sum_{B=\mathrm{T,F}} g(B,D)\right)\left(\sum_{C=\mathrm{T,F}} h(C,D)\right)$$

別々に和を計算できる

木構造をもつグラフ表現

扱いやすくない確率モデルの数理構造

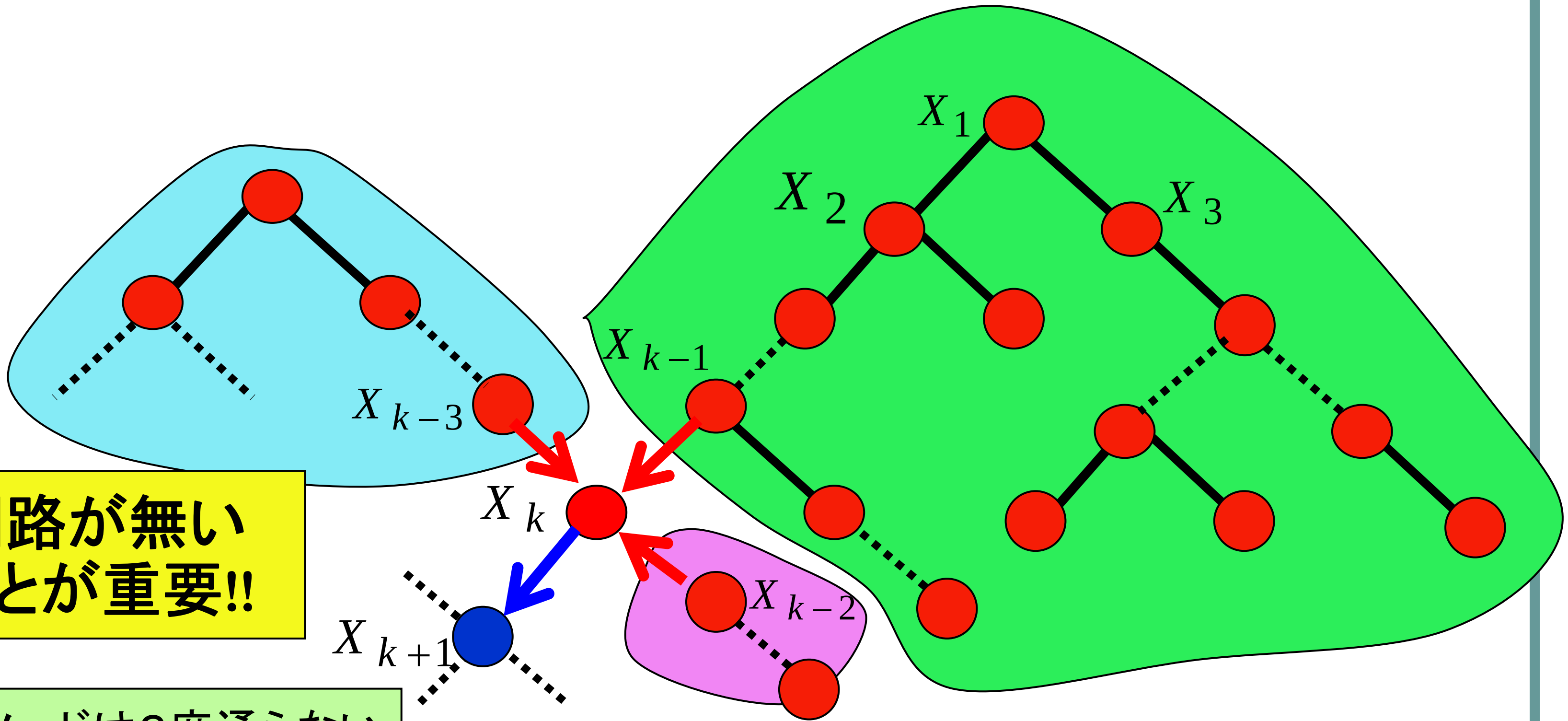
$$\sum_{A=\mathrm{T,F}}\sum_{B=\mathrm{T,F}}\sum_{C=\mathrm{T,F}} f(A,B)\,g(B,C)\,h(C,A) \longrightarrow \sum_{A=\mathrm{T,F}}\sum_{B=\mathrm{T,F}}\sum_{C=\mathrm{T,F}}$$



別々に和を計算することが難しい

閉路を含むグラフ表現

閉路のないグラフ上の確率伝搬法



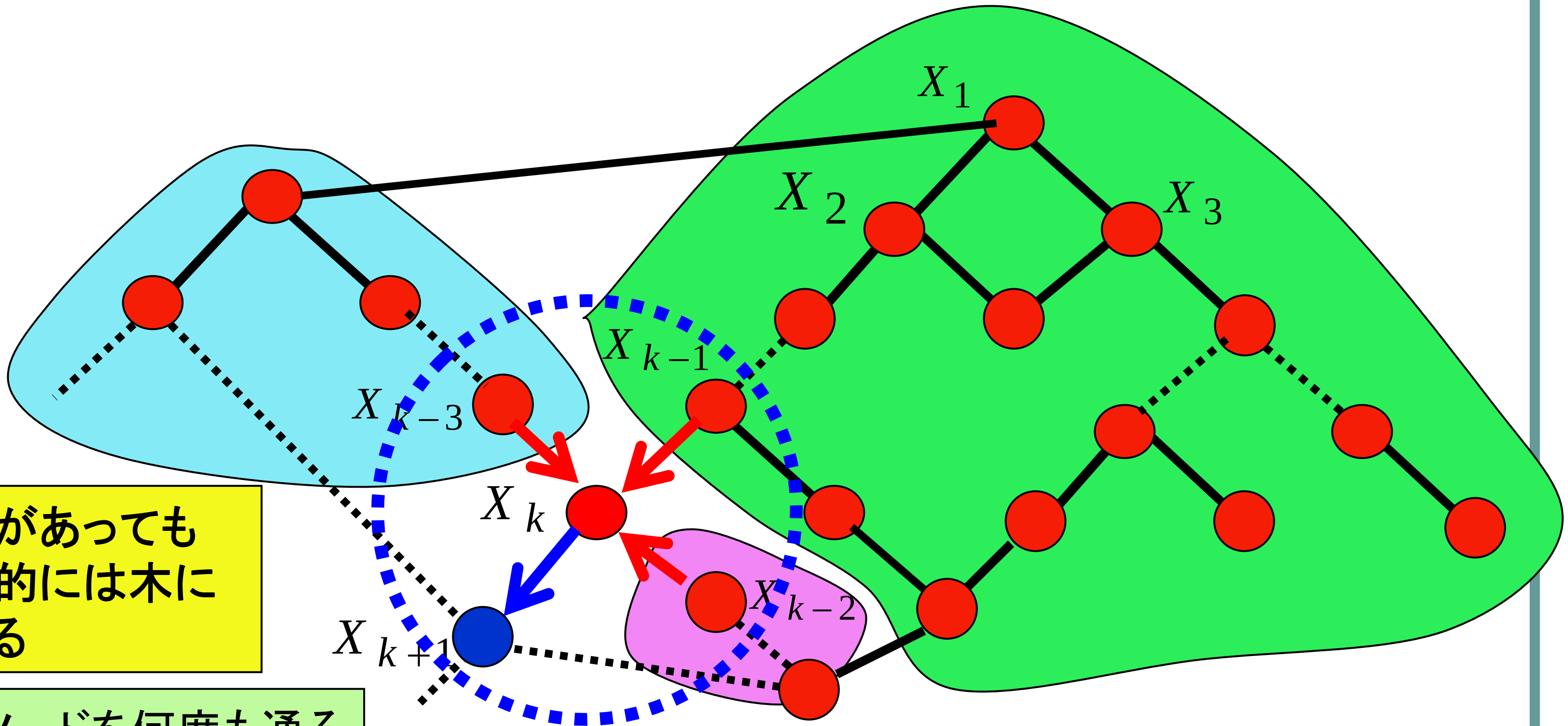
閉路が無いことが重要!!

同じノードは2度通らない

$$\boxed{M_{k\to k+1}(x_{k+1})} = \sum_{x_k} \boxed{M_{k-1\to k}(x_k) M_{k-2\to k}(x_k) M_{k-3\to k}(x_k)} \Phi_{k,k+1}(x_k, x_{k+1})$$

メッセージに対する漸化式（Message Passing Rule）

閉路のあるグラフ上の確率伝搬法



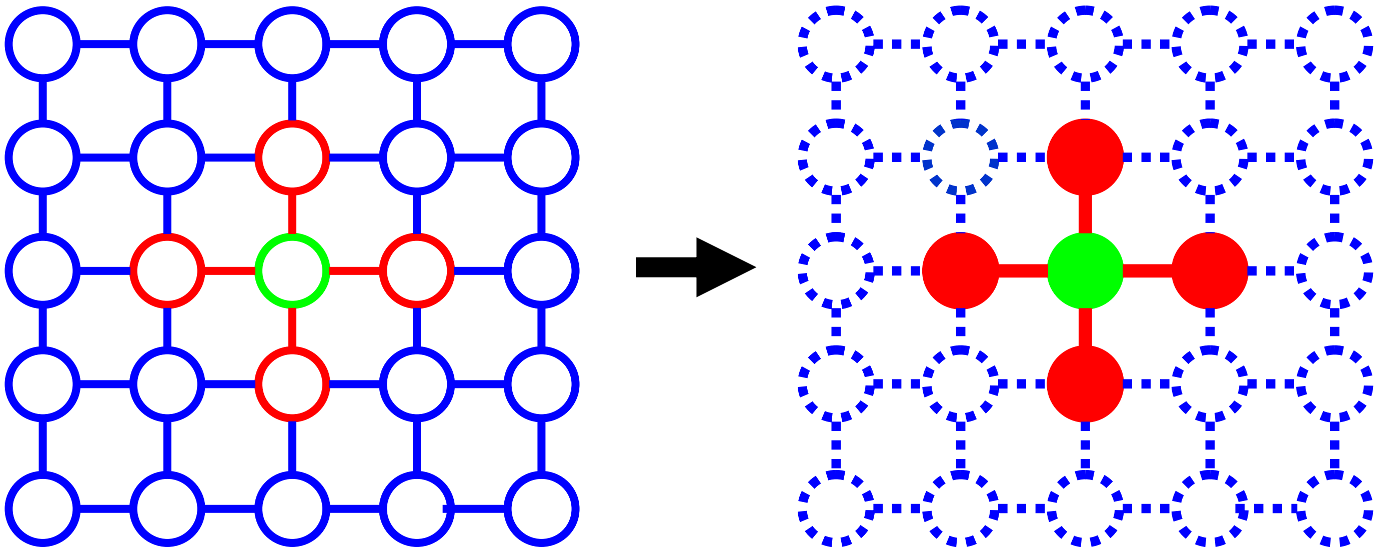
閉路があっても局所的には木にみえる

同じノードを何度も通る

$$\boxed{M_{k\to k+1}(x_{k+1})} \cong \sum_{x_k} \boxed{M_{k-1\to k}(x_k)M_{k-2\to k}(x_k)M_{k-3\to k}(x_k)}\Phi_{k,k+1}(x_k, x_{k+1})$$

メッセージに対する固定点方程式(Message Passing Rule)

閉路のあるグラフ上の確率伝搬法(Belief Propagation)

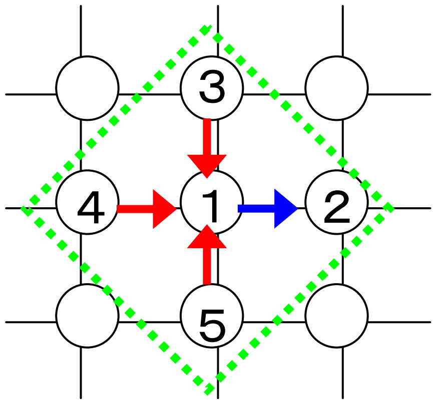


着目ノードとその近傍ノードだけを残すと木構造になる.

確率伝搬法(Belief Propagation)の統計的近似アルゴリズムとしての転用

閉路のあるグラフ上の確率伝搬法

$$\boxed{M_{1\to 2}(f_2)} = \sum_{z_1} \Phi_{12}(z_1, f_2) \boxed{M_{3\to 1}(z_1)} \boxed{M_{4\to 1}(z_1)} \boxed{M_{5\to 1}(z_1)}$$



閉路のあるグラフ上でも局所的な構造だけに着目してアルゴリムを構成することは可能.
ただし，得られる結果は厳密ではなく近似アルゴリズム

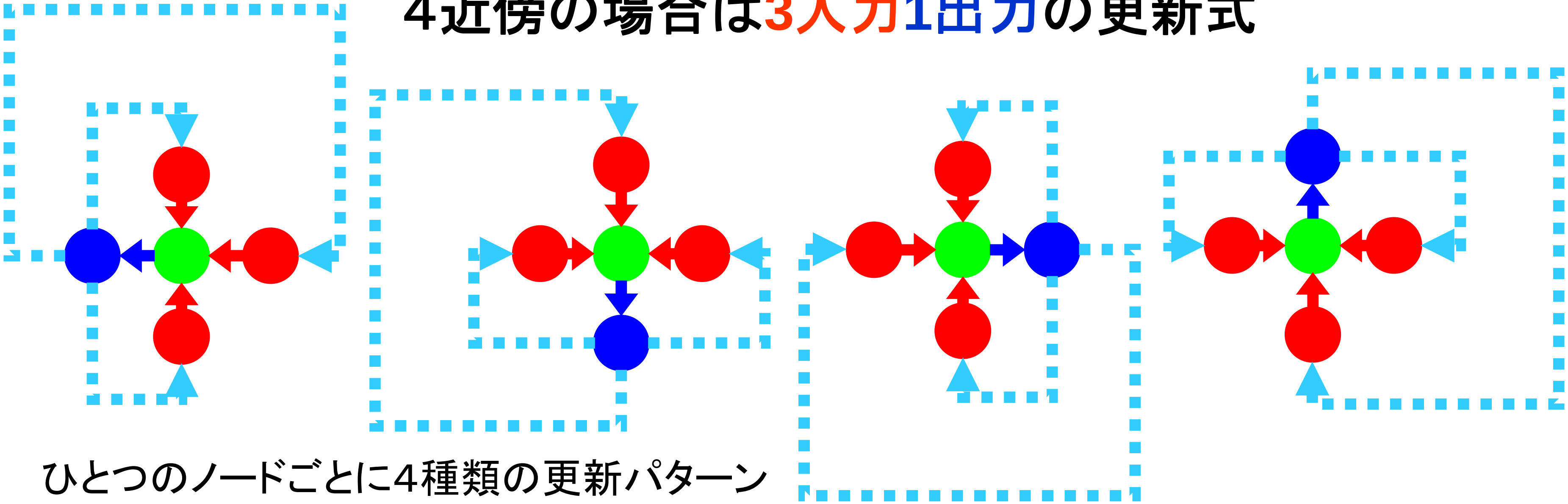
$$\boxed{\vec{M} = \vec{\Psi}\left(\vec{M}\right)}$$

メッセージに対する固定点方程式

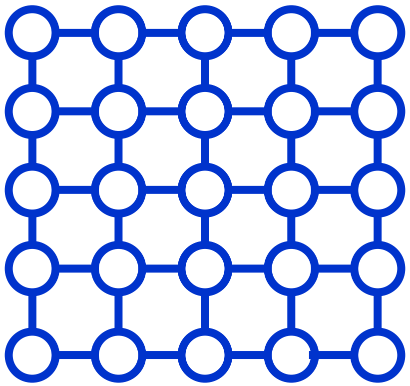
平均，分散,共分散はこのメッセージを使ってあらわされる

閉路のあるグラフ上の確率伝搬法

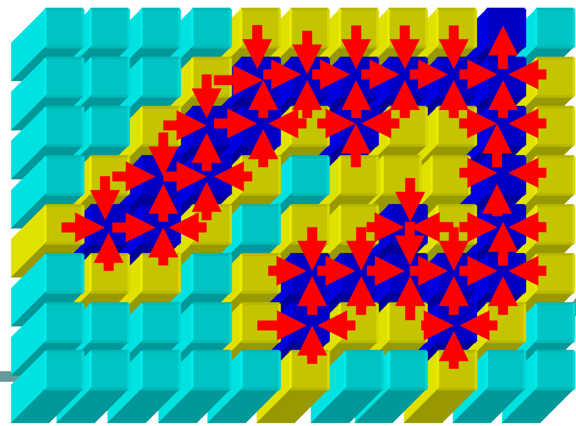
4近傍の場合は3入力1出力の更新式



ひとつのノードごとに4種類の更新パターン



ノード上での
動作の様子
の一例



●固定点方程式と反復法

固定点方程式

$$\vec{M}^* = \vec{\Psi}\left(\vec{M}^*\right)$$

反復法

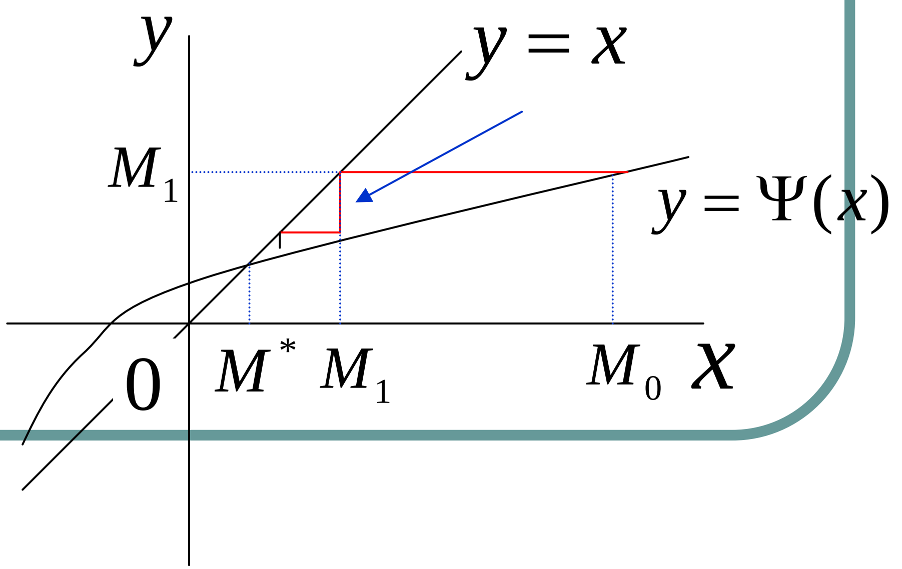
繰り返し出力を入力に入れることにより，固定点方程式の解が数値的に得られる．

$$\vec{M}_1 \leftarrow \vec{\Psi}\left(\vec{M}_0\right)$$

$$\vec{M}_2 \leftarrow \vec{\Psi}\left(\vec{M}_1\right)$$

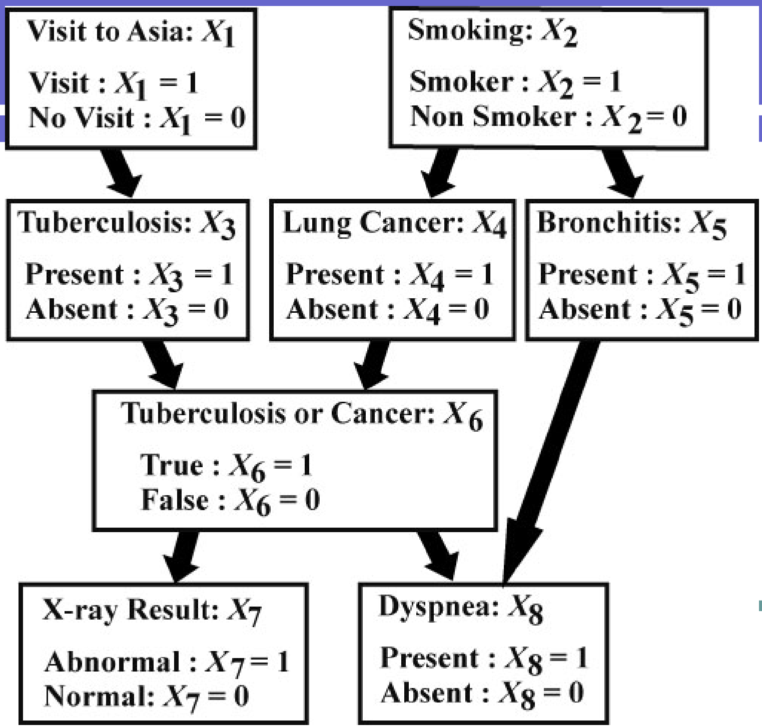
$$\vec{M}_3 \leftarrow \vec{\Psi}\left(\vec{M}_2\right)$$

$$\vdots$$



確率伝搬法

1. 閉路を持たないグラフ上の確率モデルに対して厳密な結果を与える.
2. 閉路を持つグラフ上の確率モデルでは近似アルゴリズムとなる.



$$
\begin{aligned}
\Pr\{\boldsymbol{X}\} = {} & \Pr\{X_8 | X_5, X_6\}\Pr\{X_7 | X_6\} \\
& \times \Pr\{X_6 | X_3, X_4\}\Pr\{X_5 | X_2\} \\
& \times \Pr\{X_4 | X_2\}\Pr\{X_3 | X_1\} \\
& \times \Pr\{X_1\}\Pr\{X_2\}
\end{aligned}
$$

８個程度のノードの簡単ではあるが閉路を含むグラフ上の確率モデルで確率伝搬法の構造を説明し，得られる近似結果と厳密な結果を比較してみる.

数値実験

$$P_8(1)=0.5607 \qquad P_8(0)=0.4393$$

$$P_{58}(1,1)=0.4736 \qquad P_{58}(1,0)=0.0764$$

$$P_{58}(0,1)=0.0871 \qquad P_{58}(0,0)=0.3629$$

確率伝搬法
Belief Propagation

$$P_8(1)=0.5640 \qquad P_8(0)=0.4360$$

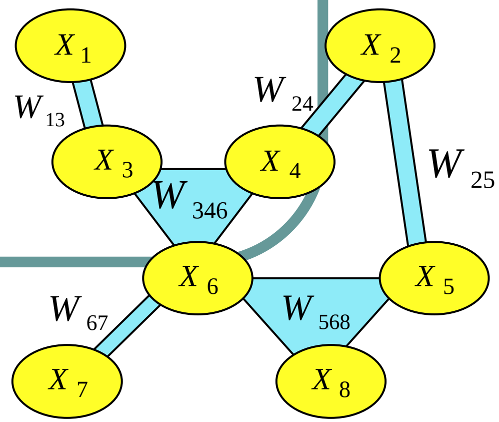
$$P_{58}(1,1)=0.4776 \qquad P_{58}(1,0)=0.0724$$

$$P_{58}(0,1)=0.0864 \qquad P_{58}(0,0)=0.3636$$

Exact

$$P_8(x_8) \equiv \sum_{\boldsymbol{x}\setminus\{x_8\}} P(x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8)$$

$$P_{58}(x_5,x_8) \equiv \sum_{\boldsymbol{x}\setminus\{x_5,x_8\}} P(x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8)$$

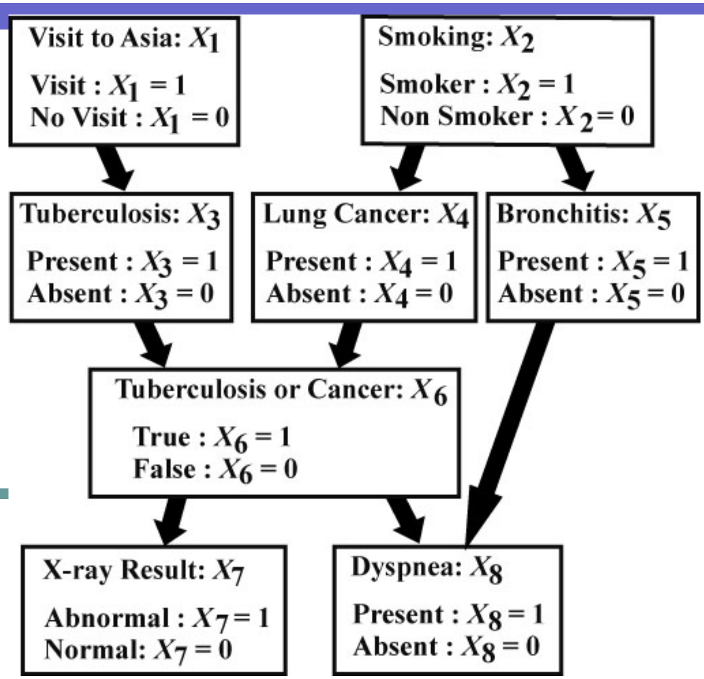


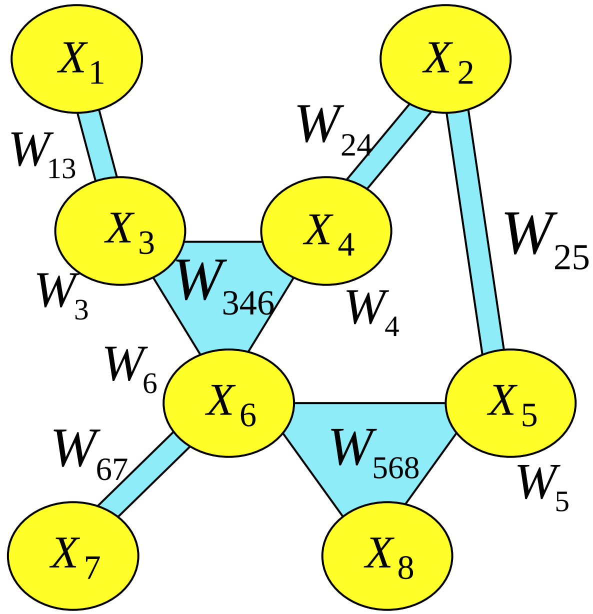
数値実験

確率伝搬法

$$\Pr\left\{X_{\text{Bronchitis}} = \text{Present} \middle| X_{\text{Dyspnea}} = \text{Present}\right\}$$

$$= \frac{\Pr\left\{X_{\text{Bronchitis}} = \text{Present}, X_{\text{Dyspnea}} = \text{Present}\right\}}{\Pr\left\{X_{\text{Dyspnea}} = \text{Present}\right\}} = \frac{0.3629}{0.4393} = 0.8261$$





確率伝搬法の定式化

- **確率伝搬法と平均場理論の類似性の指摘**

  Y. Kabashima and D. Saad, Belief propagation vs. TAP for decoding corrupted messages, Europhys. Lett. 44 (1998).
  M. Opper and D. Saad (eds), Advanced Mean Field Methods ---Theory and Practice (MIT Press, 2001).

- **一般化された確率伝搬法の提案**

  S. Yedidia, W. T. Freeman and Y. Weiss: Constructing free-energy approximations and generalized belief propagation algorithms, IEEE Transactions on Information Theory, 51 (2005).

- **確率伝搬法の情報幾何的解釈**

  S. Ikeda, T. Tanaka and S. Amari: Stochastic reasoning, free energy, and information geometry, Neural Computation, 16 (2004).

確率伝搬法の応用範囲

- **Image Processing**

  K. Tanaka: Statistical-mechanical approach to image processing (Topical Review), J. Phys. A, 35 (2002).

  A. S. Willsky: Multiresolution Markov Models for Signal and Image Processing, Proceedings of IEEE, 90 (2002).

- **Low Density Parity Check Codes**

  Y. Kabashima and D. Saad: Statistical mechanics of low-density parity-check codes (Topical Review), J. Phys. A, 37 (2004).

  S. Ikeda, T. Tanaka and S. Amari: Information geometry of turbo and low-density parity-check codes, IEEE Transactions on Information Theory, 50 (2004).

- **CDMA Multiuser Detection Algorithm**

  Y. Kabashima: A CDMA multiuser detection algorithm on the basis of belief propagation, J. Phys. A, 36 (2003).

  T. Tanaka and M. Okada: Approximate Belief propagation, density evolution, and statistical neurodynamics for CDMA multiuser detection, IEEE Transactions on Information Theory, 51 (2005).

- **Satisfability Problem**

  O. C. Martin, R. Monasson, R. Zecchina: Statistical mechanics methods and phase transitions in optimization problems, Theoretical Computer Science, 265 (2001).

  M. Mezard, G. Parisi, R. Zecchina: Analytic and algorithmic solution of random satisfability problems, Science, 297 (2002).

まとめ

- ベイズの公式による統計的モデリング
- ベイジアンネットワークのグラフ表現
- 確率伝搬法（Belief Propagation）