

階層型 RNN を用いたロボットの旗揚げタスクにおける 肯定及び否定指示の理解

Connectionist Representations of Grounding of Positive/Negative Sentences
to Robot Behavior in Flag Game

山田竜郎^{*1} 村田真悟^{*2} 有江浩明^{*3} 尾形哲也^{*3}
Tatsuro Yamada Shingo Murata Hiroaki Arie Tetsuya Ogata

^{*1}早稲田大学 理工学術院 基幹理工学研究科 表現工学専攻

Department of Intermedia Art and Science, Graduate School of Fundamental Science and Engineering, Waseda University

^{*2}早稲田大学 理工学術院 創造理工学研究科 総合機械工学専攻

Department of Modern Mechanical Engineering, Graduate School of Creative Science and Engineering, Waseda University

^{*3}早稲田大学 理工学術院

Faculty of Science and Engineering, Waseda University

Meanings of language expressions are constructed from not only words grounded on matters in the real world, but also words participating in the construction by working as logical operators instead of per se being grounded directly on the world, such as “not”. This study investigates how connectionist models learn and internally represent the functions to deal with both of these word groups and to ground the sentences constructed from them on the corresponding behaviors just by experiencing raw sequential data. In the flag game for experiment, a robot implemented with a recurrent neural network was required to ground imperative positive/negative sentences given as a sequence of words on corresponding sequences of goal-oriented behavior. The analyses of the internal representations unveiled that the network extracted XOR problems included implicitly in the target sequences and solve them by learning to represent the logical operations in its non-linear dynamics in a self-organizing manner.

1. はじめに

記号接地構造 [Harnad 90] の理解を目的として、順伝播型神経回路や再帰結合型神経回路 (recurrent neural network: RNN) を実装したロボットにおける言語表現と行動時系列の統合学習が広く行われている [Cangelosi 04, Ogata 07, Yamada 15]. 従来研究の学習実験の多くは、ロボットに与えられる指示文の各単語が、当該環境における動作対象 (名詞) や、動作時系列の低次元要素特徴 (動詞や副詞等) に対応するように設計された [Sugita 05, Tuci 11]. また [Chuang 12, Stramandinoli 12] は、すでに低次元概念として獲得済みの対応関係を明示的に利用するアルゴリズムにより、概念的に高次元行動と言語表現の統合を試みた。

一般に言語表現は、抽象度の差はあるが、実世界の対応する物事と紐づく。上記のように従来の統合学習は、単語レベルでの実世界との接地、さらにはそれらを元に構成される語句レベルでの接地の問題を扱うものであった。しかしながら一方で、言語表現を論理学の記法に還元する際に、論理演算子として置き換えられる否定語 “not” のような、それ自体は実世界内の要素に全く対応せず、論理操作として働くことのみ文の意味を構成する語もある。そのため、例えば “Close the door.” と “Don’t open the door.” により同一の行動が指示され得る。従来のコネクショニストモデルによる統合研究はこれらの語を扱っていない。

そこで本研究はコネクショニストモデルにおいて上述の、文の意味の構成に異なる仕方で関わる語群—すなわち (1) 実世界の対象や行動に接地する語群に加え、(2) 論理操作として働く語群 (今回は肯定・否定語に限定する)—を包括的に処理し、適切な行動に紐付ける能力が、いかに生の時系列データとして与

えられる教示の経験から、事前知識や外部的枠組み無しで総合的に獲得され、また内部表現され得るか、ロボット学習実験を通して考察する。

2. タスク設定

前節で述べた互いに異なる性質の語群を、最低限ではあるが顕著に含むタスクとして旗揚げゲームを扱う (図 1)。旗揚げにおいてロボットが正しい行動を生成するために解決すべき課題は以下のように分析される。

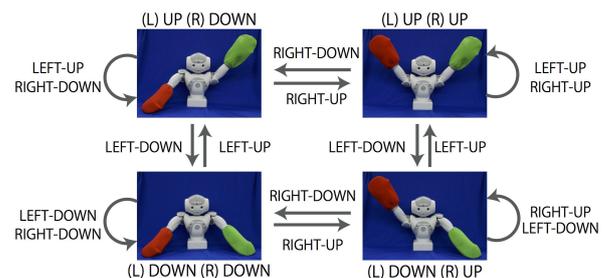


図 1: 旗揚げゲームの概要。指示は 3 語文によって与えられる。可能な指示は (赤, 緑) * (上げ, 下げ) * (て, ないで) の組み合わせからなる 8 通り。また旗の左右は入れ替わる。ロボットは指示と旗の色から、LEFT-UP, LEFT-DOWN, RIGHT-UP, RIGHT-DOWN の 4 つのゴール指向的行動のうち適切なものを選択しなければならない。ロボットの各姿勢において実際に 4 つの行動を実現する運動は、本図に示した矢印に従う。

課題 1a (環境情報との統合) ロボットは左右の手に赤と緑の旗を持つ。どちらの手に赤旗を持つかによって単語 “赤 (red)” がどちらの腕 (LEFT/RIGHT) を指すか異なる。

連絡先: 山田竜郎, 早稲田大学理工学術院, 東京都新宿区大久保 3-4-1, 03-5286-2742, t-a-t-s-u-r-o.y@fuji.waseda.jp

課題 1b (ロボット自身の状態との統合) 例として左手に赤旗を持つ際, “赤-上げて (Red up true.)”は, ゴール指向的行動 LEFT-UP を意味するが, そのときロボットがどの姿勢であるかにより, とるべき運動は異なる. 左腕を下げていれば上げる運動を生成し, すでに上げていればその姿勢を維持する.

課題 2 (論理操作) “上げて (up true)”と “下げないで (down false)”は同じ意味 UP を指す. 同様に “下げ (down true)”と “上げないで (up false)”は DOWN を指す.

課題 1a, 1b は従来の統合学習事例の範疇であり, 実世界の対象 (LEFT/RIGHT) やゴール指向行動 (UP/DOWN) と, それらに対応する言語表現 (red/green, up/down) からなる句の接地関係を学習する問題である. 本研究の新規性は, それに加え, 論理演算子として働き肯定 / 否定文を構成する “て/ないで (true/false)”の処理を, 課題 2 として要求する点である. 具体的には, 旗揚げにおいてロボットは, up/down と true/false の XOR 問題を解決し, UP/DOWN どちらの行動をとるべきか選択しなければならない. ただしこの XOR の学習に関して, 実際に生成すべき運動はロボットの現在の姿勢に依存するため, 単純な二値的教示としては与えられない. 本研究は, 異なる性質の語群の処理がそれぞれいかに表現獲得されるかを調査するものであるため, 品詞に関する予備知識なども与えない.

3. 実験

3.1 学習モデル

本研究では学習器として, コンテキスト層に応答速度の異なるニューロンセットを階層的に重ねた RNN [Yamashita 08] を用いる (図 2).

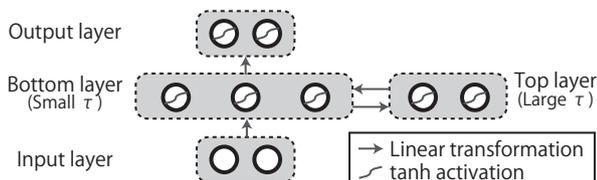


図 2: 異なる時定数のコンテキスト層を 2 つ重ねた階層型 RNN.

コンテキスト層の i 番目のニューロンの時刻 t における内部状態 $u_{t,i}$ は以下で表される

$$u_{t,i} = \left(1 - \frac{1}{\tau_i}\right) u_{t-1,i} + \frac{1}{\tau_i} \left(\mathbf{w}_i^T [\mathbf{x}_t; \mathbf{c}_{t-1}] + b_i\right). \quad (1)$$

$\mathbf{x}_t, \mathbf{c}_t$ はそれぞれ時刻 t における外部入力ベクトルとコンテキスト層の出力ベクトル, \mathbf{w}_i, b_i はそれぞれ, 学習によって最適化される, i 番目のニューロンへの結合重みとバイアス項である. ニューロンの時定数は τ_i によって定められ, 小さいほど入力に対し即応的なニューロンとなる. 今回 0 ~ 99 番目のニューロン (ボトム層とする) の時定数は 2, 100 ~ 129 番目 (トップ層とする) の時定数は 15 として設定した. またトップ層と入力層は直接結合されず, ボトム層を介して情報が交換される. 即応層と非即応層を持つことで, 言語入力や視覚情報の柔軟な受取と, 安定的な姿勢制御の双方が可能である.

3.2 教示方法

一連の旗揚げゲームのシーケンスデータを教示とする学習をおこなう. データは単語 6 次元, 腕関節 2 次元, 視覚 1 次元の 9 次元ベクトルのシーケンスとして構成されるが (図 3), 上述のように品詞情報や, また言語と行動の組を示す明示的サインなども与えない. RNN は誤差逆伝播アルゴリズム [Rumelhart 86] を用いた勾配法により, 単に, 一連の旗揚げゲームを表すマルチモーダルデータの次時間の状態を予測するよう学習する中で, ゲームを遂行する機能を自己組織的に獲得する.

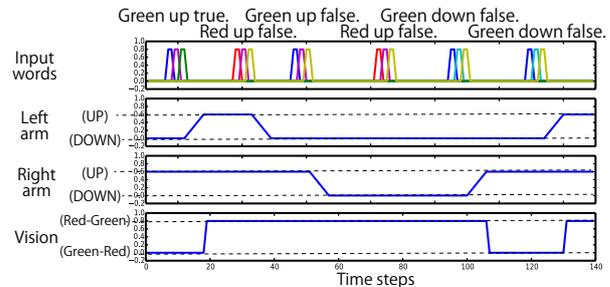


図 3: 一連の旗揚げゲームを表す教示データ

3.3 結果

学習後, 第一にパフォーマンス評価のテストを行った. テスト時は, 指示と視覚情報が外部入力として与えられるのに対し, 左右の腕の関節角度は RNN の出力自身が次時間の入力として用いられる. この順伝播計算をロボットの自律的な行動生成と解釈することができる. 実際にこの条件でロボットは, 指示入力と旗の色の情報から正しい行動を生成し, 旗揚げゲームを遂行し続けることに成功した.

その上で, 旗揚げゲームの遂行がいかに内部表現されているかを解析するため, テスト中のコンテキスト層の内部状態を主成分分析によって可視化した. まず両腕を下げている状態のロボットに, 8 パターンの指示文それぞれを入力したあとのボトム層の内部状態を調べた (図 4). 旗の持ち方は両ケース含む. PC1, 2, 5 にそれぞれ, true/false, up/down, red/green の対に従う成分が見られる. ただしここで, 旗の持ち順によって red/green の対の表現が逆転している. すなわち PC5 には, 目的語 (色語) を旗の持ち順に関する視覚情報と統合し, LEFT/RIGHT にグラウンドした成分が表現されている (課題 1a).

次に, 問題の論理操作がいかに行われているかを解析するため, 旗の持ち方を左腕に赤, 右腕に緑と固定して同様に 8 パターンの指示を行った (図 5). PC1~PC3 の方向に, 先と同様各品詞の表現が現れていることがわかる. ここで RNN は PC1,3 に表現された true-false と up-down の XOR 問題を解決し, UP と DOWN にマッピングしなければならない. 実際に PC4 の方向に意味的 UP-DOWN に対応する表現が獲得されていることがわかる. この成分は up/down, true/false と順に入ってくる単語時系列を, 再帰結合を通して非線形に変換することで得られたものである. すなわち RNN は時系列データに暗黙的に含まれた XOR 問題を抽出し, 指示入力をゴール指向的行動に適切に接地するための論理操作を, 順伝播計算の非線形ダイナミクスにおいて表現する形で学習したと言える (課題 2).

以上のように, (1a) 視覚情報と結び付いた語の処理と, (2) 非線形変換による論理操作の末得られたゴール指向行動の表現は, 最終的に, トップ層に安定的に保持された現在のロボッ

