

音列と場所の相互分節化に基づく ノンパラメトリックベイズ言語創発

Language Emergence Based on String-place Mutual Segmentation Hypothesis
Using a Nonparametric Bayes Model

萩原良信*¹ 牛見光*² 谷口彰*² 谷口忠大*¹
Yoshinobu Hagiwara Hikaru Ushimi Akira Taniguchi Tadahiro Taniguchi

*¹立命館大学 情報理工学部
Information Science and Engineering, Ritsumeikan University

*²立命館大学 情報理工学研究科
Graduate School of Information Science and Engineering, Ritsumeikan University

This study aims to construct the process of language emergence based on string-context mutual segmentation hypothesis, which was proposed by Okanoya et al. A string-place mutual segmentation method based on non-parametric bayes model is proposed for modeling the process of language emergence. Simulation experiments for language emergence using multiple robots demonstrates the effectiveness of the proposed method.

1. はじめに

岡ノ谷らは、言語進化の仮説の一つとして「音列と状況の相互分節化仮説」を提唱している [岡ノ谷 10]. この仮説は、人間の祖先是複雑な形式をもった音列を発話できるようになった後に状況にあわせて音列を発話することで、言語を形成していったというものがある。ここで、音列は音素の系列、状況は自身のおかれた内部および外部の環境を意味する。

この仮説では、人間の祖先が言語を形成するために二つの能力が必要であったとしている。一つ目の能力が音列分節化能力であり、二つ目の能力が状況分節化能力である。音列分節化能力は、他の人間が発話した音列をフレーズと呼ばれる音素のまとまりに分節化し、フレーズ列として認識する能力である。図1は、「音列と状況の相互分節化仮説」のイメージ図である。図1左上の音列分節化の例では、“あにえ〜にくいえ”という音列を“あにえ〜”と“にく”と“いえ”の三つのフレーズに分節化している。状況分節化能力は、ある状況を構成する要素（状況要素）に分節化し、その組み合わせとして認識する能力である。図1右上の状況分節化の例では、“狩猟”の状況を、“人が集まっている”と“シカがいる”と“草原にいる”などの状況要素に分節化している。

さらにこの仮説では、二つの分節化能力が独立に進化したのちに、音列と状況が相互に分節化されることで言語が生じたと考えている。図1下の例は、“あにえ〜”というフレーズと“人が集まっている”という状況要素を音列と状況の相互分節化により学習する様子を示している。状況にあわせて発話された音列において、似た状況で同じフレーズが発話されれば、状況に共通する状況要素を表すフレーズが分節化され、異なる状況で共通のフレーズが発話されれば、状況に共通する状況要素が分節化される。このように、音列と状況の相互分節化によって、現在用いられているような単語が形成されていったと推察される。

本研究の目的は、「音列と状況の相互分節化仮説」に基づく計算モデルを構成し、このモデルにより状況要素を表すフレーズを学習する事である。具体的には、音列と状況の相互分節化



図1: 音列と状況の相互分節化仮説のイメージ図

能力をノンパラメトリックベイズの枠組みでモデル化し、このモデルを実装した複数のロボットによる状況要素を表すフレーズの相互学習を実施する。本研究では、状況を場所、状況要素を物体群により構成される場所要素として、音列と場所の相互分節化能力のモデル化と場所要素を表すフレーズの学習を実施した。

2. 音列と場所の相互学習

語彙や場所を学習する計算モデルに関する研究として、谷口らは、人間の教示した発話文とロボットの自己姿勢からロボットが場所に対応する語彙を教師なしで学習する手法を提案している [谷口 15]. また、石伏らは、ロボットの自己姿勢と観測画像からロボットが場所の領域を教師なしで学習する手法を提案している [Ishibushi 15]. これらの研究では、語彙や画像と自己姿勢を統計的に扱う計算モデルにより、ロボットによる場所の学習を教師なしにより実現している。

本研究では、上記の二つの計算モデルに基づき、音列と画像と自己姿勢から場所の要素に対応するフレーズを学習する手法を構成した。ここで、場所は場所の要素によって特徴付けられた領域であり、例えば、“机の前”や“こたつの前”などがある。また、“机”や“こたつ”、“かごに入ったミカン”などが

連絡先: 萩原良信, 立命館大学, 滋賀県草津市野路東 1 丁目 1-1, 077-561-5745, 077-561-5745, yhagiwara@em.ci.ritsumeikan.ac.jp

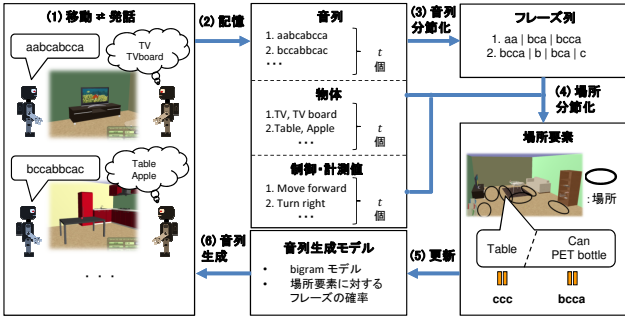


図 2: 提案する音列と場所の相互学習

場所の要素である。

図 2 に提案手法による音列と場所の相互学習の概要を示す。ここでは、二体のロボットが一つの環境において、音列と場所の相互学習を行う場合について説明する。提案手法による音列と場所の相互学習は、以下の (1) から (6) の手順で実施される。

- (1) 二体のロボットが環境中を移動しながら観測画像に基づいた発話を行う。
- (2) 各ロボットは、相手が発話した音列とその場所で認識された物体と制御・計測値を記憶する。
- (3) 記録した音列を音列分節化によってフレーズ列へ分節化する。
- (4) 記録した物体と制御・計測値とフレーズ列から場所分節化によって場所と場所要素を表すフレーズを学習する。
- (5) 学習結果に基づいて音列生成に用いるモデルの更新を行う。
- (6) 更新した音列生成モデルにより音列を生成する。

上記の手順を繰り返す事で、音列と場所の相互学習を行う。

3. 音列と場所の相互分節化モデル

音列と場所の相互学習を可能とする音列と場所の相互分節化モデルを構築した。まず、音列の分節化能力は、教師なし形態素解析手法 [Neubig 10] を用いてモデル化した。また、ロボットの発話する音列の生成には bi-gram モデルを用いた。次に、場所の分節化能力と場所の要素を表すフレーズの学習を一つのモデルとして構成した。構成したモデルのグラフィカルモデルを図 3 に示す。Table. 1 は、図 3 の各パラメータである。このモデルは、谷口らの場所の推定モデル [谷口 15] に、場所ごとに場所要素の混合比をクラスタリングする Dirichlet mixture model (DMM) と場所要素の推定と場所要素を表すフレーズの学習を行う multimodal Hierarchical Dirichlet Process-Latent Dirichlet Allocation (mHDP-LDA) [中村 14] を統合したものである。

場所の推定モデルは、時刻 t における自己姿勢 x_t と場所要素の混合比 θ_t^l と場所ごとの場所要素の混合比 θ_t^p から、時刻 t にロボットが存在した場所のインデックス l_t を推定する。このとき、自己姿勢 x_t は、MCL によって自己姿勢 x_{t-1} と制御値 u_t 、計測値 z_t から推定される。mHDP-LDA は、時刻 t に観測された物体群 v_t とフレーズ列 w_t とロボットが存在した場所のインデックス l_t から、 v_t と w_t に対応する場所要素 c_t を推定する。DMM は、場所の推定モデルと mHDP-LDA を統合する部分であり、時刻 t の場所要素の混合比 θ_t^p がどの場所 l_t を分節化したものかを推定する。

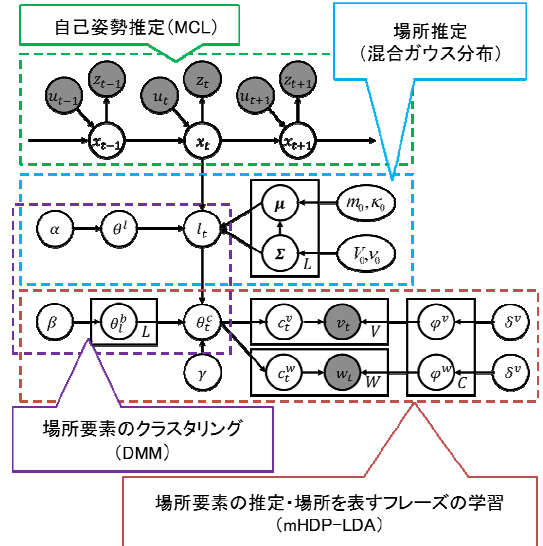


図 3: 提案手法のグラフィカルモデル

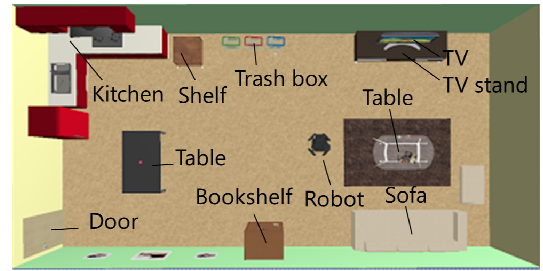


図 4: SIGVerse で構築した実験環境

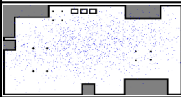
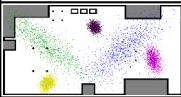
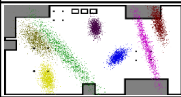
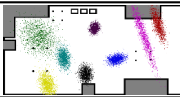
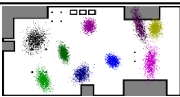
4. 音列と場所の相互学習実験

4.1 実験条件

提案した音列と場所の相互分節化モデルによる場所要素に対応するフレーズの学習をシミュレーション環境において実施した。実験のシミュレータとして、仮想環境中にロボットを構築し、距離センサや画像センサにより周辺環境を計測する機能を持った SIGVerse を用いた。図 4 は、SIGVerse によって仮想環境に構築した実験環境である。実験環境には、"Door", "TV", "Sofa" などの複数の物体が配置されており、ロボットがいる位置によって異なる物体が観測されるようになっている。この実験環境において、計算モデルを実装した二体のロボットが移動と発話を繰り返すことで、状況要素を表すフレーズを学習する。ロボットは、RGB カメラにより画像を取得し、距離センサによって MCL の計測情報、オドメトリにより制御情報を取得できる。ロボットは、事前に環境の地図を構築しており、この地図に基づいて MCL によって自己姿勢を推定する。

4.2 実験結果

図 5 に提案手法による音列と場所の相互学習の結果を示す。図 5(a) の場所の学習結果は、列が step1 から step100 までの学習の試行回数、行が学習した場所とフレーズの種類、フレーズの平均長を示している。まず、学習が進むにつれて場所が分節化され、場所を表現するガウス分布の平均と分散が学習されている事がわかる。次に、フレーズの種類は、学習が進むにつ

	step1	step5	step10	step50	step100
場所領域					
フレーズの種類	45	286	400	411	417
フレーズ平均長	2.42	3.94	4.21	4.09	4.03

(a) 場所の学習結果

	step10	step50	step100	step150	Step200
場所6でのフレーズ生成頻度上位5個	'bead' 0.030	'bb' 0.030	'cdcc' 0.030	'edcc' 0.023	'bcee' 0.030
	'aad' 0.017	'dad' 0.020	'baa' 0.017	'eda' 0.017	'ee' 0.027
	'cac' 0.017	'ca' 0.020	'edcc' 0.017	'dabe' 0.017	'dd' 0.023
	'bcbbdb' 0.017	'e' 0.020	'b' 0.017	'bedab' 0.017	'ecda' 0.020
	'bbac' 0.017	'ada' 0.017	'db' 0.015	'e' 0.017	'cdcc' 0.017

(b) 場所要素に対応するフレーズの学習結果

図 5: 音列と場所の相互学習結果

表 1: パラメータの詳細

x_t	ロボットの自己姿勢
u_t	制御値
z_t	計測値
l_t	場所の index
μ, Σ	場所 (平均と共分散行列)
θ^l	場所の混合率 (多項分布)
θ_t^c	場所要素の混合率 (多項分布)
θ_t^b	場所ごとの基底確率分布 (ディリクレ分布)
c^v, c^w	場所要素の index
v_t	観測された物体
w_t	フレーズ
ϕ^v	場所要素ごとの視覚特徴の生起確率 (多項分布)
ϕ^w	場所要素ごとの単語の生起確率 (多項分布)
$\alpha, \beta, \gamma, \delta, m_0, \kappa_0, V_0, v_0$	ハイパーパラメータ

れて減っており、フレーズの平均長さが増えていることから、場所を表現するフレーズが学習されている事がわかる。

図 5(b) は、任意の場所における場所要素に対応するフレーズの学習結果を示している。列は、step10 から step200 までの学習の試行回数を示している。各試行において、文字列が学習した場所でのフレーズの生成頻度の上位 5 個であり、数値が各フレーズの生起確率である。学習が進むにつれて、いくつかのフレーズの生起確率が高くなっているが、場所を表現する特定のフレーズの学習には至らなかった。実験結果から提案手法による場所と音列の分節化を確認したが、場所を表現するフレーズの学習には至らなかった。場所を表現するフレーズの学習に向けて、ハイパーパラメータの調整や仮想環境での実験条件について再検討したい。

5. おわりに

本研究では、「音列と状況の相互分節化仮説」に基づく計算モデルを構成し、このモデルによる状況要素を表すフレーズの学習方法を提案した。音列と状況の相互分節化能力をノンパラメトリックベイズの枠組みでモデル化し、このモデルを実装した複数のロボットによる状況要素を表すフレーズの相互学習実験を仮想環境において実施した。音列と場所の相互学習実験において、提案モデルによる場所と音列の相互学習を確認したが、場所を表現するフレーズの学習には至らなかった。今後は、実験条件および提案モデルのハイパーパラメータを再検討し、音列と場所の相互学習実験による提案も出るの有用性を言語創発のシミュレーション実験から検証したい。

参考文献

- [岡ノ谷 10] 岡ノ谷一夫：言語起源の生物学的シナリオ。認知神経科学, Vol. 12, No. 1, pp. 1-8, 2010.
- [谷口 15] 谷口彰, 稲邑哲也, 谷口忠大：実ロボットによる場所概念と言語モデルの教師なし同時学習, 計測自動制御学会 システム・情報部門 学術講演会 2015, pp. 458-463, 2015.
- [Ishibushi 15] Satoshi Ishibushi, Akira Taniguchi, Toshiaki Takano, Yoshinobu Hagiwara and Tadahiro Taniguchi : Statistical Localization Exploiting Convolutional Neural Network for an Autonomous Vehicle, 41th Annual Conference of the IEEE Industrial Electronics Society (IECON), pp1369-1375, 2015.
- [Neubig 10] Graham Neubig, 三村正人, 森信介, 河原達也. ベイズ推論を用いた連続音声からの言語モデル学習, 研究報告音声言語情報処理 (SLP), Vol. 16, pp. 1-6, 2010.
- [中村 14] 中村友昭, 西原成, 長井隆行. マルチモーダル LDA とベイズ階層言語モデルを用いた物体概念と言語モデルの相互学習, 人工知能学会全国大会論文集, Vol. 28, pp. 1-4, 2014.