



HOKKAIDO
UNIVERSITY

講義「人工知能」

第12回 進化型計算メカニズム2

意思決定システムの進化的設計

北海道大学大学院情報科学研究院
情報理工学部門 複合情報工学分野
調和系工学研究室 准教授 山下倫央

<http://harmo-lab.jp>

tomohisa@ist.hokudai.ac.jp

2024年5月16日(木)

遺伝的アルゴリズムの理論的考察

Binary Cordingとスキーマ定理に基づく
単純遺伝的アルゴリズムの個体群の集団的挙動解析

Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning, Addison Wesley, 1989.



遺伝子の部分表現であるスキーマの持つ特徴と
個体群の世代間変化の関係を明らかにした

スキーマとは

スキーマ

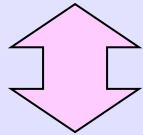
遺伝子の集団に保持される部分構造.

{1, 0, * (don't care)} よりなる文字列で構成される.

* は1か0か問わない部分を表す.

スキーマH

1	1	*	0	*
---	---	---	---	---



1	1	0	0	0
1	1	0	0	1
1	1	1	0	0
1	1	1	0	1

これらの遺伝子表現が
スキーマHで代表される.

ハミング距離

定義1

2つの個体を $a = \langle a_1 a_2 \dots a_n \rangle$, $b = \langle b_1 b_2 \dots b_n \rangle$
($a_i, b_i \in \{0, 1\}, i = 1, \dots, n$) とするとき, a と b の
ハミング距離 $d(a, b)$ は次式で与えられる.

$$d(a, b) = \sum_{i=1}^n |a_i - b_i|$$

遺伝子a

1	0	1	1	0	1	1	0	0	1
---	---	---	---	---	---	---	---	---	---

遺伝子b

1	1	1	0	0	1	0	0	1	1
---	---	---	---	---	---	---	---	---	---

$$d(a, b) = 4$$

スキーマの次数

定義2

スキーマHにおいて、具体的に値の定められている
(つまり*では無い)遺伝子座の数 $o(H)$ をスキーマHの
次数という.

スキーマH

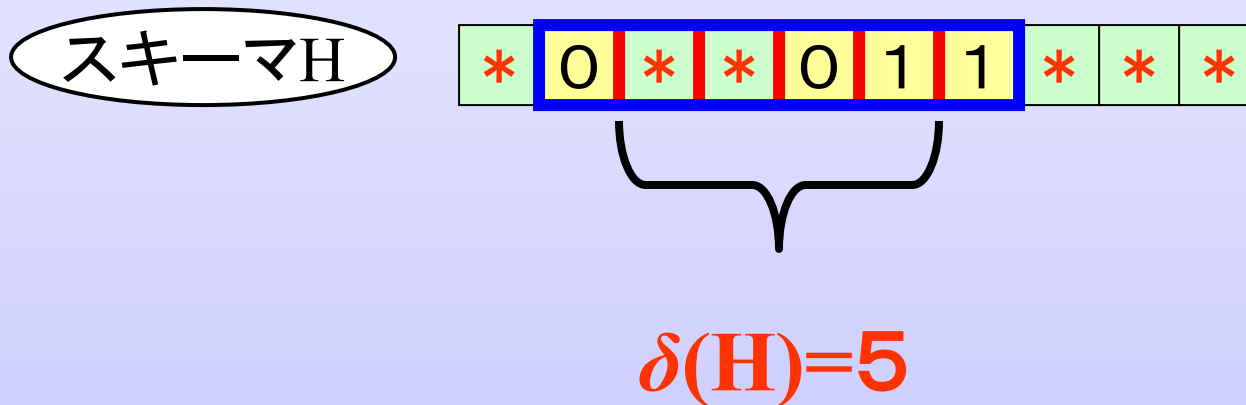
*	0	*	*	0	1	1	*	0	*
---	---	---	---	---	---	---	---	---	---

$$o(H)=5$$

スキーマの定義長

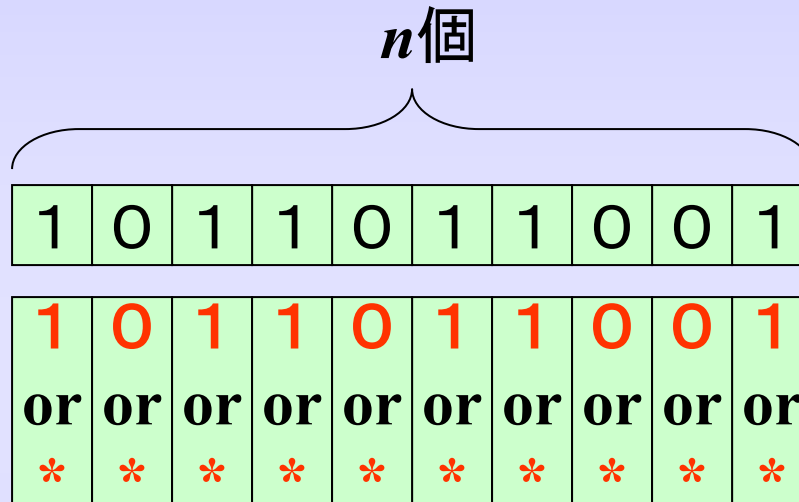
定義3

スキーマHにおいて、最初に値の定められている遺伝子座から最後に値の定められている遺伝子座までの長さ $\delta(H)$ を定義長という。ここで長さとは、その間にある遺伝子座と遺伝子座との間の個数を言う。



スキーマの性質

性質1 1つの特定の個体の属するスキーマは 2^n 個ある.



2^n 通り

例) $n=3$, 001 が属する
 $2^3=8$ 個のスキーマ

001	*01
00*	*0*
0*1	**1
0**	***

性質2 従って, サイズ N の集団内のスキーマは高々 $N \cdot 2^n$ 通り

性質3 長さ n の遺伝子表現において, 存在する全スキーマ数は 3^n 通り

スキーマ数(1)

定理1

各遺伝子座での0,1は等確率0.5で生起するものとする.
このとき, N 個の個体よりなる長さ n の個体群の中の
異なるスキーマ数の期待値は

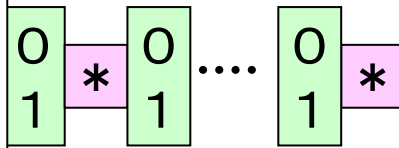
$$\sum_{i=0}^n C_i 2^i \left(1 - \left(1 - (1/2)^i \right)^N \right) \text{である.}$$

スキーマ数(2)

1次 長さ 3

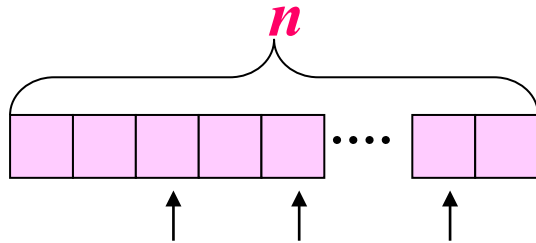
**0
**1
0
1
0**
1**

$$i \left(1 - \left(1 - (1/2)^i \right)^N \right)$$



0 か 1 の
場合の数

スキーマは 2^i 通り



0 か 1 の入る場所
の場合の数

n 個から i 個 選ぶ選び方は ${}_n C_i$ 通り

長さ n の個体に属する
次数 i のスキーマは全部で
 ${}_n C_i \cdot 2^i$

暗黙の並列性

ランダムに発生させた
ある個体が次数 i の
スキーマに属する確率は
 $(1/2)^i$

属さない確率は
 $1 - (1/2)^i$

N 個の個体がどれも
属さない確率は
 $(1 - (1/2)^i)^N$

どれかの個体が次数 i の
スキーマに属す確率は
 $1 - (1 - (1/2)^i)^N$

スキーマ評価値

世代 k におけるスキーマ H の評価値 $F(H, k)$ は

$$F(H, k) = \frac{1}{m(H, k)} \sum_{x \in H} f(x)$$

ただし $m(H, k)$ は世代 k の個体群中でスキーマ H に包含される個体数. $f(x)$ は個体 x に対する適応度関数.

スキーマ H

*	*	1	0	*	1
---	---	---	---	---	---

スキーマ評価値

$$F(H, k) = (1.0 + 4.2) / 2$$

遺伝子プール

適応度

1	1	1	0	1	1
0	1	1	1	0	1
0	1	0	1	0	1
1	0	1	0	1	1

1. 0

2. 3

3. 3

4. 2

遺伝子操作によるスキーマの変化(1)

(1) 選択

世代 k において個体群中のスキーマ H に属する個体数を $m(H, k)$ と表す. スキーマ H の評価値を $F(H, k)$, 個体群中の全個体の適応度の平均値を $F_{avg}(k)$ とすれば, スキーマ H に属する個体のルーレット選択による個体数期待値は,

$$m_s(H, k+1) = \sum_{I \in H} \frac{f(I)}{F_{avg}(k)}$$

ルーレット選択において
スキーマ H に属する
個体を選択される確率

$$= m(H, k) \cdot \left(\frac{1}{m(H, k)} \sum_{I \in H} f(I) \right) / F_{avg}(k)$$

$$= m(H, k) \frac{F(H, k)}{F_{avg}(k)}$$

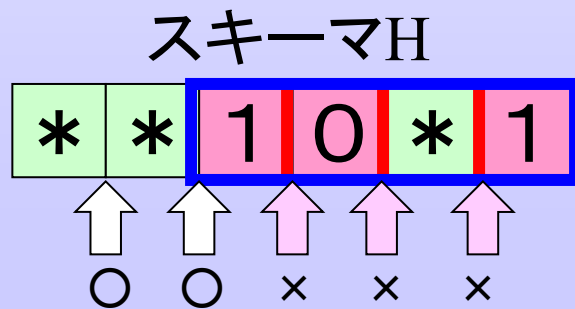
ルーレット選択において
スキーマ H に属する個体の数は
評価値に応じて増減

遺伝子操作によるスキーマの変化(2)

(2) 交叉

交叉確率を p_c とすれば、スキーマ H に属する個体が交叉で壊される確率は $p_c \delta(H)/(n-1)$ である。交叉により新たに同じスキーマに属する個体が生成されることや、またスキーマ H に属する個体同士が交叉する場合もあるので、交叉後に残る個体数の期待値は、選択による個体数の期待に $1 - \frac{p_c \delta(H)}{n-1}$ を掛けた以上の値となる。

$$m_c(H, k+1) \geq m(H, k) \cdot \left(1 - \frac{p_c \delta(H)}{n-1}\right)$$



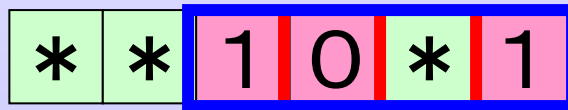
壊れる確率は $p_c \cdot 3/(6-1)$

遺伝子操作によるスキーマの変化(2)

(3) 突然変異

突然変異によりビットの値が変化する確率を p_m とすると、突然変異操作後に、その個体が同じスキーマにとどまる確率は $(1 - p_m)^{o(H)}$ となる．一般に p_m の値は小さいので、近似的に $1 - p_m \cdot o(H)$ とすることができる．これも、交叉と同様、突然変異によりこのスキーマに属する個体が新たに生成される可能性があるので、実際には個体数はさらに多くなる．

$$m_m(H, k+1) \geq m(H, k) \cdot (1 - p_m \cdot o(H))$$



1bitが変化する確率 p_m → 変化しない確率 $(1 - p_m)$

スキーマ定理

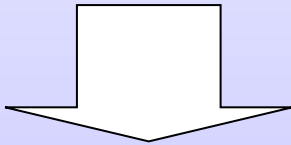
スキーマ定理

$$m_s(H, k+1) = m(H, k) \frac{F(H, k)}{F_{avg}(k)}$$

$$m_c(H, k+1) \geq m(H, k) \cdot \left(1 - \frac{p_c \delta(H)}{n-1}\right)$$

$$m_m(H, k+1) \geq m(H, k) \cdot (1 - p_m \cdot o(H))$$

$$m(H, k+1) \geq m(H, k) \frac{F(H, k)}{F_{avg}(k)} \cdot \left(1 - \frac{p_c \delta(H)}{n-1}\right) \cdot (1 - p_m \cdot o(H))$$



$$m(H, k+1) \geq m(H, k) \frac{F(H, k)}{f_{avg}(k)} \left(1 - p_c \frac{\delta(H)}{n-1} - p_m o(H)\right)$$

n : 遺伝子長

p_c : 交叉率

p_m : 突然変異率

$m(H, k)$: 世代 k におけるスキーマ H
に包含される遺伝子数

$F(H, k)$: 世代 k におけるスキーマ H
の評価値

$F_{avg}(k)$: 世代 k の適応度の平均値

$\delta(H)$: スキーマ H の定義長

$o(H)$: スキーマ H のオーダー

スキーマの
評価値

スキーマの
定義長

スキーマの
次数

Building Block仮説

スキーマ定理より

短くて低いオーダであり, しかも適応度が平均以上のスキーマは飛躍的に増大していく.

静的なBuilding Block(積み木)仮説

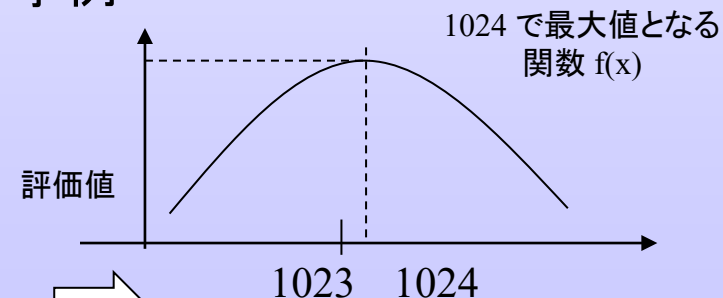
与えられた任意の低次, 短定義長の超平面(即ちスキーマ)パーティションにおいて, GAは(そのパーティション内の)最良な静的平均適応度(勝者)をもつ超平面に収束すると期待される.

ただし、表現型が近い個体は遺伝子型も類似していることが前提

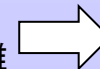
例: Building Block仮説が適用されにくい事例

表現型1023: 遺伝子型 0111111111

表現型1024: 遺伝子型 1000000000



左から探索した場合、
最適解にたどり着くのが困難



一般スキーマ定理

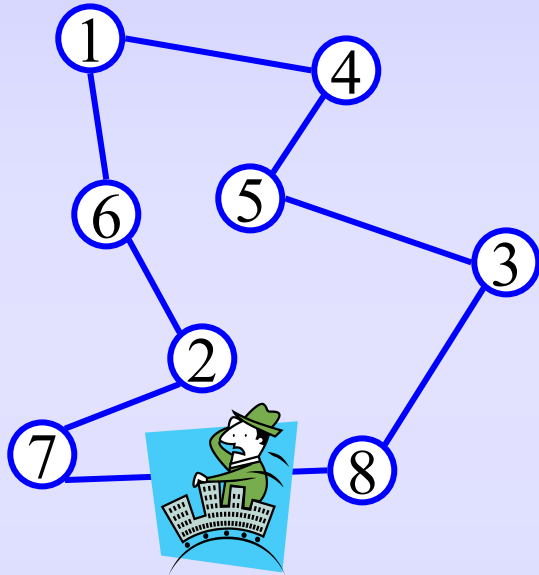
一点交叉のスキーマ定理を, m点交叉に拡張

一般スキーマ定理(m点交叉)

$$m(H, k+1) \geq m(H, k) \frac{F(H, k)}{F_{avg}(k)} \left(1 - p_c \frac{{}^nC_m - {}^{n-1-\delta(H)}C_m}{{}^{n-1}C_m} - p_m o(H) \right)$$

遺伝的アルゴリズムの応用例(1)

巡回セールスマン問題(TSP)



すべての都市を一巡する総移動距離
が最小の巡回路を求める問題

都市数が n のとき, 全組み合わせは
 $(n-1)!/2$

遺伝子表現:

都市の巡回順

例: 1 6 2 7 8 3 5 4

適応度:

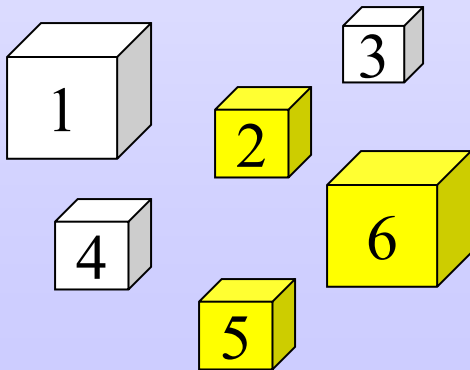
総巡回路長

遺伝的アルゴリズムの応用例(2)

ナップザック問題



複数の物体が与えられたとき、重さがある範囲内でいくつかの物体を選択し、その価値の合計が最大になるような組み合わせを求める問題



遺伝子表現:

物体*i*が選択されているとき1, されていないとき0であらわす.

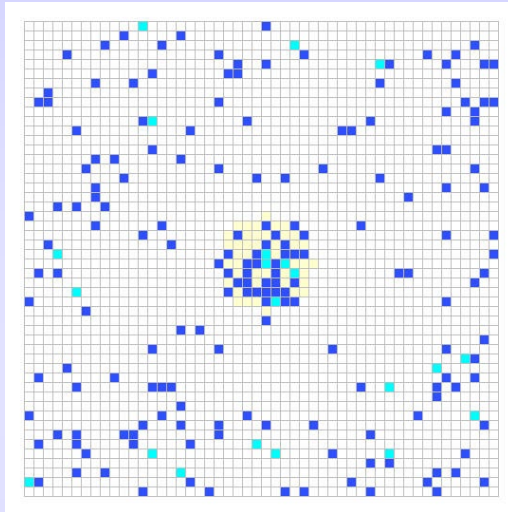
例: 010011

適応度:

価値の合計

遺伝的アルゴリズムの応用例(3)

人工社会のエージェント設計



上下左右に砂糖の多い場所があるか？
0000 ~ 1111



上下左右の移動
1 ~ 4

0000	1
0001	1
0010	4
0011	2
0100	3
...	

遺伝子表現

個々のエージェントは生き残るために砂糖を獲得しようと行動するが、砂糖を多く持っているエージェントが必要以上に砂糖を集めると、砂糖を手に入れられないエージェントが死滅してしまう。

適応度：
生き残りエージェント数

果たしてどういう行動ルールがよいのか？