

# EQUAL SUM SUBSETS に対する遺伝的アルゴリズムを用いた 学習者適応型 e ラーニングシステムの開発

田淵宏太 陰山大貴 名古屋孝幸 豊田寿行  
鳥取環境大学 環境情報学部 情報システム学科

## 1. はじめに

従来の e ラーニングは、学習者の属性や知識レベルに関係なく、全ての学習者に対して同一の教材を配信するシステムが中心となっており、学習者ごとに最適な問題を提供することが困難である。そのため、学習者の理解度を逐次測定し、課題を動的に学習者に提示する学習者適応型 e ラーニングシステムが提案されている [2]。

本研究では、年代等の属性によりクラス分けされた各学習者カテゴリに対して、各カテゴリに適切な問題を配信するシステムを提案する。本システムでは、学習者カテゴリごとの習熟度を逐次測定し、その結果にもとづいて配信する問題集合を動的に決定する。その際、配信する各問題に対する難易度情報を前提としておらず、各問題の難易度の設定が困難な状況にも適応可能である。また、各カテゴリに対する適切な問題集合の割り当ては、計算困難な組合せ最適化問題により定式化されるが、遺伝的アルゴリズムを用いることで適切な問題集合の割り当てを効率的にすることができる。

## 2. 学習者適応型 e ラーニングの提案

### 2.1 動的な問題配信システム

学習者適応型 e ラーニングシステムとして、難しすぎる問題や易しすぎる問題を出題しないように学習者の能力水準に近い問題を動的に配信するシステムが提案されている [1][2]。

そのような学習者適応型 e ラーニングシステムの多くは、配信する各問題の難易度等の設定が必要である。しかしながら、問題の難易度の適切な設定が困難な状況では、そのようなシステムの適用は難しい。また、各問題の難易度はすべての学習者に対して一様とは限らない。そのため学習者ごとに、あるいはなんらかの属性によりクラス分けされた学習者カテゴリごとに難易度が設定されるべきである。これらの状況を改善するために本研究では、学習者は年代等の属性により  $k$  個のカテゴリにクラス分けされているものとする。また、配信する教材としてあらかじめ  $kn$  個の問題からなる集合を用意し、その  $k$  分割を求めることで各学習者カテゴリに  $n$  個の問題を割り当てるものとする。その際、それまでの解答状況により定まる各問題の学習者カテゴリごとの難易度にもとづいて、問題の割り当てを遺伝的アルゴリズムを用いて最適化する。

### 2.2 システムの概要

システム利用者は、まず自分が該当するカテゴリを選択する。選択したカテゴリに該当する問題が問題集合 DB から  $n$  問出題され、解答するたびに解答結果 DB へ解答結果が挿入される。解答結果は学習者カテゴリごとに分析

され、各問題に対する学習者カテゴリごとの難易度が設定される。その結果をもとに、学習者カテゴリに対する問題の割り当てを最適化する。本システムのデータフローダイアグラムを図 1 に示す。

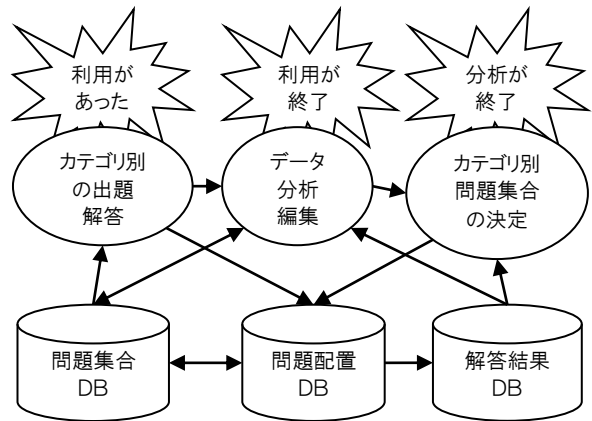


図 1 システム全体図

## 3. 遺伝的アルゴリズムを用いた最適化

本節では、各学習者カテゴリに適切な問題を割り当てる問題を組合せ最適化問題として定式化する。また、その問題の近似解を高速に求めるための遺伝的アルゴリズムについて述べる。

### 3.1 問題の定式化

各学習者カテゴリに問題集合を割り当てる問題は EQUAL SUM SUBSETS を一般化した EQUAL SUM SUBSETS WITH DISJOINT INDICES FROM  $k$  SETS として定式化される。配信する問題の集合を  $P = \{p_1, \dots, p_{kn}\}$ 、学習者カテゴリの集合を  $C = \{c_1, \dots, c_k\}$ 、各問題  $p_i$  に対して学習者カテゴリ  $c_j$  における問題  $p_i$  の重み (難易度) を定めるための重み関数を  $r_{c_j}$  とする。また、すべての  $c_j \in C$  とすべての  $p_i \in P$  に対して  $r_{c_j}(p_i)$  の和をとったものを  $T$  で表す。すなわち

$$T = \sum_{c_j \in C} \sum_{p_i \in P} r_{c_j}(p_i)$$

とする。このとき、各学習者カテゴリに対する問題の割り当ては、次のような組み合わせ最適化問題として定式化される。

EQUAL SUM SUBSET WITH DISJOINT INDICES FROM  $k$  SETS  
 入力：問題集合  $P = \{p_1, \dots, p_{kn}\}$ 、学習者カテゴリの集合  $C = \{c_1, \dots, c_k\}$ 、および各  $c_j$  に対する重み関数  $r_{c_j}$ 。

実行可能解：  $P$  の  $k$  分割  $S = \{S_1, S_2, \dots, S_k\}$ ,  $|S_j| = n$ 。

Development of an Adaptive E-learning System with Genetic Algorithm for EQUAL SUM SUBSETS  
 Kouta TABUCHI, Takayuki NAGOYA, Department of Information Systems, Tottori University of Environmental Studies

目的関数：

$$\max_{1 \leq j \leq k} \left\{ \frac{T}{k^2} - \sum_{p_i \in S_j} r_{c_j}(p_i) \right\}$$

この問題は NP - 困難であることが知られており、多項式時間アルゴリズムは存在しないであろうと強く予想されている。そのため、厳密な解を求めるためのアルゴリズムを設計することは現実的ではない。この問題を解決するため、本研究では、近似解を高速に求めるための遺伝的アルゴリズムを設計した。

### 3. 2 遺伝的アルゴリズムの設計

遺伝的アルゴリズムは、生物の遺伝と進化の仕組みを工学的にモデル化して、様々な問題解法やシステムの学習などに応用しようとするものである。母集団の中から適応度の高い個体を選び、選択、交叉、突然変異を繰り返すことで、より適応度の高い個体を得る。得られる解の精度に保証はないが、多くの場合は最適解に近い解が得られる。

この節では、本研究で設計した EQUAL SUM SUBSETS WITH DISJOINT INDICES FROM  $k$  SETS に対する遺伝的アルゴリズムについて述べる。

**遺伝子型の決定：** 遺伝的アルゴリズムを設計するにあたり、まず遺伝子型を決める必要がある。遺伝子型とは、各実行可能解を表現するためのデータ構造のことで、本研究では  $\{1, 2, \dots, k\}$  の各要素をちょうど  $n$  個ずつ含むような長さ  $kn$  の記号列とする。図 2 に 1 つの実行可能解を表現する遺伝子  $S$  の例を示す。  $i$  番目の要素  $S[i]$  が  $j$  であるとき、問題  $p_i$  は学習者カテゴリ  $c_j$  に割り当てていることを意味する。

	1	2	3	4	...	$kn$
$S$	1	2	3	4	...	1

図 2 一つの実行可能解の遺伝子型

**実行可能解の適応度の決定：** 各実行可能解の良さを表す尺度を適応度という。本研究では、3. 1 節で定めた目的関数を適応度として用いる。目的関数値の小さいものは適応度が高く、大きいものは適応度が低いものとする。

**初期遺伝子集団の決定：** 本研究では、ランダムに  $N$  個の実行可能解を生成しそれを初期世代とする。

**選択：** 現在の遺伝子集団から、エリート戦略により適応度の高い個体を  $e$  個選び次の世代の遺伝子集団の要素とする。

**交叉：** 現在の遺伝子集団から、適応度に応じた確率で 2 つの遺伝子  $S_1, S_2$  を選び、以下の方法により  $S_1, S_2$  を交叉させ、新しい解  $S_3$  を生成する。

- 各  $i = 1, 2, \dots, kn$  に対して、 $S_1[i] = S_2[i]$  ならば  $S_3[i] = S_1[i]$  とする。
- 上記 1 においてどのカテゴリにも割り当てられなかった各問題に対して、ランダムに学習者カテゴリを割り当てる。ただし、各カテゴリに割り当てられた問題数が同じになるようにする。
- 局所探索により上記 1, 2 で得られた遺伝子  $S_3$  を改善

する。

上記 1, 2, 3 の操作を  $N - e$  回繰り返すことで、 $N - e$  個の新しい遺伝子を生成し、次の世代の遺伝子集団に加える。

**突然変異：** 新しく得られた遺伝子集団に生まれる遺伝子の要素をランダムに変更し、遺伝子集団としての多様性、つまりばらつきを大きくする。

上記の選択、交叉、突然変異により新しい世代の遺伝子集団を繰り返し生成する。それにより得られた遺伝子の中から、最も適応度の高いものを近似解として出力する。

### 4. システム適用例

鳥取県は 2005 年に鳥取県環境立県アクションプログラムを制定し環境立県を推進している。2007 年には施策の一つである「環境教育・学習者を 20 万人にします」を達成し、新たなフェーズへ進む必要がある。

しかし、年代によって教育機関で環境教育を受けている世代と受けていない世代が存在する。そのため環境に配慮することを突然、強制された世代と環境に配慮することが当たり前な世代とで環境に対する知識及び関心の差がある。この差から環境教育の内容を年代に応じた教育内容にすべきだが、どの年代にどのような教育をすべきか分かっていない。

そこで、学習者に最適な問題を割り当てると同時に各年齢層ごとのデータを収集するために、今回提案した e ラーニングシステムを構築した。このシステム的设计条件を下記 1. ～6. に記す。

- ユーザは「ようちえん・ほいくえん・小学生」、「中学生・高校生」、「18 才～27 才」、「28 才～27 才」、「28 才～64 才」、「65 才～」の 5 個のカテゴリに分ける。
- 一つのカテゴリに対して出題数は 20 問とする。
- 各問題の重みには正解率を使用する。
- 初期世代で生成する実行可能解は 100 個とする。
- 遺伝的アルゴリズムの世代数は 100 世代とする。
- エリート戦略によって選ぶ個体数は 10 個とする。

### 5. まとめ

本研究では、遺伝的アルゴリズムを用いて各学習者カテゴリに最適な問題を動的に割り当てる e ラーニングシステムを構築した。これにより学習者カテゴリごとの習熟度を逐次測定し、各問題の難易度の設定が困難な場合でも遺伝的アルゴリズムを用いることで適切な問題集合を動的に決定することが可能である。今後の課題として今回構築した環境教育のための e ラーニングシステムを評価する必要がある。

### 参考文献

- 稲葉, 松永, 中村, 学習者適応型 e ラーニングの構築・実験, 情報技術フォーラム講演論文集, 7(4), pp. 55 - 58, 2008.
- 延原, 小山, 三宅, 庄司, 劉, 横田, 学習者の適応度に対応した適応型 e ラーニングシステムの考察, 日本データベース学会 Letters Vol. 3 No. 2, pp. 85-88, 2004.
- 新原将文:ニューロファジィ・遺伝的アルゴリズム, 産業図書株式会社, 1994.