

【調査報告】スキーマ定理
森 隆史

1 はじめに

GA には積木仮説というものがある。積木仮説とは、最適解に近い良好な解は、その部分構造(積木: building block) が組み合わさることで生成されるという仮説である。積木を考える 1 つの方法に、スキーマがあり、このスキーマの数が世代間でどう変化するかを見積もるための定理をスキーマ定理と呼ぶ。

2 スキーマ

スキーマとは、個体の構成要素である遺伝子の部分集合のことである。

2 つの記号 {0, 1} のどちらかを取る l 個の記号からなる記号列(染色体)の集合を S_l とする。すなわち、 S_l は 1 つの個体を表す。 S_l の要素いくつかからなる個体集合があったとき、この個体集合のスキーマとは、{0, 1, *} のいずれかを取る l 個の記号からなる記号列で定義され、その記号列に一致する個体の集合を表す。ただし、記号 "*" は、ワイルドカード記号(wild card symbol, don't care symbol と呼ばれる)であり、"0", "1" のいずれにも一致する。たとえば、記号列 10*1* で定義されるスキーマは、集合 {10010, 10011, 10110, 10111} を表す¹⁾。

また、スキーマには、定義長とオーダが定義されている²⁾。定義長は、左から見て最初の "*" 以外の記号と最後の "*" 以外の記号間の距離である。オーダ(order)は、 "*" 以外の記号の数を示す。

3 スキーマ定理²⁾

本節で用いる変数の一覧を Table 3 に示す。

Table 1 使用する変数

n	集団サイズ
f_1, f_2, \dots, f_n	各染色体の適合度
$f(i_j \dots k)$	記号列 $i_j \dots k$ の適合度
$f(H)$	スキーマ H の適合度
f	集団の平均適合度
$m(H, t)$	世代 t においてスキーマ H に含まれる記号列の数
l	染色体長
$\delta(H)$	スキーマ H の定義長
$O(H)$	スキーマ H のオーダ
p_c	交叉率
p_m	突然変異率

次のそれぞれの場合について、次世代のスキーマ H の数を考える。

交叉や突然変異が無い場合

交叉や突然変異が無い場合、個体 i は全個体の適合度の総計中で個体 i の適合度が占める割合の確率で次世代に残る。したがって、次世代でのスキーマ H の数は、式

(1) のようになると考えられる。

$$\begin{aligned}
 m(H, t+1) &= m(H, t) \cdot n \cdot \frac{f(H)}{\sum f_i} \\
 m(H, t+1) &= m(H, t) \times \frac{f(H)}{f}
 \end{aligned}
 \tag{1}$$

交叉を行う場合

定義長が長いスキーマは、その間に交叉点が入り破壊される可能性が大きい。このことを定式化すると式(2)のようになる。交叉が起こった上でスキーマ H が破壊されるのは、交叉点が $\delta(H)$ 内にあるときであることを示している。なお、不等号となるのは同じスキーマ H に含まれる記号列同士が交叉をしてもスキーマが破壊されることがないなどの場合があるためである。

$$\begin{aligned}
 m(H, t+1) \\
 \geq m(H, t) \cdot \frac{f(H)}{f} \cdot \{1 - p_c \cdot \frac{\delta(H)}{l-1}\}
 \end{aligned}
 \tag{2}$$

突然変異を行う場合

突然変異によりスキーマが破壊されるのは、スキーマ H の $O(H)$ 個の 0 か 1 が突然変異した場合である。したがって、次世代におけるスキーマ H の数は、式(3)で見積もれる。

$$\begin{aligned}
 m(H, t+1) \\
 \geq m(H, t) \cdot \frac{f(H)}{f} \cdot \{1 - p_c \cdot \frac{\delta(H)}{l-1} - O(H) \cdot p_m\}
 \end{aligned}
 \tag{3}$$

式(3)のことをスキーマ定理と呼ぶ。また、式(3)から、短くて低いオーダであり、かつ適合度が平均以上のスキーマは飛躍的に増大することが分かる。このようなスキーマのことを積木 (building block) と呼ぶ。

4 まとめ

本報告では、スキーマ定理について調査を行った。

スキーマ定理によると、短くて低いオーダであり、かつ適合度が平均以上のスキーマは飛躍的に増大ということがいえる。GA では、このようなスキーマを組み合わせることで良好な解を得るという積木仮説などもあり、スキーマ定理は GA において非常に重要な定理であるといえる。

参考文献

- 1) 三宮信夫, 喜多一, 玉置久, 岩本貴司. 遺伝的アルゴリズムと最適化. 朝倉書店, 1998.
- 2) 伊庭斉志. 遺伝的アルゴリズムの基礎. オーム社, 1994.