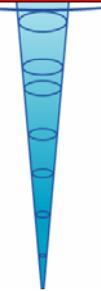


## 知識獲得における 精度と複雑性のトレードオフ

大阪府立大学大学院 工学研究科  
電気・情報系専攻 知能情報工学分野 教授  
大阪府立大学 計算知能研究所 所長

石淵 久生



### 略歴

- 1963 熊本県生まれ
- 1981 京都大学 工学部機械系 入学
- 1985 京都大学 工学部精密工学科 卒
- 1987 京都大学大学院 工学研究科 精密工学専攻 修了
- 1987 大阪府立大学 工学部 経営工学科 助手 (ファジィシステム)
- 1993 同 講師
- 1994 同 助教授
- 1999 同 教授
- 2000 大阪府立大学大学院 工学研究科 経営工学分野 教授
- 2005 大阪府立大学大学院 工学研究科 知能情報工学分野 教授
- 2006 大阪府立大学 計算知能研究所 所長(兼任)

## 学会活動

### 日本知能情報ファジィ学会

2003年度 & 2004年度 理事  
2005年度 & 2006年度 評議員  
2007年度 & 2008年度 副会長

### IEEE (CI Society and SMC Society)

#### Associate Editor

Trans. on SMC Part B (2002-)  
Trans. on Fuzzy Systems (2004-)  
Computational Intelligence Magazine (2005-)  
Trans. on Evolutionary Computation (2007-)

#### Technical Committee

Fuzzy Systems Technical Committee Vice-Chair (2007-)

## 発表内容

### 進化型多目的最適化

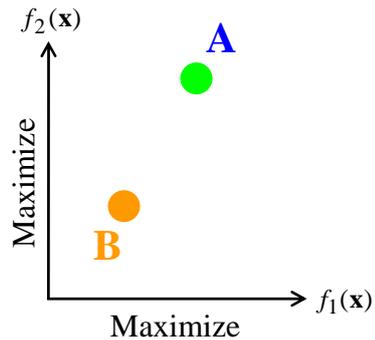
- ・ 遺伝的アルゴリズムなどの進化計算を用いた多目的最適化
- ・ 進化計算の分野で最もホットな研究分野
- ・ 特徴は『多数のパレート最適解が同時に獲得される』こと

### 多目的知識獲得

- ・ 複数の評価基準に基づく知識獲得
- ・ 今後の進展が期待される萌芽的な研究分野
- ・ 特徴は『複数の知識が同時に獲得される』こと

## 多目的最適化における解の比較

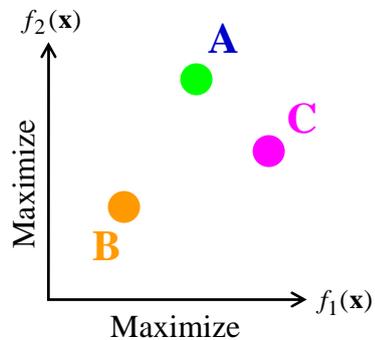
Maximize  $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}))$



A は B を優越する  
B は A に優越される  
(A は B よりも良い解)

## 多目的最適化における解の比較

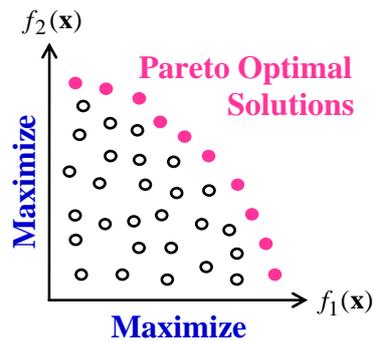
Maximize  $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}))$



A と C は互いに非劣  
(どちらが良いか言えない)

## 多目的最適化における解の概念

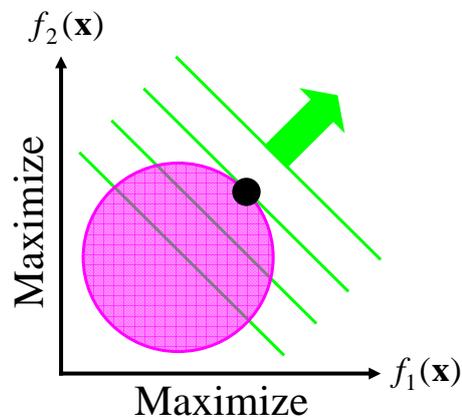
**パレート最適解**：他のどの実行可能解にも  
優越されない解（非劣解）



## 多目的最適化へのアプローチ

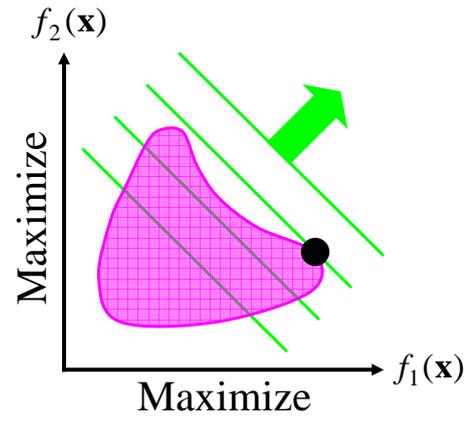
**複数の目的を統合（加重和最大化）**

$$\text{Maximize } g(\mathbf{x}) = w_1 f_1(\mathbf{x}) + w_2 f_2(\mathbf{x})$$



### 加重和最大化の問題点

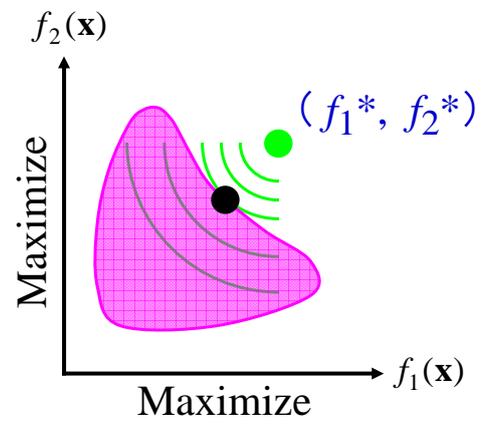
重みベクトルの設定に敏感  
非凸領域の解が獲得できない



### 多目的最適化へのアプローチ

参照解(目標とする解)との距離の最小化

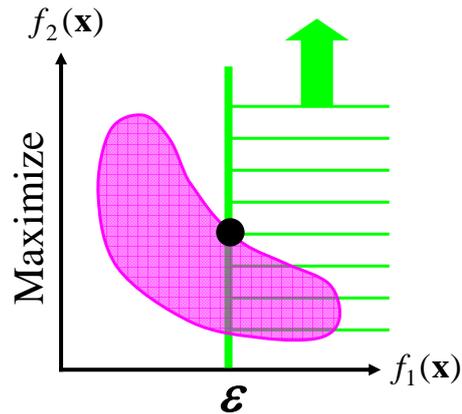
$$\text{Minimize } g(\mathbf{x}) = \| (f_1(\mathbf{x}), f_2(\mathbf{x})) - (f_1^*, f_2^*) \|$$



## 多目的最適化へのアプローチ

### 目的関数を制約条件に変更

Maximize  $f_2(\mathbf{x})$  subject to  $f_1(\mathbf{x}) \geq \varepsilon$



## 共通の問題点

重みベクトルや参照点, 制約の事前設定が困難

### 多目的最適化の手順

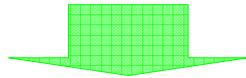
- ステップ1. 多目的最適化問題の定式化
- ステップ2. 意思決定者から何らかの情報を得る
- ステップ3. 単一目的最適化問題として最適解を求める

**問題点:** ステップ2で重みベクトルや参照点, 制約などの情報を意思決定者から獲得することが困難

## 進化型多目的最適化

### 多目的最適化の手順

- ステップ1. 多目的最適化問題の定式化
- ステップ2. **意思決定者から何らかの情報を得る**
- ステップ3. 単一目的最適化問題として最適解を求める

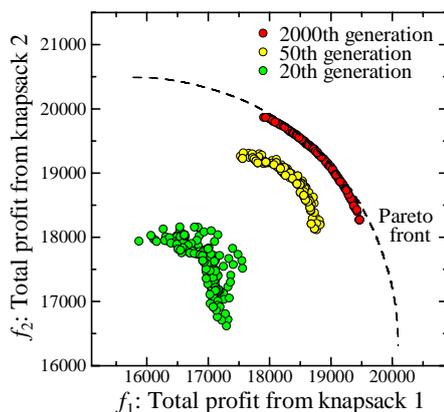


### 進化型多目的最適化の手順

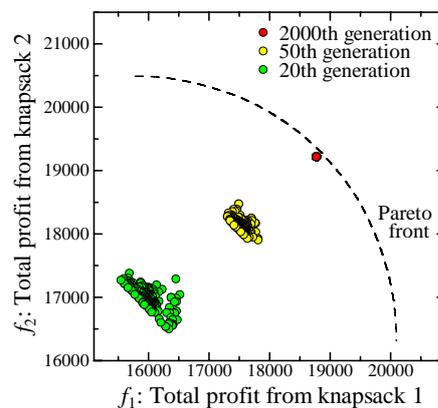
- ステップ1. 多目的最適化問題の定式化
- ステップ2. **多数のパレート最適解を求める**
- ステップ3. パレート最適解から1個の選好解を選ぶ

## 進化型多目的最適化と加重和最大化

### 2目的最大化問題



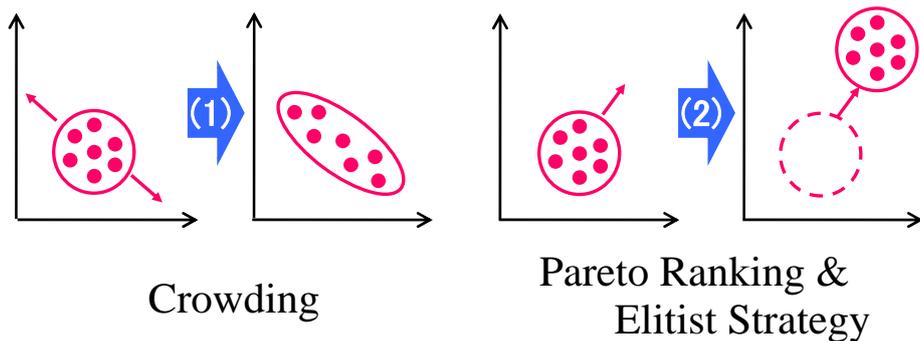
進化型多目的最適化



加重和最大化

## 進化型多目的最適化のポイント

- (1) 解集合の多様性を増加させる
- (2) パレートフロントへの収束性を改良する

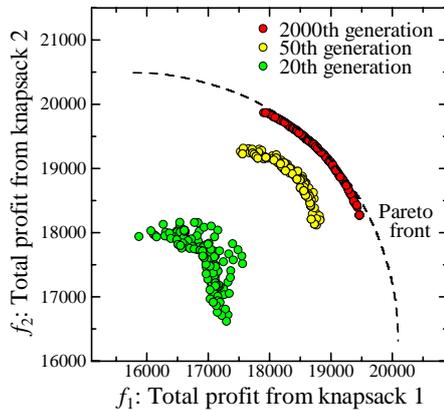


## 進化型多目的最適化の研究方向

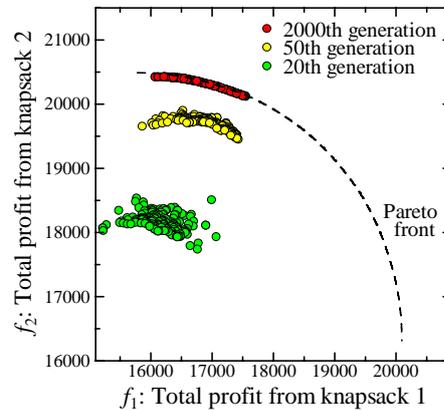
- (1) 意思決定者の持つ選好情報の活用
- (2) 多数の目的を持つ多目的最適化 (Many-Objective)
- (3) 解集合に対する評価指標を用いたアルゴリズム
- (4) 獲得された多数の解の意思決定者への提示方法

## 進化型多目的最適化の研究方向

### (1) 意思決定者の持つ選好情報の活用



進化型多目的最適化



選好情報の利用の例

## 進化型多目的最適化の研究方向

### (2) 多数の目的を持つ多目的最適化(5目的や10目的)

- ・ 目的の数が多いと解と解の間で優越関係が成立しない
- ・ 全てのパレート最適解を求めることは非現実的

### (3) 解集合に対する評価指標を用いたアルゴリズム

- ・ 多目的最適化アルゴリズムの比較に関する議論
- ・ 獲得された非劣解集合の評価指標の提案
- ・ 評価指標を直接的に用いたアルゴリズムの提案

## 進化型多目的最適化の研究方向

### (4) 獲得された多数の解の提示方法

- ・ 2目的最適化問題であれば簡単に図示できる
- ・ 3目的以上の場合には困難（10目的最適化問題の解）

### 進化型多目的最適化の手順

- ステップ1. 多目的最適化問題の定式化
- ステップ2. 多数のパレート最適解を求める
- ステップ3. **パレート最適解から1個の選好解を選ぶ**

## 発表内容

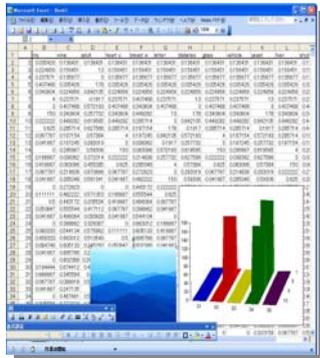
### 進化型多目的最適化

- ・ 遺伝的アルゴリズムなどの進化計算を用いた多目的最適化
- ・ 進化計算の分野で最もホットな研究分野
- ・ 特徴は『多数のパレート最適解が同時に獲得される』こと

### 多目的知識獲得

- ・ 複数の評価基準に基づく知識獲得
- ・ 今後の進展が期待される萌芽的な研究分野
- ・ 特徴は『複数の知識が同時に獲得される』こと

## 知識獲得



知識獲得

- If  $x$  is ... then Class 1
- If  $x$  is ... then Class 2
- If  $x$  is ... then Class 3
- If  $x$  is ... then Class 2
- ⋮
- If  $x$  is ... then Class 1
- If  $x$  is ... then Class 2

### 我々の研究の出発点

大規模データから獲得された大量のルール集合は、  
利用者が簡単に理解できる使いやすい知識なのか？

## 容易に理解できる知識の獲得

### 進化型ルール選択

遺伝的アルゴリズムを用いて、獲得された大量のルール  
から**必要最小限のルール**だけを選び出す。

- If  $x$  is ... then Class 1
- If  $x$  is ... then Class 2
- If  $x$  is ... then Class 3
- If  $x$  is ... then Class 2
- ⋮
- If  $x$  is ... then Class 1
- If  $x$  is ... then Class 2

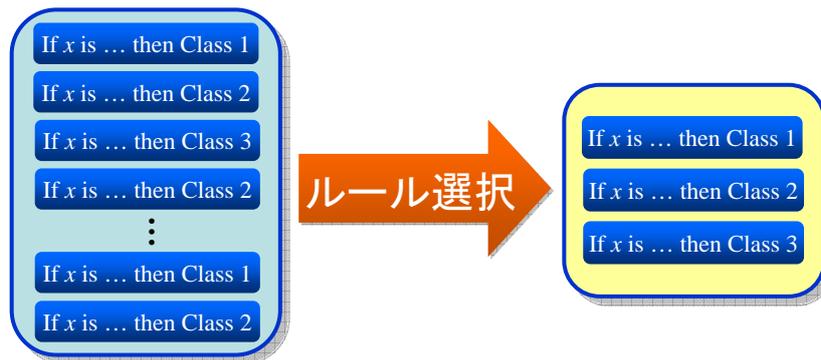
ルール選択

- If  $x$  is ... then Class 1
- If  $x$  is ... then Class 2
- If  $x$  is ... then Class 3

## 必要最小限のルールとは？

### 問題点

全く同じデータベースから知識獲得を行う場合でも、状況により、必要とする知識の精度や量は異なるのでは？

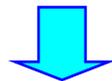


## 本研究のアプローチ

個々の利用者が必要とする知識は異なる



個々の利用者の要求に応じた知識を獲得する  
(Personalized Knowledge Extraction)



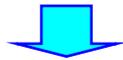
どのような要求を持った利用者にも対応できるように、精度と複雑性の異なる様々な知識を獲得する

## 精度の最大化と複雑性の最小化

### 知識獲得の目的

1. 精度の最大化: 精度の高い知識を獲得したい  
⇒ 問題点: 獲得された知識は複雑で分かりにくい
2. 複雑性の最小化: 分かりやすい知識を獲得したい  
⇒ 問題点: 獲得された知識は単純で精度が低い

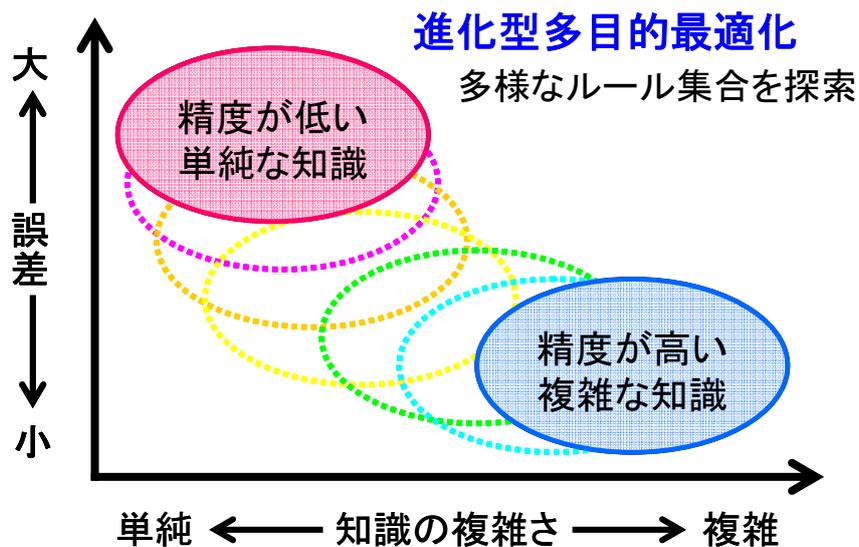
両者を同時に最適化することはできない



### 本研究でのアプローチ

精度と複雑性に関する多目的知識獲得

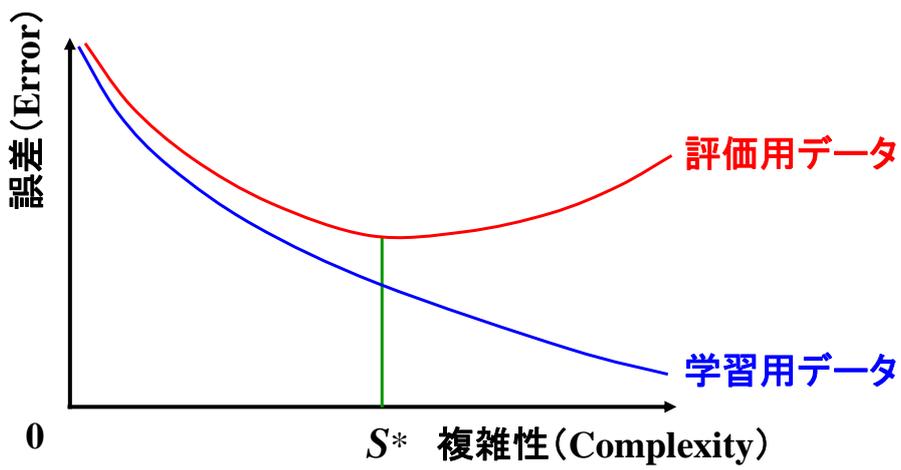
## 本研究のアプローチ



### 最適な構造の決定

評価用データに対する誤差最小化

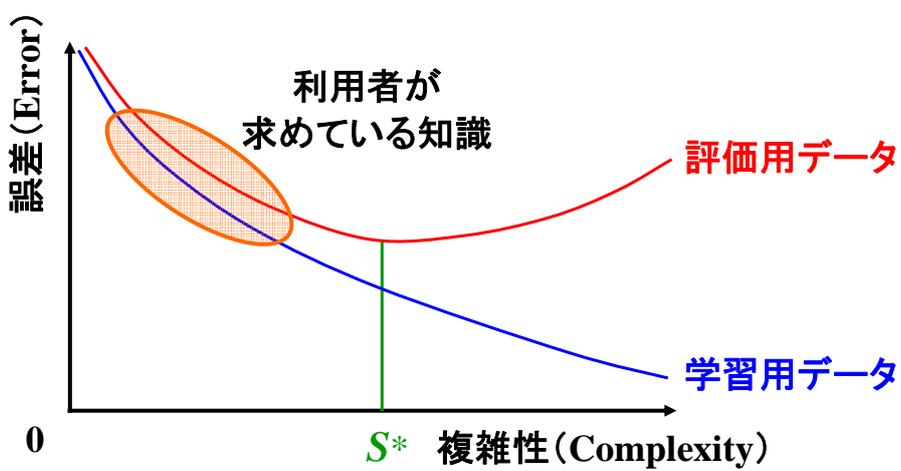
→ 汎化能力が最大化となる  $S^*$  を求める



### 汎化能力最大化の問題点

分かりやすさを考慮しても  $S^*$  が最適なのか？

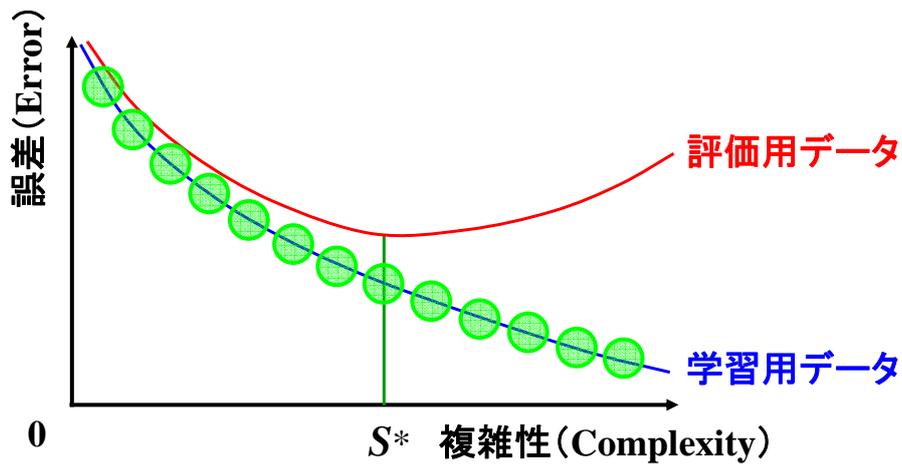
→  $S^*$  が複雑すぎる状況も考えられる



## 本研究のアプローチ

### 精度と複雑性間のトレードオフ解析

→ 多数のパレート最適解を求める

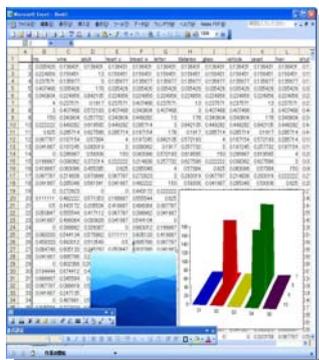


## 多様なルール集合の探索を行うための手法

### 2段階ルール選択

ステップ1: ルール獲得(データマイニング)

ステップ2: 多目的ルール選択(遺伝的アルゴリズム)



ルール獲得

If x is ... then Class 1

If x is ... then Class 2

If x is ... then Class 3

If x is ... then Class 2

⋮

If x is ... then Class 1

If x is ... then Class 2

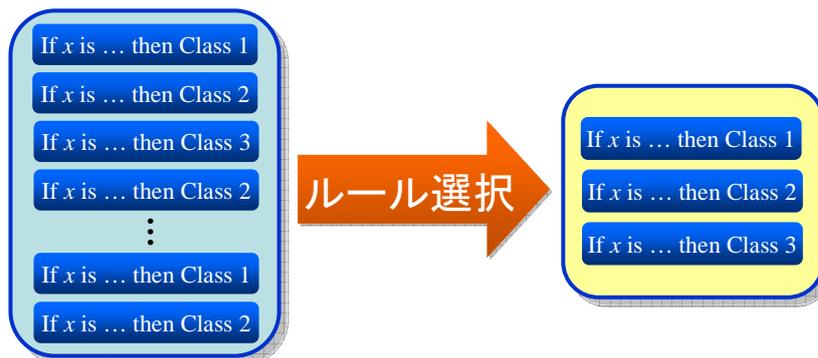
## 多様なルール集合の探索を行うための手法

31

### 2段階ルール選択

ステップ1: ルール獲得(データマイニング)

ステップ2: 多目的ルール選択(遺伝的アルゴリズム)



## 目的関数

32

### 3目的組合せ最適化問題

Maximize  $f_1(S)$ , minimize  $f_2(S)$  and  $f_3(S)$

$f_1(S)$ : ルール集合Sによる正当パターン数

$f_2(S)$ : ルール集合Sに含まれるルール数

$f_3(S)$ : 条件の数の総数(総ルール長)

## 進化型多目的最適化を用いた数値実験

個体群サイズ: 200個 (NSGA-II)

終了基準: 1000世代

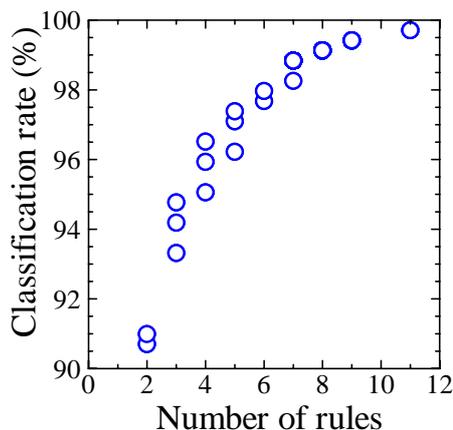
交叉確率: 0.9 (一様交叉)

突然変異確率:  $1/N$  ( $N$ : 候補ルール数)

評価方法: 10x2CV (Two-fold cross-validation)

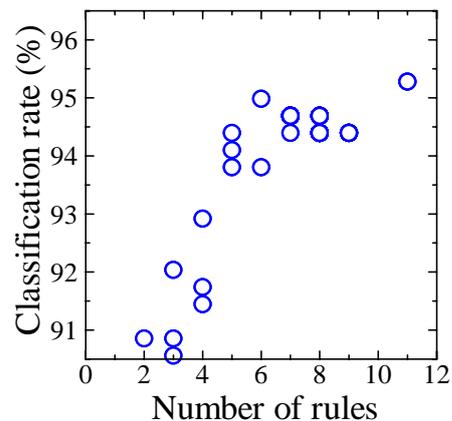
## 精度と複雑性との関係

### 3目的遺伝的ルール選択の結果 (Breast W)



学習用データに対する正答率

明確なトレードオフ

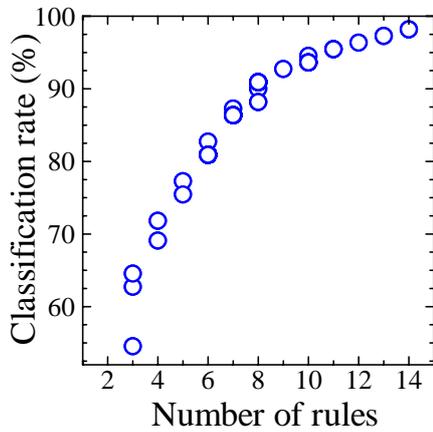


評価用データに対する正答率

ルール数6個以上では微妙

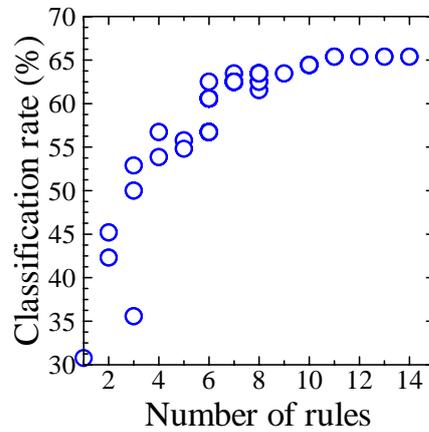
## 精度と複雑性との関係

### 3目的遺伝的ルール選択の結果 (Glass)



学習用データに対する正答率

明確なトレードオフ

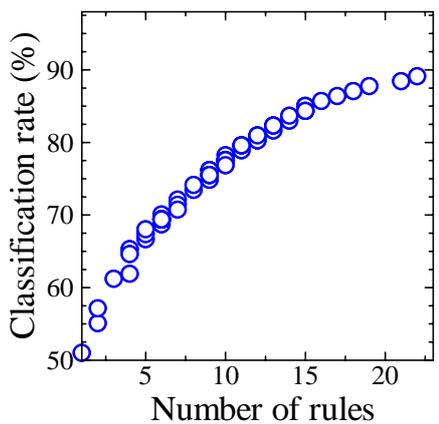


評価用データに対する正答率

ルール数10個以上では微妙

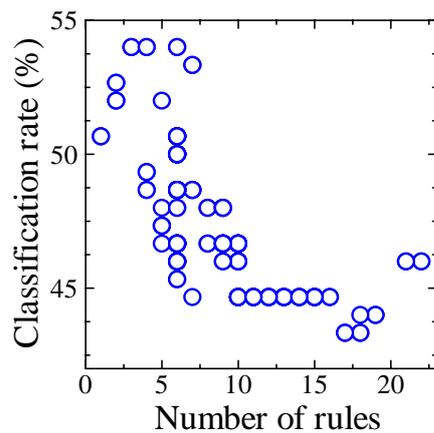
## 精度と複雑性との関係

### 3目的遺伝的ルール選択の結果 (Heart C Data Set)



学習用データに対する正答率

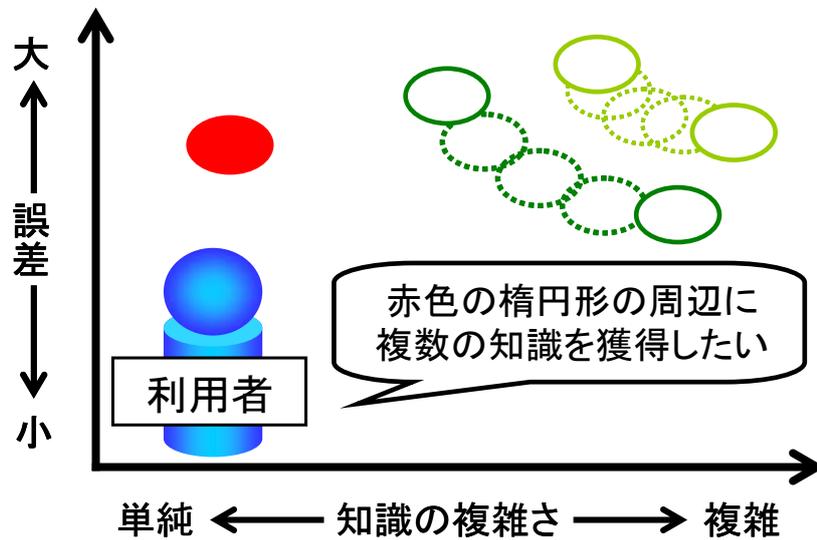
明確なトレードオフ



評価用データに対する正答率

ルール数5個以上で過剰適合

## 拡張(利用者の選好情報の導入)



## 利用者の選好情報の導入

