

Human-Agent Interaction を導入した 強化学習エージェントによる人工知能デモシステム

岡野 卓矢[†] 北越 大輔[†] 鈴木 雅人[†]
東京工業高等専門学校[†]

1. はじめに

若年層の「理科離れ」を背景に、科学や工学分野に関心のある小中学生に対し当該分野へのさらなる興味を抱かせ、勉学意欲の向上を促すことが求められている[1]。本稿では、人工知能分野の中でも人間の活動と関連づけやすい強化学習に注目し、専門知識のない人々にも当該分野に対する興味を抱いてもらうためのデモシステムの開発を目的とする。

提案システムでは、Human-Agent Interaction (HAI) の枠組を導入し、人間が強化学習エージェントの学習過程に関わることによる当該分野への興味・理解の促進を目指す。また、情報工学分野に関心のある学生を被験者とした主観評価実験によって提案システムの特性評価を行う。

2. 提案するデモシステム

専門知識のない人々にも人工知能分野に対する興味を抱いてもらうため、日本の伝統的な遊びである「だるまさんがころんだ」を題材としたデモシステムを開発した(図1)。

2.1 デモシステムにおけるエージェントの学習

強化学習とは、環境との相互作用にもとづき、エージェントを環境に適応させる機械学習である[2]。提案システムにおける強化学習エージェントは、だるまさんがころんだにおける非鬼(鬼に近づく)プレイヤーとして振舞う(表1)。

エージェントはゲーム開始当初、「鬼が10数えている間のみ動ける」といった知識を持たず、試行錯誤的な行動の結果得られた報酬をもとに適切な振舞(方策)を学習する。方策は、エージェントが観測した状態 s と s に対する行動 a の対 (s, a) に重みを定める関数 ω で表現され、エージェントは一般的に ω の値に比例した割合で行動選択する。本研究におけるエージェントの方策学習には、Profit Sharing が適用される。時刻



図1 提案システムの概観

表1 提案システムにおける強化学習の設定

状態 s	鬼が発する音声とエージェント自身の速度
行動 a	鬼へと向かう方向を正とした加減速
報酬 r	鬼に近づいた場合 ($r > 0$)
	鬼が無言の時に動いてしまった場合 ($r < 0$)

$t = T$ においてエージェントが状態行動対 (s_T, a_T) を選択した結果 r が与えられた時、 ω の値は(1)式に従って更新される (γ : 減衰率, $0 \leq \gamma \leq 1$)。

$$\omega(s_t, a_t) \leftarrow \omega(s_t, a_t) + \gamma r^{T-t} \quad (t = 0, 1, \dots, T) \quad (1)$$

2.2 システムへのHAIの導入

近年、人間-エージェント間の対話的やりとりを通してエージェントへの親近感を抱かせるHuman-Agent Interaction (HAI) [3]の枠組が注目され、HAIによるエージェントの学習支援に関する研究成果も報告されている。

提案システムでは、人間の「褒める」や「叱る」といった働きかけをエージェントへの報酬 r_H という形式で実現する(表2)。これは実世界における躾や教育に相当するため、人間は違和感なくエージェントに働きかけることができ、エージェントの学習過程を自然に受け入れられると予想される。また、エージェントは人間の働きかけの適切さに応じて、アイコンを用いて

表2 HAIにもとづく学習支援

人間の働きかけ	褒める	叱る
与える報酬 r_H	$r_H > 0$	$r_H < 0$

An AI Demonstration System Employing Reinforcement Learning Agent based on the concept of Human-Agent Interaction

[†]Takuya Okano, Daisuke Kitakoshi, Masato Suzuki, Tokyo National College of Technology



図2 エージェントの感情表現

感情表現する(図2)。人間はエージェントの感情変化にもとづき、エージェントに適切な振舞を獲得させる方法を学習することとなる。

3. デモシステムの特性評価

情報工学分野に関心のある本校の学生(1,2学年)に対して提案システムを用いたデモを実施し、主観評価実験を行った。なお、本校において1,2学年を対象とした人工知能分野に関する授業は実施されていない。アンケート結果の考察には、データ間の依存関係を有向グラフで表した確率モデルであるベイジアンネットワーク(BN)を用いた。BNでは、関連性のある2ノード間にリンク(矢印)が描かれる。表3に、構築するBNにおけるノード(=アンケートの質問項目)の一例を示す。表中の各ノードは4段階評価だが、「客観理解」のみ自由記述方式のため、被験者がどれだけ強化学習について理解しているのか、回答をもとにこちらで採点・評価した。BNの構造は、アンケート結果をサンプルとして情報理論的な学習法により決定した(図3)。

アンケート結果から、エージェントの感情表現によって、エージェントの学習を促進する方

表3 構築するBNの各ノード(抜粋)

ノード名	ノードの意味(質問への回答)
事前知識	デモを見る以前に、人工知能の分野についてどれだけ知っていたか
興味変化	デモを見る前後で、人工知能の分野への興味はどのように変化したか
感情表現	エージェントの感情表現により、学習を促す働きかけ方はわかりやすくなったか
客観理解	強化学習がどのようなものかわかったか

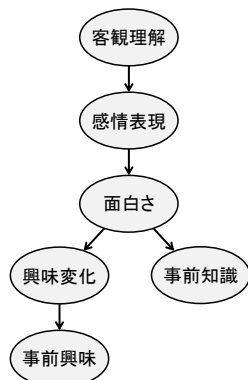


図3 主観評価実験から構築したBN

法が被験者にとってわかりやすくなるという傾向が示された。さらに、被験者がエージェントの感情表現(=働きかけ)を受け入れられれば、当該分野への理解をより深めることや、デモをより面白く感じて当該分野への興味喚起に繋がると期待できる。また、被験者がデモを面白いと感じれば、事前知識の有無に関係なく当該分野への理解や興味の喚起ができることも確認された。なお、当該デモシステムにおける過去の研究では、人間の働きかけのみを導入したシステムについて検証し、

- 被験者の年齢層や事前知識の有無が当該分野への理解に直接的には影響しない
- 当該分野への興味を喚起できれば、被験者の持つ事前知識が乏しくても理解を促進可能

といった結果が確認されていたが、エージェントの感情表現を加えた本実験でも同様の傾向が示された。

4. まとめ

本稿では、人工知能分野に関する専門知識のない人々にも当該分野に対する興味を抱かせるため、HAIを導入した強化学習エージェントによる「だるまさんがころんだ」を題材としたデモシステムを開発した。また、デモ対象者を被験者とした主観評価実験により、デモシステムにHAIを導入した場合の効果を検証した。

実験結果から、HAIの導入によってデモ対象者の個人的な特徴に関わらないデモの提供が可能となることに加え、デモ対象者がエージェントの働きかけを受け入れられれば、当該分野への理解や興味をより深められることを確認した。その一方、現在の設定ではエージェントの働きかけを自然に受け入れられない被験者も少なくないため、より多くの人に受け入れられるようなHAIの枠組を実装していく必要がある。

参考文献

[1] 丸田拓和, 西村亮, 北越大輔, 鈴木雅人, “エージェントの行動獲得過程を効果的に提示するデモシステムの開発とその評価”, 第61回人工知能学会先進的学習科学と工学研究会予稿集, pp. 55-60, 2011.

[2] Richard S.Sutton, Andrew G.Barto (三上貞芳, 皆川雅章 訳), “強化学習”, 森北出版株式会社, 2000.

[3] 山田誠二, 角所考, “適応としてのHAI”, 人工知能学会論文誌, Vol.17, No.6, pp. 658-664, 2002.