

公平性が保たれた複数のアプリオリアルゴリズム演習問題の自動生成

大谷 紀子 研究室

1972039 北野 守唯

1. 背景と目的

アプリオリアルゴリズムとは、トランザクションデータと呼ばれる取引記録をもとに、「事象 A が発生すると事象 B も発生しやすい」などの相関ルールを求めるアルゴリズムであり、本学では「人工知能とデータマイニング」という講義で扱っている。理解度促進のためには学生ごとに内容の異なる問題をレポート課題として配布することが有用であるが、人手で学生数分の問題を作成するのは難しい。担当教員の負担を減らすために、森は演習問題自動生成システムを開発した[1]。森のシステムでは、生成した数十個の問題の中で、解答過程の要所に基づいて算出した評価値が最も高いものを出力している。一度の実行で生成される問題は1つであるため、現在は繰り返し実行することで複数の問題を生成しているが、各問題が異なるものになる保証がない。また、生成される問題間に難易度のバラつきが生じるため、学生が不満を感じるという問題点がある。

本研究では、「人工知能とデータマイニング」を履修する学生のアプリオリアルゴリズムレポート課題に対する満足度向上を目的として、難易度に大きなバラつきがなく、それぞれ内容の異なる複数の問題を自動生成する手法を提案する。

2. 提案手法

本手法では、アイテム数が x の演習問題を生成するとき、解候補を 2^x 次元のベクトルで表現する。アプリオリアルゴリズムにおいては、 2^x 通りのトランザクションからどのパターンがいくつずつ使われているかが、問題内容を区別する上で重要で

ある。第 $1 \sim 2^x$ 成分が各トランザクションパターンの選択回数を表すことで、成分を比べることによる問題の重複の有無の判定が可能になる。

また、演習問題の学習効果の度合いを計る基準として、総合評価値、個別評価値を用いる。演習問題の最終的な総合評価値は $0 \sim 1$ の値をとり、 1 に近づくほど学習効果が高いことを意味する。総合評価値を計るために、最初に「枝刈り」「サポートカウント計算」「再帰処理」「結論部生成」「相関ルール生成」の5つの観点における個別評価値を算出する。「枝刈り」と「サポートカウント計算」の個別評価値は、 n を処理前のアイテム集合数、 x を処理後のアイテム集合数としたとき式(1)、式(2)で求められる。「相関ルール生成」の個別評価値は、 n をルール生成試行回数、 x をルール成立数としたとき式(3)で求められる。

$$\text{枝刈り評価値} = 1 - \frac{|0.5n - x|}{0.5n} \quad (1)$$

$$\text{サポートカウント計算評価値} = 1 - \frac{|0.8n - x|}{0.8n} \quad (2)$$

$$\text{相関ルール生成評価値} = 1 - \frac{|0.5n - x|}{0.5n} \quad (3)$$

評価値で学習効果を計る手法は森の研究でも採用されていた。本研究ではサポートカウント計算評価値を算出する箇所を増やすほか、ルール成立数が多すぎたり少なすぎたりする場合に、相関ルール生成評価値を 0 にするなど、森の評価値算出方法を本研究の目的に沿った形に修正を加えた上で使用する。「再帰処理」と「結論部生成」の個別評価値は、再帰処理の回数と結論部の最大アイテム

ム数をそのまま各個別評価値とする。すべての個別評価値を足し合わせ、取り得る最大値で割ったものが演習問題の最終的な総合評価値となる。

問題生成手順を以下に記す。

1. 演習問題の元となるトランザクションデータ、最小支持度、最小確信度を生成する。
2. トランザクションデータの一部をランダムに変化させ、問題の総合評価値を算出する。
3. 2.を50回繰り返す。
4. 得られた中で最も高い総合評価値が、問題内容の有用性を保証する最低ライン以上、かつすでに出力された問題と内容の重複がなければ出力する。
5. 2.~4.を5回繰り返し、1.に戻る。
6. 5.の実行中、あらかじめ設定された数の問題が生成された時点で動作を終了する。

3. 評価実験

演習問題としての有用性が評価値に表れているかを検証する実験 A と、同時に生成された演習問題の難易度に大きなバラつきが生じていないかを検証する実験 B の2つの評価実験を実施した。

実験 A では、本学の「人工知能とデータマイニ

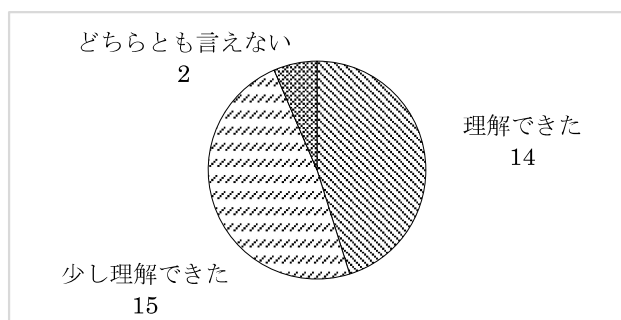


図1 設問解答後のアプリオリ理解度

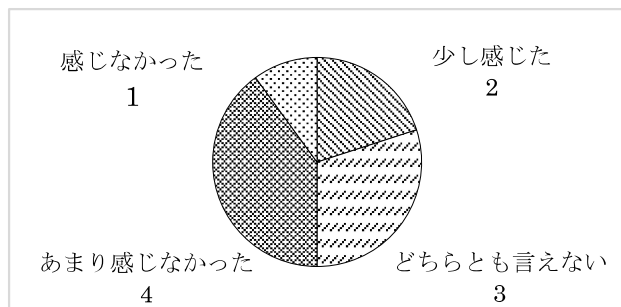


図2 設問内容の不公平さ

ング」を履修中の学生 31 人にシステムで生成された問題を解かせたうえで、アプリオリアルゴリズムに対する理解の度合いを5段階で回答させた。また、問題を解く際に悩んだ箇所を自由記述形式で回答させた。5段階評価における回答の割合を図1に示す。

実験 B では、本学の「人工知能とデータマイニング」を履修済みの学生 12 人を2人ずつの6組に分け、「易しい」「普通」「難しい」の3つの難易度についてそれぞれ2組ずつ取り組ませた。各組に割り振られた難易度の問題を2種類配布し、組内での相談可としたうえで2人に1問ずつ解かせた。解答終了後、2種類の問題間に感じた不公平さの度合いを5段階で回答させ、詳しい理由を自由記述形式で回答させた。5段階評価における回答の割合を図2に示す。

4. 考察

実験 A では、「理解できなかった」と答えた人はおらず、アンケート回答者のほとんどから「アプリオリアルゴリズムの勉強になった」という回答が得られた。よって、本システムで生成される問題は演習問題として有用であるといえる。一方実験 B では、「不公平さを感じなかった」という回答が得られたのは半数程度であり、特に「計算量に差を感じた」という意見が多かった。問題点として、頻出アイテム集合の生成回数に関する評価値の算出方法の設定が不適切である可能性が挙げられる。例えば難易度「普通」の場合、最初のサポートカウントによる枝刈り終了後、アイテム数1の頻出アイテム集合の要素数が4になるか5になるかだけでも、アイテム数2の頻出アイテム集合を求める計算量に差が出てくる。枝刈りの有無による評価値の差を大きくする等、計算量の違いをさらに抑える工夫が必要である。

参考文献

- [1] 森健太郎, “データマイニングの演習問題自動作成,” 武蔵工業大学卒業論文, 2009.