

# 動作と言語のシンボルグラウンディングへの取り組み： 自動運転車の対話的操作を目指して

小林 一郎

お茶の水女子大学 基幹研究院 自然科学系

koba@is.ocha.ac.jp

## 1 はじめに

本稿では、自然言語の意味を制御対象の動作へ結びつけるシンボルグラウンディングについて、株式会社デンソーアイティラボラトリーの塚原裕史氏、および、慶應義塾大学の大前学氏との共同研究である自動運転車の対話的操作に関する取り組み [10] を紹介する。

## 2 発話操作の概要

図1に我々が構築している自動運転車の発話操作手法の概要を示す。

システムに与えられる入力情報は、自動運転車の操作者による自然言語発話、および、自動運転車が自らのセンサーを用いて観測する環境情報となる。自然言語発話を与えられると、空間的意味カテゴリに基づく構文解析パーザにより、Kollar ら [2] により提案されたSDC (Spatial Description Clause) と呼ばれる空間関係を表すことが出来る意味構造に変換される。その情報と車に取り付けられたセンサーにより認識される環境表現の対応関係を捉えるために、SDCの構造に基づいた確率的グラフィカルモデルを生成し、確率伝

搬法により各ノードの確率値を計算することにより、自然言語と実世界空間のグラウンディングを行う。グラフィカルモデルは、Tellex ら [5] により提案された  $G^3$  (Generalized Grounding Graph) を使用する。  $G^3$  には、車の操作も確率変数として含まれており、自然言語での発話から操作対象や目的場所の選択と同時に車の操作も選択される。

以下、各部の処理について概説する。

## 3 運転指示文の空間意味表現変換

### 3.1 空間意味表現を用いた構文解析

本研究では、Combinatory Categorical Grammar (CCG) [4] による構文木の統語カテゴリに相当する部分に、SDCの空間意味カテゴリを採用する。それらのカテゴリを文法規則としたCCGを用い、Shift-Reduce法に基づくパーザを構築した [8]。図2, 3に「赤い車の隣に停めて」という駐車指示をSDC木に変換した例とCCGによる構文木に変換した例をそれぞれ示す。

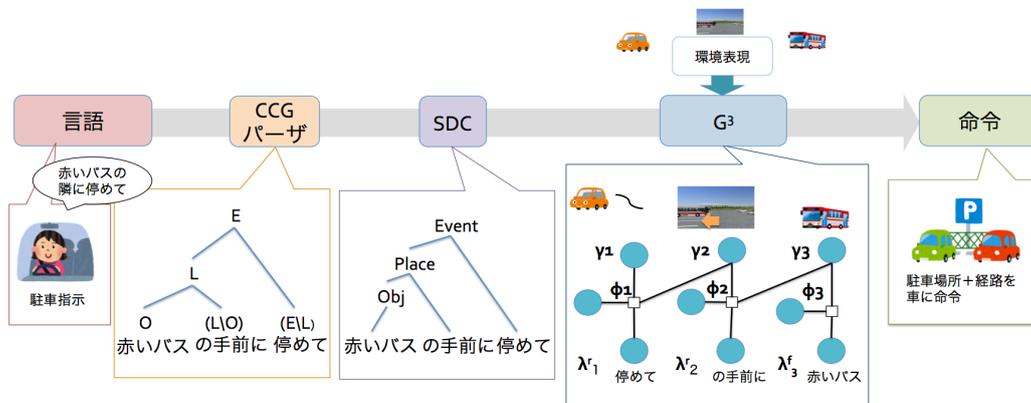


図1: 自動運転車への発話操作の概要

EVENT<sub>1</sub>(r= 停めて  
 l= PLACE<sub>2</sub>(r= の隣に  
 l= OBJ<sub>3</sub>(f= 赤いバス)))

図 2: SDC 木

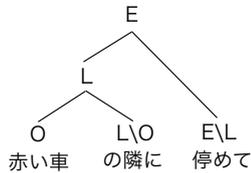


図 3: SDC の意味カテゴリによる構文木

### 3.2 SDC 木への変換

パーザによって解析された木を SDC へ変換において、深さ 1 の部分木ごとに行い、日本語の駐車指示コーパス [8] を参照して得たヒューリスティックな知識に基づいた変換規則を用いる [9, 1]. 図 4 は深さ 1 の構文木 L ( O (赤い車 the red car) (L\O) (の隣に next to)) を SDC PLACE(r=の隣に next to l=OBJ(f= 赤い車 the red car)) へ変換する例である。

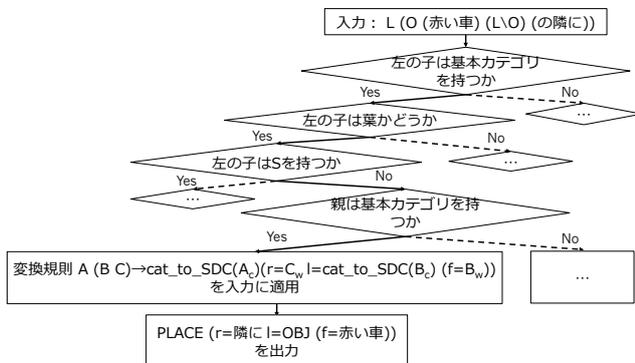


図 4: 深さ 1 の部分木から SDC への変換の例

## 4 確率伝搬法を用いたグラウンディング

$G^3$  の確率の計算は、確率伝搬法と呼ばれる手法を使用する。Tellex らの手法 [5] では、それぞれの環境と言葉の組み合わせを独立に扱い、それを Conditional Random Field を用いてモデル化し、部分グラフの確率を最大にするような近似的な計算を行っていたため、

グラフ全体の確率について必ずしも最大化されているとは限らなかった。本研究では、確率伝搬法による近似を用いてグラフ全体の繋がりを考慮した周辺確率を計算し最大化を行う [3, 6].

### 4.1 グラウンディンググラフの生成

グラウンディンググラフは、SDC へ変換された駐車指示から自動的に生成する [7]. SDC の要素を前から順に見て行き、ルールに従って  $G^3$  を生成することで、SDC の階層的な構造を反映している。

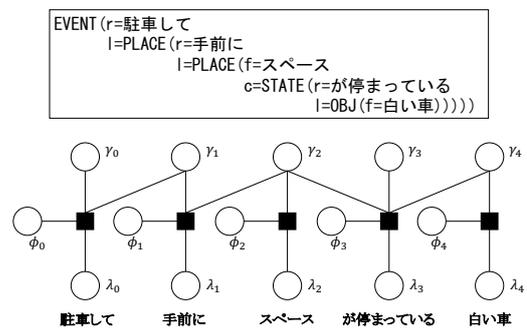


図 5: 例 : SDC と生成された  $G^3$

### 4.2 グラフにおける素性の設定

素性は、空間に関する素性と、言語に関する素性をそれぞれ用意する。空間に関する素性は、因子ノードに接続する環境変数ノードの数と種類ごとに考える。言語に関する素性では、辞書を人手で作成し、異なる表現ではあるが同じ意味を表す言葉を吸収する。

### 4.3 確率伝搬法による周辺確率分布の計算

確率伝搬法は、ノードからノードへメッセージを伝搬させることで、観測されているノードの情報を元に、観測されていないノードの周辺確率を計算する手法である [6]. Algorithm 1 が確率伝搬法のアルゴリズムである。

### 4.4 重みパラメータの学習

式 (6) で与えられる重み因子により  $G^3$  の周辺確率を計算し、得られた周辺確率を用いてパラメータ  $\theta_{\alpha,s}$  の更新を行う。

---

**Algorithm 1** 確率伝搬アルゴリズム
 

---

**1** メッセージの初期化

メッセージの初期化

全てのメッセージを 1.0 に初期化する.

$$m_{\alpha \rightarrow i}^{(0)} = 1, \quad m_{i \rightarrow \alpha}^{(0)} = 1 \quad (1)$$

**2** メッセージの計算

以下の式で, 因子ノードから確率変数ノードへのメッセージの伝搬 (2) と, 確率変数ノードから因子ノードへのメッセージの伝搬 (3) を行う.

$$m_{\alpha \rightarrow i}^{(t+1)}(x_i) \propto \sum_{x_{\alpha \setminus i}} \Psi_{\alpha}(x_{\alpha}) \prod_{j \in x_{\alpha} \setminus i} m_{j \rightarrow \alpha}^{(t)}(x_j) \quad (2)$$

$$m_{i \rightarrow \alpha}^{(t+1)}(x_i) \propto \prod_{\beta \in \partial i \setminus \alpha} m_{\beta \rightarrow i}^{(t)}(x_i) \quad (3)$$

**3** 周辺確率の計算

 収束したメッセージを  $m_{\alpha \rightarrow j}^*$  として, 確率変数ノードの周辺確率 (4) と因子ノードの周辺確率 (5) を以下の式で計算する.

$$p(x_i) \propto \prod_{\alpha \ni i} m_{\alpha \rightarrow i}^*(x_i) \quad (4)$$

$$p(x_{\alpha}) \propto \Psi_{\alpha}(x_{\alpha}) \prod_{j \in x_{\alpha}} m_{j \rightarrow \alpha}^*(x_j) \quad (5)$$


---

$$\Psi_{\alpha}(x_{\alpha} | \theta_{\alpha}) = \exp \left( \sum_s \theta_{\alpha,s} t_s(x_{\alpha}) \right) \quad (6)$$

$\alpha$  は因子ノードを区別するインデックスである. パラメータの更新には, IPF アルゴリズム (Iterative Proportional Fitting Algorithm) を用いる. 更新式は, 式 (7) で与える.

$$\theta_{\alpha,s}^{(q+1)} \leftarrow \theta_{\alpha,s}^{(q)} + \eta \log \frac{\hat{p}_{\alpha}(t_s(x_{\alpha})) - \mu(\theta_{\alpha,s}^{(q)} - 1)}{p_{\alpha}(t_s(x_{\alpha}) | \theta_{\alpha}^{(q)}) + \mu} \quad (7)$$

ここで,  $\hat{p}_{\alpha}(t_s(x_{\alpha}))$  は学習データから与えられる経験分布,  $p_{\alpha}(t_s(x_{\alpha}) | \theta_{\alpha}^{(q)})$  は, 確率伝搬法により与えられる, 因子の周辺分布である.  $\mu$  は, パラメータがオーバーフローすることを防ぐ正則化項であり, この値を導入することで安定した学習を行うことが出来る. また,  $\eta$  は学習率である.

## 5 実験

実験では, 駐車指示内に含まれる物体や位置と, 実環境に存在する物体の対応付けを行った. 素性は定性的な素性のみを用いた場合, 定量的な素性のみを用いた場合, 両方を用いた場合の 3 種類について結果を出した.

図 6 に示す実験結果では, 縦軸が精度, 横軸が学習回数である. 色がそれぞれの素性についての結果を表していて, マゼンタが定量的な素性のみ, 黄色が定性的な素性のみ, シアンが定量的な素性と定性的な素性を用いた場合である. また, all は駐車指示内の全てのグラウンディングが正しい場合の精度, place は場所 (駐車位置) に関するグラウンディングが正しい場合の精度, object は物体 (車) に関するグラウンディングが正しい場合の精度である. 精度においては両方の素性を組み合わせた場合が最も正解率が高く, 次に高かったのが定量的な素性を用いた場合であった. 定性的な素性を用いた場合が最も低かった. 定性的な素性だけを用いた学習で, 正解率が下がってしまった原因は, 素性の数が少ないことが挙げられる. 定量的な素性では, 細かい位置の違いまで区別することが可能だが, 定性的な素性では順序関係を求めるだけであるので, 実世界の特徴を捉えるための情報が不足していたと考えられる.

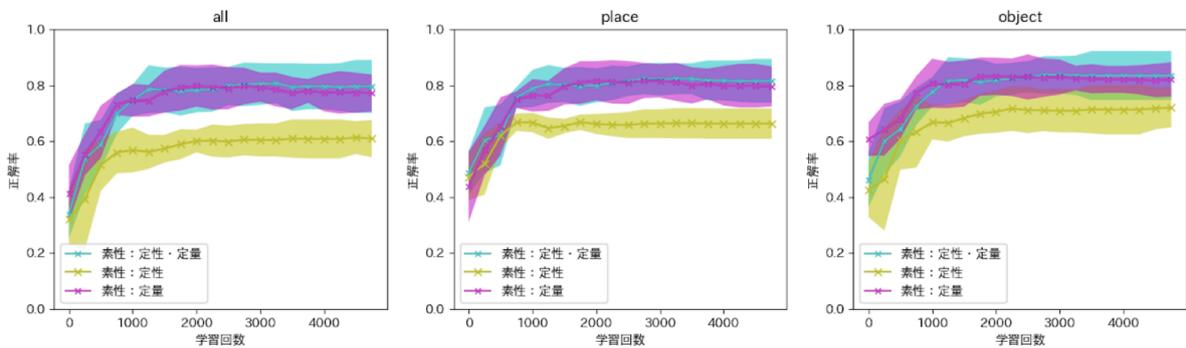


図 6: 言語指示におけるグラウンディングの実験結果 (縦軸: 精度; 横軸: 学習回数)

## 6 おわりに

本稿において、動作と言語におけるシンボルグラウンディングの取り組みを自動運転車への対話的操作を例に示した。表現学習が可能になったことで、複数のモダリティの対応関係をとることが容易になった。しかし、その対応関係に対する明確なメカニズムに対する解析は難しく、かつ、確実であるという保証がない。本稿で紹介した手法は、統計的機械学習の手法を基礎とし、グラウンディングのメカニズムが明確であるものを紹介した。

## 謝辞

本稿における研究内容は、(株)デンソーアイティラボラトリーの塚原裕史氏、慶應義塾大学の大前学先生との共同研究の成果であり、共同研究に関わってくださった大田原菜々さん、稲子明里さん、日垣博考さん、他皆様に感謝申し上げます。

## 参考文献

- [1] Akari Inago, Hiroshi Tsukahara, and Ichiro Kobayashi. Parsing parking instructions for self-driving cars into spatial semantic descriptions. *the 8th International Conference on Intelligent Computing and Applications*, 2019.
- [2] Thomas Kollar, Stefanie Tellex, Deb Roy, and Nicholas Roy. Toward understanding natural language directions. In *HRI '10: Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, pp. 259–266, 3 2010.
- [3] Nana Otawara, Hiroshi Tsukahara, Ichiro Kobayashi, and Atsushi Keyaki. Understanding natural language parking instructions and grounding for self-driving cars. *NeuralPS 2019 Workshop: Machine Learning for Autonomous Driving*, 2019.
- [4] Mark Steedman. *The Syntactic Process*. MIT Press.
- [5] Stefanie Tellex, Thomas Kollar, Steven Dickerson, Matthew R. Walter, Ashish G. Banerjee, Seth Teller, and Nicholas Roy. Understanding natural language commands for robotic navigation and mobile manipulation. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pp. 1507–1514, San Francisco, CA, 8 2011.
- [6] 大田原菜々, 塚原裕史, 小林一郎. 自動運転の対話的操作を実現するための自然言語の実環境へのグラウンディング. 人工知能学会全国大会論文集, Vol. JSAI2019, .
- [7] 大田原菜々, 塚原裕史, 小林一郎. 自動運転の対話的操作を実現するための自然言語の空間意味表現に基づくグラウンディンググラフの生成. 日本知能情報フェジ学会 フェジシステム シンポジウム 講演論文集, Vol. 34, pp. 446–449, 2018.
- [8] 稲子明里, 塚原裕史, 小林一郎. 自動運転の言葉による指示を対象にした空間的意味表現の構造化への取り組み. 人工知能学会全国大会論文集, Vol. JSAI2017, .
- [9] 稲子明里, 塚原裕史, 小林一郎. 自動運転の駐車指示を対象にした空間的意味記述の生成への取り組み. 人工知能学会全国大会論文集, Vol. JSAI2018, .
- [10] 塚原裕史, 日垣博考, 大前学, 大田原菜々, 稲子明里, 小林一郎. 音声によって運転指示可能な自動運転車両の開発. 人工知能学会全国大会論文集, Vol. JSAI2018, .