The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

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

Generating Spatial Semantic Descriptions of Parking Instructions for Self-driving Cars

稲子 明里 *1 Akari Inago　塚原 裕史 *2 Hiroshi Tsukahara　小林 一郎 *1 Ichiro Kobayashi

*1お茶の水女子大学 Ochanomizu University　*2株式会社デンソーアイティーラボラトリ Denso IT Laboratory, Inc.

This paper describes a method to transform a combinatory categorial grammar(CCG) parse tree into a spatial description clause(SDC) to be grounded onto the objects or the phenomena in the real environment. We especially designed the CCG for handling spatial meanings contained in parking instructions for self-driving cars. In previous studies, we developed a parser for generating parse trees of parking instructions according to the CCG we defined. We propose a rule-based converter to perform a conversion from these parse trees into SDCs using heuristic knowledge gained from our analysis on a careful investigation on individual examples. We built the converter and evaluated it on the results of our parser. The result of conversion shows fairly high accuracy for parse trees generated by our parser from parking instructions in the corpus used in this study.

1. はじめに

近年, 車の自動運転の実用化に向けた動きが活発化している. 車を運転できない人でも自動運転車を操作するために, より操作を容易にする手法の開発が期待されている. なかでも, 口頭指示による対話的な操作は, 車への意思疎通を行う手段として今後必要になることが予想される. そこで我々は, 自然言語で表現された操作指示内容と実世界との対応づけ (グラウンディング) を目的とし, 駐車指示内容を車の操作へと結びつける空間意味記述の生成を目指す. 先行研究において, 我々は駐車指示を空間意味記述の前段階となる構文木へ変換する手法を提案し, そのパーザを実装した [Inago 17]. その後パーザの n-best 解のリランキングによって, 精度を向上する手法を提案し, リランカーを実装した [Inago 17-2]. これを踏まえて本研究では, 生成した構文木を空間意味記述に変換する手法を提案する.

2. 関連研究

ロボットへの自然言語の経路指示からロボットが行うべき行動を推定する研究は盛んに行われてきた. Kollar らは自然言語から実世界の空間意味関係を抽出するために Spatial Description Clause(SDC) という意味構造を提案した [Kollar 10]. SDC の構造は, Jackendoff (1983) [Jackendoff 83], Landau and Jackendoff (1993)[Landau 93], Talmy (2005)[Talmy 05] らによって提案された意味構造に基づいている. 統計的機械翻訳に始まり, 言葉と実世界を結ぶグラウンディングの問題を言葉の構文構造に従ったグラフィカルモデルとして扱う取り組みが増えている. Tellex ら [Tellex 11] は SDC の構造に従って生成される Generalized Grounding Graphs(G^3) と呼ばれる確率的グラフィカルモデルを提案した. 近年では, Paul ら [Paul 16] が抽象概念を階層構造に反映させた Adaptive Distributed Correspondence Graph というモデルを提案した. Arkin ら [Arkin 16] は複数の連なった指示を分割して, 個々の文に意味を割り当てる Monologic Distributed Correspondence Graph を提案した. 一方, 最近の自然言語を論理式に変換するための取り組みにおいては, Guu らが周辺尤度最大化を用いた系統的な探索に強化学習のランダムな探索を組み合わせた学習アルゴリズム (RANDOMER) を提案し, ニューラル意味解析器に適用した. この手法は, indirect supervision のみの学習が可能な場合において間違ったプログラムから偶然正しい結果が出力されるという事例の防止を課題として提案された.

連絡先: 稲子明里，お茶の水女子大学大学院人間文化創成科学研究科理学専攻情報科学コース，〒 112-8610 東京都文京区大塚 2-1-1 小林研究室，g1320504@is.ocha.ac.jp

3. 提案手法

我々の手法は, Kollar ら [Kollar 10] と Tellex ら [Tellex 11] の研究に基づいている. 駐車指示に空間的な背景知識を組み込むための意味構造として, Kollar によって提案された SDC を用いた. SDC の生成に関して, 先行研究で Kollar らは条件付き確率場を用いてラベル付けをして直接変換していたが, この手法では文全体の要素の依存関係を捉えることができなかった. また Tellex らは Stanford parser[Marneffe 06] を用いて直接 SDC に変換していたが, 統語カテゴリを用いた構文解析を行っていたため, 空間意味構造への変換には, 制約を伴った入力文しか受け付けられず, ほぼテンプレートマッチングと同様の処理になってしまっていた. そこで我々は, Combinatory Categorial Grammar (CCG) [Steedman 00] において統語範疇のカテゴリを空間意味のカテゴリに変更した文法を新しく導入することでより複雑な入力文から正確な空間意味の依存木を抽出し, それを SDC に変換する手法を提案する. CCG を用いることで, 空間的な意味をラムダ式で表現することが可能になる. ラムダ式を用いることによって, 文章が空間的な意味を正しく捉えられているか否かに対する確認を行うことができる. 本研究では, Kollar らによって定義された SDC における 4 つのタイプに加え, STATE と VIEW の 2 つが追加されている [Inago 17]. これにより, SDC には EVENT, OBJECT, PLACE, PATH, STATE, VIEW の 6 つのタイプが存在し, それらに対応して E(EVENT), O(OBJECT), L(PLACE), P(PATH), S(STATE), V(VIEW) を CCG における基底カテゴリとして扱う. また指示内容と観測した物体とのグラウンディングには, Tellex らが提案した G^3 を拡張して用いる. 図 1 に提案手法の概要を示す. 本研究では, 点線で囲まれた言語,

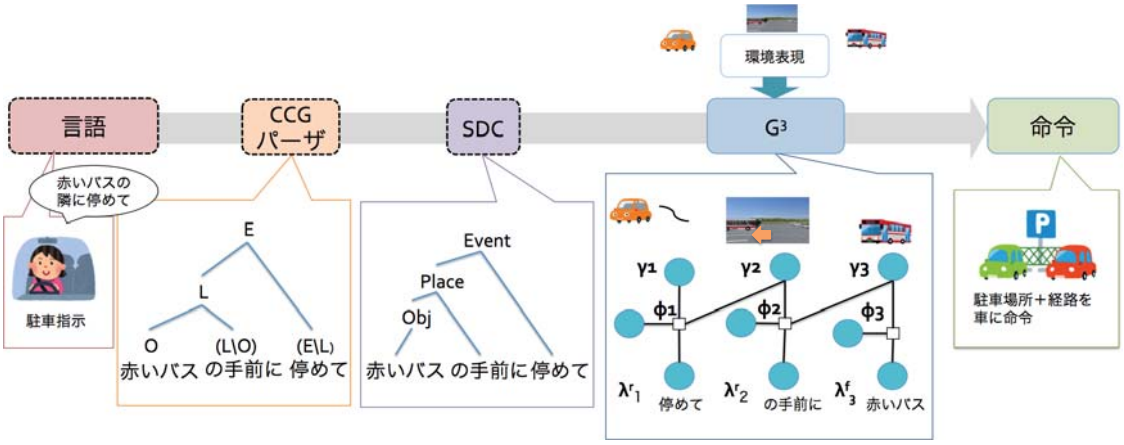

図 1: 提案手法の概要

CCG パーザ, SDC の部分に焦点を当て, 構文木から SDC への変換手法の提案を行う.

4. 構文木から空間意味記述への変換

4.1 変換規則

このセクションではパーザによって解析された木を SDC に変換する方法について説明する. 変換は深さ 1 の部分木ごとに行い, 日本語の駐車指示コーパス [Inago 17] 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. この変換規則は条件に合致するかどうかに従って, 49 個存在する. 条件は大きく 6 個に分かれる. ノードがどのカテゴリをもつか, 子が葉かどうか, 全体の木の深さはいくらか, 親のカテゴリと右の子のカテゴリが同じか, 子がスペースを表す言葉（例えば「駐車スペース」,「場所」,「そこ」など）を持つどうか, 左の子が relation フィールドのテキストを持つかどうかがある. 構文木から SDC への変換は, 以下の 4 ステップに分けられる.

ステップ 1. 深さ 1 の部分木を選択する.

ステップ 2. 部分木の情報に従って, 適切な変換規則を選択する.

ステップ 3. 変換規則に従って, 部分木を SDC に変換する.

ステップ 4. 全ての部分木を変換するまでステップ 1 から 3 を繰り返す.

図 2 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)) に変換する例である. このプロセスは, ステップ 2 と 3 に相当する.

4.2 省略された SDC の補完

駐車指示には, グラウンディングに必要な情報が暗黙的に省略されており, そのまま SDC をへ変換した際に, グラウンディンググラフを生成するために必要な部分が抜けている場合が存在する. そこで, 我々は省略された SDC の部分を補完するための変換規則も定義した. 省略された SDC は,

1. 根の SDC が PLACE である場合
2. 根の SDC が STATE である場合

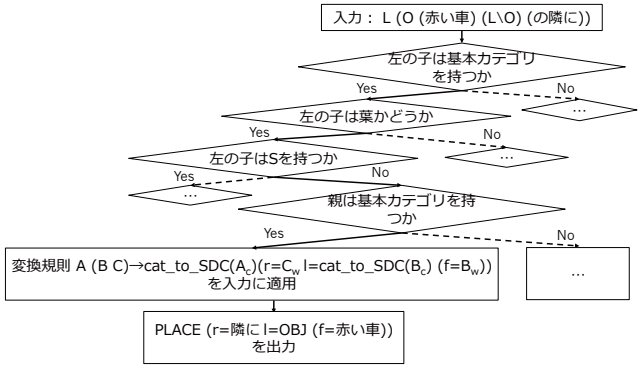

図 2: 深さ 1 の部分木から SDC への変換の例

3. PLACE が relation フィールドに属するテキストを持っているが, landmark フィールドの SDC は持っていない場合
4. PLACE が relation フィールドに属するテキストと condition フィールドに属する SDC を持っているが, landmark フィールドの SDC は持っていない場合

の 4 種類に分類される. 表 1 に省略された SDC の補完規則を示す. *は簡略化した任意の SDC またはテキストである. 式の左辺は省略がある SDC を表し, 右辺は補完後の SDC を表す. 図中では, A が親カテゴリを持つノードを表し, B と C が単語とカテゴリまたは SDC とカテゴリを持つ子ノードを表す. 変換規則の→の左辺が変換前の深さ 1 の部分木, 右辺が変換後の SDC を表す. N_w はノード N が持つ単語 w, N_c はノード N が持つカテゴリ c を表す. cat_to_SDC(C) は基底カテゴリ C から変換された SDC を表す.

5. 実験

先行研究で提案したパーザから生成された構文木から SDC への変換に対する評価実験について述べる.

5.1 実験設定

対象データには, クラウドソーシングによって集められたコンビニエンスストアの駐車場での駐車シーンにおける 6,019 文

表 1: 省略された SDC の補完のための変換規則

Type	Conversion rule	Example
1	PLACE(*) → EVENT(r=停めて l=PLACE(*))	PLACE(f=スペース r=前の l=OBJ(f=歩行者)) → EVENT(r=停めて l=PLACE(f=スペース r=前の l=OBJ(f=歩行者)))
2	STATE(*) → EVENT(r=停めて l=PLACE(f=スペース c=STATE(*)))	STATE(r=が空いてる l=PLACE(r=右端 l=OBJ(f=駐車場))) → EVENT(r=停めて l= PLACE(f=スペース c=STATE(r=が空いてる l=PLACE(r=右端 l=OBJ(f=駐車場)))))
3	(*PLACE(r=*)*) → (*PLACE(r=* l=PLACE(f=スペース c=STATE(r=空いてる)))*)	EVENT(r=停めて l=PLACE(r=真ん中に c=STATE(r=空いてる))) → EVENT(r=停めて l=PLACE(r=真ん中に l=PLACE(f=スペース c=STATE(r=空いてる))))
4	(*PLACE(r=* c=STATE(*))*) → (*PLACE(r=* l=PLAE(f=スペース space c=STATE(*)))*)	EVENT(r=停めて Park l=PLACE(r=右側 c=STATE(r=空いてる))) → EVENT(r=停めて l=PLACE(r=右側 l=PLCAE(f=スペース c=STATE(r=空いてる))))

表 2: 駐車指示コーパスにおける構文木のタイプと事例数

NO.	CCG ベースの文法による構文解析パターン	例文	事例数
1	(E (L (O (L\O)) (E\L)))	入り口の近くに停めて	586
2	(L (O (L\O)))	コンビニ前	562
3	(E (L (O (((O/O) (O (O/O)\O)) O) (L\O)) (E\L)))	黒い車の右の車の右側に停めて	451
4	(L (O (((O/O) (O (O/O)\O)) O) (L\O)))	コンビニの前の白い車の右隣	417
5	(E (L ((L/L) (O (L/L)\O) L) (E\L)))	黒い車から一番離れた場所に停めて	295
6	(L ((L/L) (O (L/L)\O) L))	トイレから反対側の右側	292
7	(E (L (L (O (L\O)) (L\L))) (E\L))	黒いクルマからずっと奥の一番右に停めて	137
8	(L (L (O (L\O)) (L\L)))	コンビニとは反対側の真ん中	131
9	(L ((L/L) (O (L/L)\O) L (O (L\O))))	白い車の横でコンビニ前	128
10	(L (O (S (L (O (L\O)) (S\L)) (O\S)) (L\O)))	歩行者の右側にある白いクルマのさらに右側	117
...	...	...	...
372	(L (S (L\S) ((L\S)/L L)))	空いてる右端のスペース	1
373	(L ((L/L) (S (L (O (L\O)) (S\L)) (L/L)\S) (L (S (L\S)))))	トイレの前ではなくて右側の空いてるスペース	1
374	(E (O ((O/O) (O (S (L (O (L\O)) (L\S)) (O\S)) (O/O)\O) O) (E\O)))	トイレの隣にあるコンビニの入口に停めて	1
375	(L ((L/L) (O (L/L)\O) (L (S (O (S\O)) (L\S)) ((L\L) (O (L\L)\O)))))	白いクルマの歩行者がいない方, コンビニ側	1
376	(L (S (O (S\O)) ((L\S) (V (O (V\O)) (L\S)\V))))	歩行者がいない白いクルマの向かって右側	1
377	(E (S (O (S\O)) ((E\S) (L (O ((O/O) (O (O/O)\O) O) (L\O)) (E\S)\L))))	歩行者は避けてコンビニの入口の前に停めて	1
378	(L (S (O ((O/O) (O ((O/O) O) (O/O)\O) O) (S\O))) (L\S))	一番右側のコンビニ入口の前の歩行者がいないところ	1
379	(E (L ((L/L) (S (L (S\L)) (L/L)\S) (L (S (L\S)))) (E\L)))	真ん中にある目の前の空きスペースに停めて	1
380	(S (L ((L/L) L) (S\L)))	真ん中の駐車場が空いています	1
381	(E (L (S (O ((O/O) (O ((O/O) (O (O/O)\O) O) (O/O)\O) O) (S\O)) (L\S)) (E\L)))	黒いクルマと白いクルマの間の歩行者がいるところに停めて	1

の日本語の駐車指示コーパスを用いた [Inago 17]. クラウドソーシングのワーカー数は 168 で, 自車から見える駐車シーンの画像とその画像上に示された駐車位置のマークを提示し, ランドマークとして, コンビニ, コンビニの入り口, コンビニに隣接する屋外トイレ, 駐車している白色と黒色の自動車, 歩行者 1 人 (2 台の自動車の間) を利用して, 2 台の車の右, それらの間にある 2 つの駐車スペースについて, それぞれ以下の条件の時における, 駐車指示を入力するように依頼した.

1. 自車と駐車位置との関係のみを利用
2. 各ランドマーク 1 つと駐車位置との関係を利用
3. ２つのランドマークと駐車位置との関係を利用（3 種類の異なる組合せを選択）

このようにして収集された駐車指示文には, 人手で CCG の構文木と SDC がアノテーション付けられており, 構文木の種類は全部で 381 種類存在する (表 2) . コーパスのアノテーションは一人で行い, もう一人が内容確認を行った. これらの構文木中で出現する意味カテゴリの数は, 45 となった (表 3). また, 駐車指示の文の長さの平均は 9.54, 分散は 3.09 となった. 駐車シーンの中の物体の数が少ないため, 形態素の語彙数は 348 と少なくなっているが, 空間的な意味を表す CCG Bank の語彙数は形態素とは必ずしも一致せず, それらの組み合わせから成り, 文脈に依存して幾つかの空間意味カテゴリを表しうるため, 全部で 1262 となった.

5.2 実験結果と考察

まず, コーパスの駐車指示を [Inago 17-2] のパーザを用いて生成した構文木について, SDC への変換を行う実験結果について述べる. 表 4 はパーザから出力された構文木から SDC への変換精度を表している. 「全ての木」は, パーザから出力された全ての構文木を SDC に変換した場合の精度であり, 「正

表 3: 駐車指示コーパスの構文木アノテーションに含まれる意味カテゴリ

カテゴリ
S, (S/S), (L\S)\O, (L/L)\S, (S\L), (L\S)\S, (L\L)\S, (S/S)\O, V, (L\S)\V, (L\V)\O, (V\O), (V\L), (L\V), (L\V)/L, (L\S), (L\S)/L, (S\L)\L, (O/O)\S, (E\S)\L, (O\S), (S\O), (E\S), (O\V), (L\V)\L, (O/O)\V, (L/L)\V, (E\L)/S, (E/S)\S, (S/E), (L\L)\O, (S\L)\O

解木」はパーザから出力された構文木の中で正解だった構文木のみを SDC に変換した場合の精度である．精度においては, SDC の構造が正解データと完全に一致していたら正しいとみなす．括弧内は実験データの数である．パーザから出力された構文木が正しい場合には, SDC に正しく変換できたことを示している．運転指示文から直接的に複雑な分岐構造を持つ SDC へ変換するのではなく, その中間表現として, シンプルな分岐から成る CCG の構文木を得ることの有効性を示していると考えられる．今後は駐車指示データの拡充し, より多様な表現への対応についても評価を行っていきたい．

表 4: 構文木から SDC への変換精度

全ての木	正解木
0.847(6019)	1.0(5098)

6. まとめと今後の課題

本研究では, 自動運転車に向けた駐車指示から空間意味記述を生成することを目的とし, 構文木から SDC への変換手法についての提案と, 実装した変換器の日本語の駐車指示のコーパスによる評価を行った．ヒューリスティックな知識に基づき定義した変換規則を用いることにより, 今回使用した実験データについては, パーザから出力された全ての構文木が正しい場合には, 正しい SDC に変換することができた．これは, 中間表現として CCG の構文木を得ることの有効性を示していると考えられる．高い精度で運転指示を SDC に変換できたのは, 固定された駐車シーンに含まれるランドマークの数が少なく, 表現が多様でなかったためだと考えられる．そこで今後は, 多様な駐車シーンでかつ様々なランドマークが存在する状況を含めて, より多くの駐車指示文に柔軟に対応するために, 言語データの拡充や未知語処理の実装を目指し, 駐車指示以外への運転指示への拡張を考えていきたい．さらにはラムダ式の導入により SDC に背景知識の組み込みを行う予定である．

参考文献

[Inago 17] 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」, 2B3-OS-07a-5in2, 愛知, 2017 年度 人工知能学会全国大会 (第 31 回), 2017.

[Inago 17-2] 稲子明里, 塚原裕史, 小林一郎,「自動運転のための駐車指示から空間意味記述表現への変換パーザの精度向上に関する取り組み」, P1-11, 岡山, 言語処理学会第 24 回年次大会, 2018.

[Kollar 10] T.Kollar, S.Tellex, D.Roy, N.Roy: Toward Understanding Natural Language Directions, ACM/IEEE Int ’l Conf. on Human-Robot Interaction (HRI), pp.259-266, 2010.

[Steedman 00] T.Kollar, M.Steedman: The Syntactic Process, MIT Press, 2000.

[Marneffe 06] M.Marneffe, B.MacCartney, C.D.Manning: Generating Typed Dependency Parses from Phrase Structure Parses, Proceedings of the International conference on Language Resources and Evaluation (LREC), 2006.

[Jackendoff 83] R.Jackendoff: Semantics and Cognition, MIT Press, 1983.

[Talmy 05] L.Talmy: The fundamental system of spatial schemas in language, Hamp, B., ed., From Perception to Meaning: Image Schemas in Cognitive Linguistics, 2005.

[Landau 93] B.Landau and R.Jackendoff: “What” and “where” in spatial language and spatial cognition, Behavioral and Brain Sciences 16, pp.217-265, 1993.

[Tellex 11] S.Tellex, T.Kollar, S.Dickerson, M.R.Walter, A.G.Banerjee, S.Teller, N.Roy: Understanding Natural Language Commands for Robotic Navigation and Mobile Manipulation, AAAI, 2011.

[Paul 16] Rohan Paul, Jacob Arkin, Nicholas Roy, and Thomas M. Howard: Efficient Grounding of Abstract Spatial Concepts for Natural Language Interaction with Robot Manipulators, RSS, 2016.

[Arkin 16] Jacob Arkin, Matthew R. Walter, Adrian Boteanu, Michael E. Napoli, Harel Biggie, Hadas Kress-Gazit, and Thomas M. Howard: Contextual Awareness: Understanding Monologic Natural Language Instructions for Autonomous Robots, IEEE, 2017.

[Guu 17] Kelvin Guu, Panupong Pasupat, Evan Liu, Percy Liang : From Language to Programs: Bridging Reinforcement Learning and Maximum Marginal Likelihood, ACL, 2017.