

算数文章題の単文統合型作問プロセスの機械学習によるシミュレーションに関する研究

A Study on Simulation of Single Sentence Integrated Writing Process for Arithmetic Sentences by Machine Learning

伊藤大智^{*2}, 林 雄介^{*2}, 平嶋 宗^{*2}

Daichi ITO^{*2}, Yusuke HAYASHI^{*2}, Tsukasa HIRASHIMA^{*2}

^{*1} 広島大学情報科学部

^{*1}School of Informatics And Data Science, Hiroshima University

^{*2} 広島大学大学院先進理工系科学研究科

^{*2}Graduate School of Advanced Science and Engineering

Email: b181045@hiroshima-u.ac.jp

あらまし: 算数文章題の作問活動において単文統合型作問学習システム「モンサクン」が提案されており、実際の教育現場でも用いられており有用性が認められている。本研究では、機械学習をもとにモンサクンによる作問プロセスをシミュレートして、学習者の理解が不足している作問の成立条件を可視化する。

キーワード: 作問学習, 作問プロセス, 機械学習, シミュレーション

1. はじめに

問題を解くことではなく作ることによる学習を作問学習という⁽¹⁾。作問学習は有効な手段であるとされているものの、解が一意に定まらず教授者や学習者にとって負担が大きいために、実際の教育現場ではあまり行われていない⁽²⁾。この問題を解決するために、作問学習システム「モンサクン」の開発が行われてきた⁽³⁾。モンサクンでは、あらかじめ与えられた量的概念を表す単文を取捨選択し、3つの単文を並び替え、作問学習を行っていく。これまでの研究で、モンサクンを用いて教育現場での実践的利用が行われてきており、学習能力の向上が確かめられている⁽⁴⁾。

一方、モンサクンの課題として、学習者の作問課題を解くプロセスがわからない、学習者が作問課題を間違えたときに、何を理解できずに誤答しているかわからないという課題がある。

本研究では、モンサクンを機械学習で解くモデルを作成して、学習者がモンサクンを学習するプロセスをシミュレートさせる。そのうえで、シミュレーションでつくったモデルと、実際の学習者の解答を照らし合わせることで、学習者モデルを作る。

2. 算数文章題の制約

算数文章題は、ある量の存在を表す文(存在文)2つと、その2つの量の関係を表す文(関係文)1つの3つの文(単文)で表現することができる。この構造を三文構成モデルという⁽⁵⁾。算数文章題の成立においては三文構成モデルに基づいて、「数量」「オブジェクト」「文構成」を満たしている必要があり、また課題の条件である「計算式」「物語」を満たしておく必要がある。つまり、算数文章題における問題構造を理解しているということはこれらの5つの制約条件を理解していることである。

3. 機械学習によるモデルの構築

5つの制約条件と機械学習を用いて、モンサクンを機械学習で解くモデルを作成する。

3.1 強化学習

本研究では強化学習を用いてモデルを構築する。強化学習では、与えられた環境のなかで、エージェントが行動により、最終的に最大の報酬を得られるような方策を習得するための学習を行う。

3.2 「方策」の設定

「方策」は、「ある状態である行動を採る確率」である。本研究では、「方策」を、表1のような「状態数×行動数」の配列の表形式で保持する。「状態数」とは、モンサクンで選択しうるカードの組み合わせで、カードを上から順番に1,2,3,5,6としたとき、カードを3つまで選択できるとしたときのすべての組み合わせを言う。「行動数」とは、「カード番号1を加える」、「カード番号1を減らす」などのモデルが採りうるすべてのカードの増減を言う。

3.3 価値反復法のQ学習^[6]

「価値反復法」は、ある行動を採るたびに、次の状態価値と今の状態価値の差分を計算し、その差分だけ今の状態価値を増やす手法である。

本研究では価値反復法の中でも、Q学習というアルゴリズムを用いてモデル作成をする。

価値を計算する関数として、「行動価値関数」がある。「行動価値関数」は、ある状態である行動を採る価値を計算する関数である。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left(r_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) \right)$$

行動価値関数Qは即時報酬 R_{t+1} と未来に得られる報酬の最大値 $\max_{a_{t+1}} Q(s_{t+1}, a_{t+1})$ に割引率 γ をかけた値の和を学習率 α で調整しながら最適化する。

	カード1 を加える	カード1 を引く	カード2 を加える	カード2 を引く	カード3 を加える	...
0	0.1666	0	0.1666	0	0.1666	
1	0	0.1666	0.1666	0	0.1666	
2	0.1666	0	0	0.1666	0.1666	
3	0.1666	0	0.1666	0	0	
⋮						
123	0	0.3333	0	0.3333	0	
⋮						
456	0	0	0	0	0	

表1 方策の配列

3.4 報酬の設定

本研究では報酬は制約条件を満たしている数に応じて報酬を設定する。したがって、制約条件を5つ全て満たしているカードの組み合わせは最大の報酬、制約条件が不足しているカードの組み合わせはその分減じた報酬を与える。

また、学習者がある制約条件を理解していないモデルを作成しなければならない。

その場合、たとえば「オブジェクト」を理解していないことを想定したモデルを作成するときは、正解のカードの組み合わせの報酬は変更しない。

しかし、制約条件が不足しているカードの組み合わせは、「オブジェクト」の制約を満たしているカードの組み合わせであっても、「オブジェクト」の制約分報酬をへらす。

3.5 実行結果

表1は2つのモデルを100エピソード×100回実行したときの、それぞれのカードの組み合わせが出現した回数である。一つはすべての制約条件を満たしたモデル、もう一つは「オブジェクト」制約を抜きにしたモデルである。カイ二乗検定を行い、2つのモデルの各状態の出現回数に有意な差があるかどうかを検証した。

すべての制約を組み込んだモデルでは、「オブジェクト」制約抜きのモデルより、「オブジェクト」制約を満たすカードの組み合わせが有意に多く出現した。また、「オブジェクト」を抜きにしたモデルでは、「オブジェクト」制約を満たさないカードの組み合わせの出現回数が、すべての制約を組み込んだモデルより有意に多くなった。

このことから、「オブジェクト」制約抜きのモデルで学習させると、「オブジェクト」制約を満たさないカードの組み合わせが出現しやすくなるといえる。

4. まとめ

本研究では、作問課題を解く機械学習のモデルの設計・開発を行った。今後の課題としては、制約をすべて理解している学習者モデルとある制約を理解していないモデルとの学習過程の比較、分析を行う必要がある。

カードの組み合わせ	制約変更なしモデル	制約「オブジェクト」をへらす
123	10000	10000
124	45	26
125	10	3
126	1	27
134	11	5
135	21	30
136	6	6
⋮	⋮	⋮

表1 実行結果

参考文献

- 山元翔, 神戸健寛, 吉田祐太, 前田一誠, 平嶋宗, “算数の文章題を対象とした問題構造の教授とその確認としての作問”, JSISE2012 第37回教育システム情報学会全国大会, C5, 2012
- 平嶋宗, “作問学習のモデル化”, 2009年度人工知能学会全国大会(第23回), 2009
- 山元翔, 神戸健寛, 吉田祐太, 前田一誠, 平嶋宗, “教室授業との融合を目的とした単文統合型作問学習支援システムモンサクン Touch の開発と実践利用”, 電子情報通信学会論文誌 D, Vol. J96-D, No. 10, pp. 2440-2451(2013)
- 神戸健寛, 山元翔, 吉田祐太, 林雄介, 平嶋宗, “単文統合型作問学習支援システムの利用効果の問題構造把握の観点からの評価”, 電子情報通信学会論文誌 D, Vol. J98-D, No.1, pp.153-162 (2015)
- Hirashima, T., Hayashi, Y., Yamamoto, S., : Triplet Structure Model of Arithmetical Word Problems for Learning by Problem-Posing, Proc. Of HCI2014 (LNCS8522), pp.42-50 (2014)
- CJCH Watkins, P Dayan - Machine learning, 1992 - Springer Q-learning