

# ベイジアンネットワーク 概説

---

## 5.2 グラフィカルモデリングとしての応用

茨城大学工学部

佐々木稔

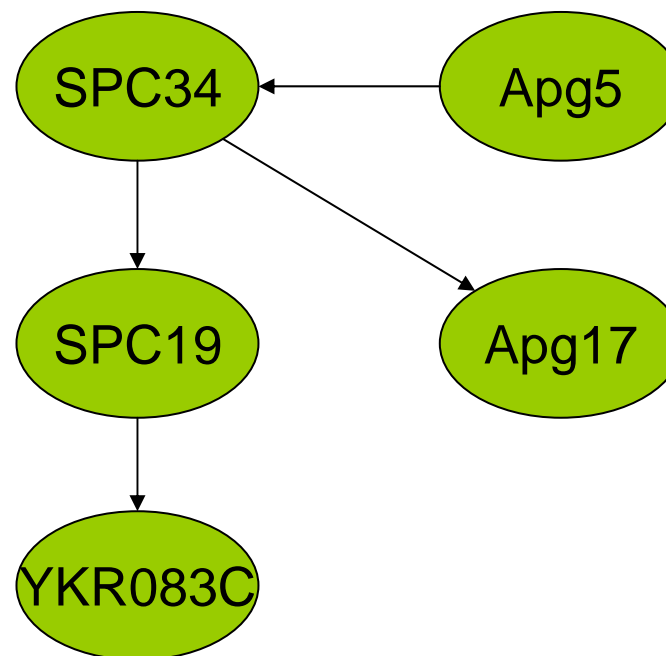
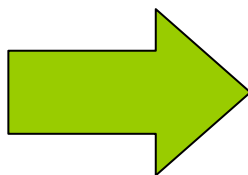
# 定性的なモデリング

- 多量のデータから得られたベイジアンネットワーク
  - 問題構造をグラフ構造で出力したい
- 例: 遺伝子情報処理

CCCTGTGGAG  
CCACACCCTA  
GGGTTGTGGA  
GCCACCCAAA  
ACCCCTGCTG  
TGGAGCCACA

ゲノム情報

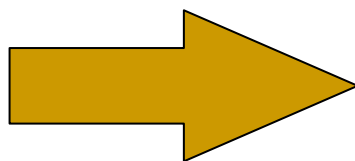
グラフィカル  
モデリング



# ベイジアンネットワークの統計的学習

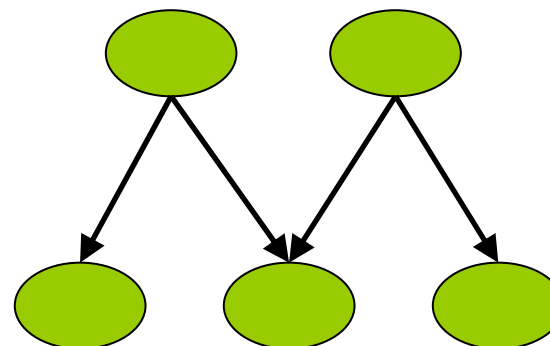
大量の  
統計  
データ

最もよく説明できる  
ネットワークを学習



AIC、BIC、MDLなどの  
情報量基準で探索

ベイジアンネットワーク



# モデルの構築

---

## □ 発生データ

### ■ 真のモデルが明らかでない場合

- 正しいモデルになる漸近一致性を持つ情報量基準を選ぶ

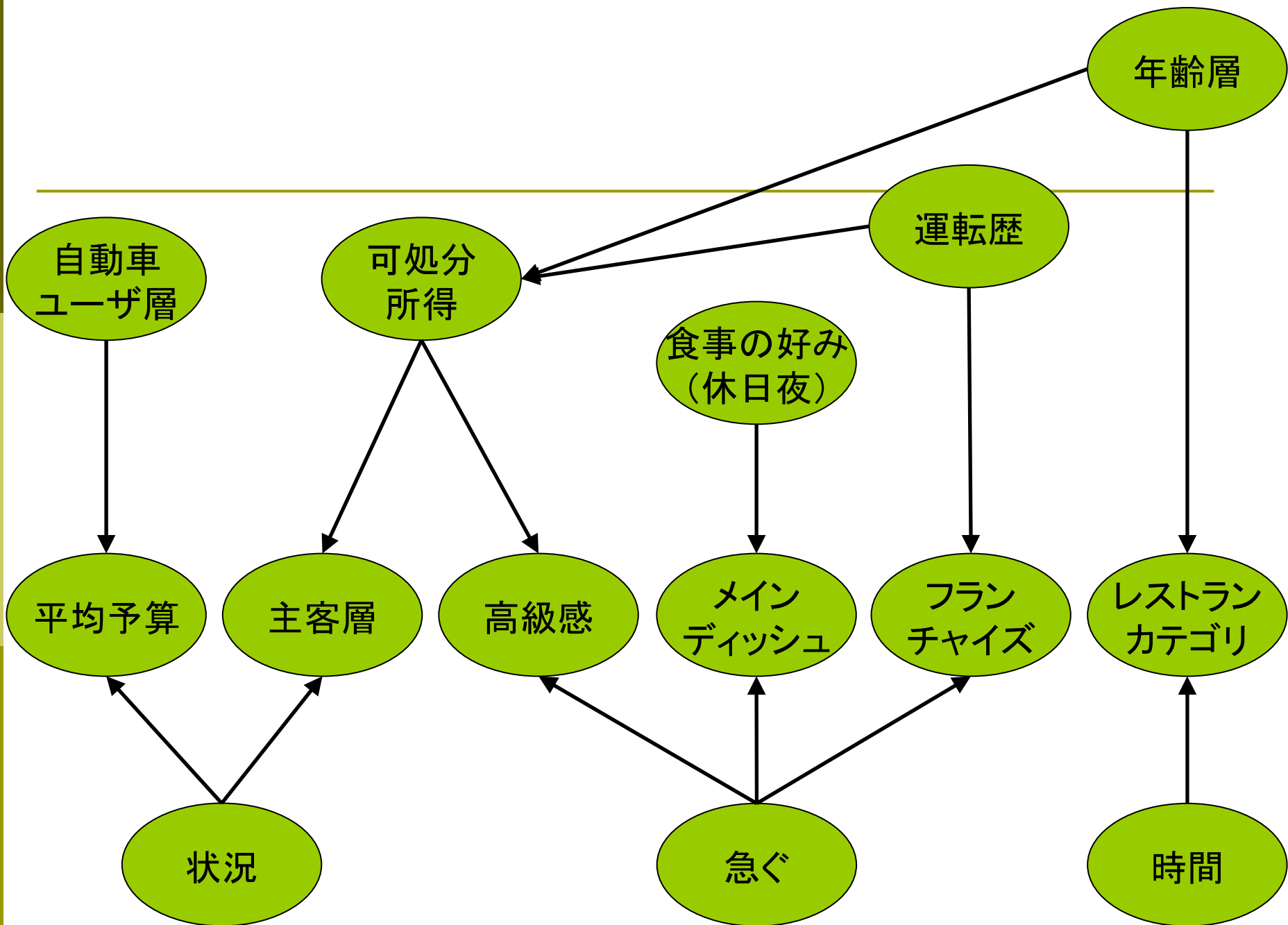
### ■ 真もモデルが明らかでない場合

- グラフ構造と変数集合がデータの背後にある有用な知識を提供
- 望ましい結果が得られない可能性がある
- ノード数が増えると、探索と情報量基準計算も増加

# モデルの構築の例

---

- アンケート
  - 品川周辺の182か所のレストランが対象
  - 6つの異なる状況で候補となるレストランを選択
  - 被験者 300人
- アンケート結果からベイジアンネットワークを構築



# 定量的なモデリング

---

## □ 確率推論

- 任意の変数に関する確率分布を計算
- 様々な条件下での知識の定量的な評価が可能
- 構築したモデルの有効活用が可能

## □ モデリング

- 目的変数と説明変数
- 膨大なデータの分析
- 変数の解釈

# 目的変数と説明変数

---

- ベイジアンネットワーク
  - 条件付き確率表
- 統計、多変量解析
  - 回帰モデル
  - 因子分析
  - 共分散構造解析
- 人工知能分野
  - 決定木
  - ニューラルネットワーク

# ベイジアンネットワークモデル

---

- 離散的な条件付き確率表で表現
  - 線形、非線形の依存関係も近似可能
  - 説明変数と目的変数の区別もない
  - 実際のクラス分類などを潜在クラス変数との解釈も可能

# 膨大なデータの分析

---

- ITの普及による情報爆発
  - 会員情報
  - 購買履歴
  - アクセス履歴
- 膨大なデータからの知識獲得
  - 優良顧客の識別
  - 顧客の購入傾向
- 従来のデータマイニングとの比較
  - 知識(モデル)の再利用性

# 変数の解釈

- 必要に応じて使う側が意味を設定
  - ある変数を潜在クラスとみなす
    - その変数が影響する部分モデル(混合モデル)が得られる
  - 潜在クラスの事後確率を計算
    - 事後確率最大のカテゴリを同定(潜在クラスの推定)
    - そのカテゴリからの部分モデルが得られる(階層ベイズ法)

