

Seq2seq 学習による論理語を含む 言語指示の理解とロボット行動の生成

Understanding of Linguistic Instructions Including Logic Words
and Generation of Robot Behavior by Seq2seq Learning

山田竜郎 村田真悟 有江浩明 尾形哲也
Tatsuro Yamada Shingo Murata Hiroaki Arie Tetsuya Ogata

早稲田大学 理工学術院
Faculty of Science and Engineering, Waseda University

Existing studies that attempted to solve symbol grounding problem on robots have dealt with words such as verbs, objectives, adjectives and adverbs, which directly correspond to objects in the environments, robot motions, or their certain features. The current study builds a seq2seq learning model that translates linguistic instructions to robot behaviors, and in our experiment, not only such grounded words but also logic words, which are not grounded in the world directly, such as “true”, “false”, “and”, and “or”, are learned together. By combinatorially synthesizing these words and visual inputs, the trained model generates appropriate robot behaviors. We also report the model’s internal representations of the relationships between instructions and robot behaviors.

1. はじめに

人間の指示に応じて行動するロボットを目指して、言語と行動の関係性をニューラルネットワーク (NN) に獲得させる試みがこれまで行われている。モデルの性能向上を図る一方で、言語と行動の関係性、すなわち語や文の意味が、モデルによっていかに表象されるかという知見を得ることも、記号接地構造 [Harnad 90] の理解の観点から重要であり、詳細な解析が行われてきた。既存研究は、動詞、名詞、形容詞といった直接実世界との接地を持つ単語群のみを扱ったものが殆どであった [Sugita 05, Heinrich 14]。しかし言語表現には “not”, “and”, “or” のように、それ自体は接地を持たないが論理操作として働くことにより意味に寄与する語も含まれるため、これらの語の表現の解析も必要である。

[Yamada 16] は、論理語を含むタスクとして、NN にロボットの旗揚げゲームを学習させ表現の解析を行ったが、扱われた論理語は “true”, “false” の二つのみであった。そこで、本研究は、これに “and”, “or” を加えたタスクを設計する。その上で、近年主に自然言語処理の分野で用いられる Seq2seq 学習 [Sutskever 14] を応用し、言語指示からロボット行動への変換学習を行う。学習後のモデルにいかにか名詞や動詞などの接地を持つ語群と、論理語の双方の表現が獲得されるか報告する。

2. 言語行動変換 Seq2seq 学習モデル

本研究が用いる Seq2seq モデルの概要を述べる。モデルは図 1 に示す 3 層構造の NN であり、中間層は LSTM 層である。各時刻において入力層は、現在のロボットの関節角度 $j(t)$ 、視覚情報 $v(t)$ 、単語 $w(t)$ を受け取る。LSTM 層はこれらの入力と、前時刻のメモリセルの状態および自身の出力を元に、現在時刻の出力を計算する。出力層は LSTM 層からの全結合で、次の時間のロボットの関節角度 $j(t+1)$ を生成する。この値は、次の時刻のモデル入力としてもフィードバックされる。

指示の受理と行動の生成は共に連続する順伝播計算によって行われる。指示、視覚情報、ロボット姿勢が中間層のメモリセルの状態としてエンコードされ、指示入力後、それをデコードすることで対応する行動が生成される。本モデルはエンコー

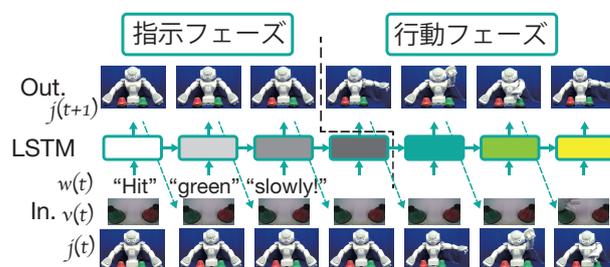


図 1: Seq2seq モデルによる指示からロボット行動への変換。

ダとデコーダの切分けを持たず単一の NN のみからなり、また指示フェーズと行動フェーズの明示的な切替えも行わない。フェーズ切替えも含め、時系列データの経験から総合的に学習する (指示フェーズも関節角度を生成し続けるが、待機姿勢を正解として学習を行う)。学習は、Back propagation through time アルゴリズムによって行われる。

3. 学習実験

3.1 タスク設定

論理語を含むタスクの一つとして旗揚げゲームを扱った (図 2)。ロボットは左右の手にそれぞれ RGB の旗のいずれかを持つ。指示のパターンとして、基本形の ["Red", "Green", "Blue"] * ["up", "down"] * ["true", "false"] に加え、目的語を “and”, “or” によって繋ぐケースを含めた。例えば、赤と青の旗を持っている際に “Red and blue up true.” と指示された場合、両方の腕を上げる。 “Red or blue up true.” の場合、どちらか一方の腕のみ上げるのを正解とする。旗の持ち方、待機姿勢、言語指示の組み合わせとして、576 通りの可能な状況を含むタスクである。

3.2 データ表現

タスクの実行を 13 次元の時系列データとして表現した。ロボットの関節は左右の Shoulder Pitch のみを用いた。UP の場合 0.8, DOWN の場合 -0.8 の値をとる。視覚情報は 3 次元で、各ノードが R, G, B に対応し、その色の旗が左腕にある

連絡先: 尾形哲也, 早稲田大学理工学術院, ogata@waseda.jp

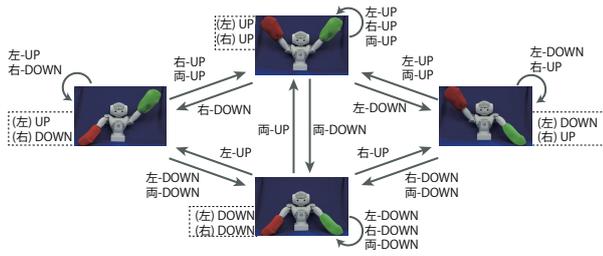


図 2: 旗揚げゲームの概要.

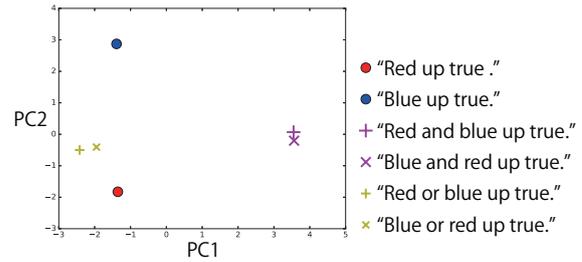


図 4: “and”, “or”の表現.

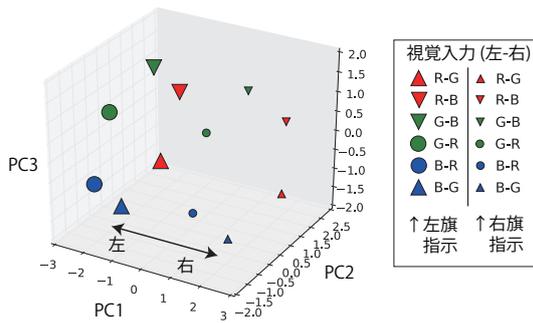


図 3: 視覚情報と色語の入力を総合することによる左右の表現.

場合は 0.8, 右であれば -0.8 , いずれにもない場合は 0.0 の値をとる. 言語は 8 次元であり, 各ノードが 1 単語に対応し, 指示文は 1-hot ベクトルの系列として表現される.

3.3 結果

中間層のノード数を 100, 最適化アルゴリズムを Adam (学習率: 0.001) として 10,000 回学習を行った. 学習後のモデルが, 全ての可能な状況で正しい行動を生成可能であることを確認した. その上で, 異なる指示を与えたときのメモリセルの状態を主成分分析によって可視化した.

まず図 3 に, 持つ旗の色が異なる状況のロボット (姿勢は DOWN-DOWN) に, “(左の旗の色語)-up-true”, “(右の旗の色語)-up-true” と指示を与えた後のメモリセルの状態を示す. PC2-3 の空間にはそのとき入力として受けている色の情報がそのまま埋め込まれているが, 与えられた色語が左右いずれの旗を示すのかが PC1 の方向に表現されていることがわかった.

次に R, B の旗を持っているロボットに and 指示, or 指示を含む異なる指示を与えたときの, メモリセルの状態を図 4 に示した. 両腕を上げることを指示する “and” は PC1 の方向に離れて表現され, “Red” と “Blue” は PC2 に表現されている. そして R, B いずれかを上げることを示す “or” は, その中間として表現された. これより, “or” がモデルのダイナミクス上のサドルとして表現されていることが示唆される. 実際に何度か同じ or 指示を繰り返したとき, ロボットはあるケースでは左腕を, 別のケースでは右腕を上げることが確認された.

最後に, “up”, “down” と “true”, “false” からなる X-OR 問題については, [Yamada 16] と同様, より低次の PC4-5 の空間にこれを解決する成分が見られた.

4. まとめと展望

本研究は, 設定した旗上げゲームにおいて, 接地語と論理語の双方からなる文章の意味が, 学習からいかに NN モデルの内部状態として構造化されるかを解析した. 前者である色語は視覚情報との総合により, 左右に対応する表現に到達した. 論理語もそれぞれ, 今回のタスクにおいて生成すべき行動に合わせた形式で構造化されていることが明らかになった. より複雑なタスクにおける解析や, ダイナミクスとしての解析を今後の展望とする.

謝辞

本研究は, 文科省科研費基盤研究 (A) (No. 15H01710), 科研費若手研究 (A) (No. 16H05878), JST, CREST の助成を受けた.

参考文献

- [Harnad 90] Harnad, S.: The Symbol Grounding Problem, *Physica D*, Vol. 42, No. 1-3, pp. 335-346 (1990)
- [Heinrich 14] Heinrich, S. and Wermter, S.: Interactive Language Understanding with Multiple Timescale Recurrent Neural Networks, *Artificial Neural Networks and Machine Learning — ICANN 2014, Lecture Notes in Computer Science (LNCS)*, Vol. 8681, pp. 193-200 (2014)
- [Sugita 05] Sugita, Y. and Tani, J.: Learning Semantic Combinatoriality from the Interaction between Linguistic and Behavioral Processes, *Adaptive Behavior*, Vol. 13, No. 1, pp. 33-52 (2005)
- [Sutskever 14] Sutskever, I., Vinyals, O., and Le, V. Q.: Sequence to Sequence Learning with Neural Networks, in *Neural Information Processing Systems 2014 (NIPS2014)* (2014)
- [Yamada 16] Yamada, T., Murata, S., Arie, H., and B, T. O.: Dynamical Linking of Positive and Negative Sentences to Goal-Oriented Robot Behavior by Hierarchical RNN, *Artificial Neural Networks and Machine Learning ICANN 2016. Lecture Notes in Computer Science*, Vol. 9886, pp. 339-346 (2016)