

CCG に基づく時間解析の実装に向けて

大西 舞子 (指導教員:戸次 大介)

1 はじめに

自然言語処理では、イベントの前後関係など時間情報が付与されたコーパス (Timebank [1]) の開発、及び、それに基づく時間情報解析が盛んに研究されている [2, 3]。一方、組合せ範疇文法 (CCG) [4] に基づく構文解析・意味解析の発展によって、論理式による意味表現や含意関係認識の研究が進んでいる。CCG に基づく意味解析・自動推論システム *ccg2lambda* [5] では、高階論理 (HOL) による意味表示を用いて論理推論を扱う含意関係認識で高い精度を達成している。しかし、現在のところ、*ccg2lambda* を含めて、CCG に基づく意味解析に時間情報を組み込む研究は十分に進展していない。

英語を対象とする *ccg2lambda* の解析では、現在や過去、未来、過去完了などの時制は十分に捉えられていない。例えば、表 1 の前提 1 と前提 2 から結論を導く推論は、人間にとっては容易であるが、現在のシステムでは含意関係の成立を予測することはできない。その原因は、動詞が表す出来事 (eventuality) の位置付けにある。特に、動詞 *met* や *lived* が表す event と state という出来事タイプの区別、並びに、*in* や *during* が表す時間関係を捉えることが必要となる。

表 1: event と state の関係性の推論の例

前提 1	Bob met Mary in 1993.
前提 2	Bob lived in Paris during 1993.
結論	Bob lived in Paris when he met Mary.

そこで本研究では、こうした時間推論と時間表現の意味解析を実現可能とするために、Kamp and Reyle [6] に基づく eventuality の分析と *ccg2lambda* で用いられていた意味表示を踏まえて、CCG の枠組みに時間に関する推論を扱える機構を組み込む手法を提案する。具体的には、CCG に基づく意味解析・含意関係認識の自動システムに、時制の区別を考慮した動詞の意味割り当てを行い、それに付随する型や述語の定義、推論に必要な公理を導入する。

本稿では、2 節で eventuality を紹介し、文全体の意味表示を用意する。3.1 節で eventuality が state の場合、3.2 節で event の場合を Kamp に基づき説明し、3.3 節で 3.1, 2 節の分析を踏まえて、推論を行える意味表示を提案する。3.4 節で、2 節と 3.3 節の意味表示から、表 1 の推論が導出可能であることを示す。

2 出来事の追加

Kamp and Reyle [6] によると、eventuality は (state/event) に区別され、様々な時間区間と比較することで位置付けられる。そこで時間区間を意味する **Time** 型を導入する。eventuality を **Event** 型を持つ項 e として、**Time** に写像する dur 関数を用いることで時間区間を $\text{dur}(e)$ と表す。発言時を **Time** 型の定項 now で表す。 e が生じる location time を **Time** 型を持つ項 t とする。参照時を **Time** 型を持つ項 r とする。 e が state の場合、 $\text{dur}(e) \circ t$ 、event の場合、 $\text{dur}(e) \subseteq t$ という関係を持つ。つまり、state の例である 1 では、 t は 1993 年にあたり、 $\text{dur}(e)$ は Bob が Paris に住んでいる期間にあたり、 t と $\text{dur}(e)$ が重複する期間が存在することを意味する。event の例である 2 では、 t は 1993 年にあたり、 $\text{dur}(e)$ は Bob が大統領に会う時点が 1993 年に含まれることを意味する。

- (1) Bob lived in Amsterdam in 1993.
- (2) Bob met the president in 1993.

ccg2lambda の意味表示に、eventuality を位置付ける条件を加える。その図式を以下に示す。

- (3) a. Bob had lived in Paris.
b. $\exists e \exists t \exists r. (\text{live}(e) \wedge \text{Subj}(e) = \text{bob} \wedge \text{in}(e, \text{Paris}) \wedge \boxed{\varphi_{\text{had}}(\text{now}, \text{dur}(e), t, r)} \wedge \boxed{\varphi_{\text{lived}}(\text{now}, \text{dur}(e), t, r)})$

ここで、 $\boxed{\varphi(n, \text{dur}(e), t, r)}$ は、束縛変数 e, t, r と定数 now で構成される論理式の表記とする。(3a) の意味表示を (3b) のように、ある eventuality が成立する時間区間は、 $\text{now}, \text{dur}(e), t, r$ 間の関係を表す条件の連言で表す。

3 時制の分析

Kamp and Reyle [6] に基づいて、英語の時制を表 2 に示す 8 種類に分類する。それぞれの時制は 4 つの素性で表現できる。素性の役割は表 3 に示す。

表 2: 時制と素性の関係

時制	参照位置	e の位置	e の性質	完了
present	-PAST	pres	+STAT	-PERF
future	-PAST	fut	+/-STAT	-PERF
simple past	-PAST	past	+/-STAT	-PERF
	+PAST	pres	+STAT	-PERF
past future	+PAST	fut	+/-STAT	-PERF
present perfect	-PAST	pres	+STAT	+PERF
future perfect	-PAST	fut	+STAT	+PERF
past perfect	+PAST	past	+/-STAT	-PERF
	-PAST	past	+STAT	+PERF
	+PAST	pres	+STAT	+PERF
past future perfect	+PAST	fut	+STAT	+PERF

表 3: 素性が与える意味表示

素性	意味表示	素性	意味表示
+PAST	$r < \text{now}$	-PAST	$\text{now} = r$
pres	$t = r$	past	$t < r$
fut	$r < t$		
+STAT	$\text{dur}(e) \circ t$	-STAT	$\text{dur}(e) \subseteq t$
+PERF	$\exists e'. \exists t'. \text{end}(\text{dur}(e)) \supseteq \text{dur}(e') \wedge \text{dur}(e') \circ t'$	-PERF	—

以下では、それぞれの時制に対する文を構成的に導出する方法についてその概略を述べる。

3.1 eventuality が state の例

表 2 で分類した時制に対する単純な例を (4)–(11) に示す。

- (4) Bob lives in Paris. present
- (5) Bob will live in Paris. future
- (6) Bob lived₁ in Paris. simple past
- (7) Bob would live in Paris. past future
- (8) Bob has lived₂ in Paris. present perfect
- (9) Bob will have lived₂ in Paris. future perfect
- (10) Bob had lived₂ in Paris. past perfect
- (11) Bob would have lived₂ in Paris. past future perfect

これらの文の意味表示を導出するのに必要な時制に関わる CCG の語彙項目を表 4 に示し、素性を割り振る。統語範疇は、C&C パーザ [7] など現在の CCG パーザの出力に基づくものである。lived₁ と lived₂ の区別は、動詞の過去形と過去分詞形に対応する。統語範疇に含まれる *base* は、その語彙が原形であることを表し、*pp* は、過去分詞形であることを表す。また、VB、VBP、VBD は、CCG パーザがサポートする POS タグであり、表中の「—」は、その語彙では関係を決定できないことを指す。

(8) の意味表示の構成を例に挙げると、2 つの $\boxed{\varphi(\text{now}, \text{dur}(e), t, r)}$ に入る意味表示が、*has* と lived₂ の意味表示であり、それらは表 4 の項目 6, 4 が満たす素性の意味表示を連言で結合したものである。

表 4: CCG と Kamp の素性の対応

項目	語彙項目	統語範疇	参照位置	e の位置	e の性質	完了
1	live	VB, $S_{base} \setminus NP$	—	—	+STAT	-PERF
2	lives	VBP $S \setminus NP$	-PAST	pres	+STAT	-PERF
3	lived ₁	VBD $S \setminus NP$	-PAST	past	+STAT	-PERF
			+PAST	pres	+STAT	-PERF
4	lived ₂	$S_{pp} \setminus NP$	—	—	+STAT	-PERF
5	have	$(S_{base} \setminus NP) / (S_{pp} \setminus NP)$	—	—	—	+PERF
6	has	$(S \setminus NP) / (S_{pp} \setminus NP)$	-PAST	pres	—	+PERF
7	had	$(S \setminus NP) / (S_{pp} \setminus NP)$	+PAST	past	—	-PERF
			-PAST	past	—	+PERF
			+PAST	pres	—	+PERF
8	will	$(S \setminus NP) / (S_{base} \setminus NP)$	-PAST	fut	—	-PERF
9	would	$(S \setminus NP) / (S_{base} \setminus NP)$	+PAST	fut	—	-PERF

3.2 eventuality が event の例

表 2 の -STAT に対する時制の例 (12)–(15) に示す。

- (12) Bob will meet Mary. future
 (13) Bob met₁ Mary. simple past
 (14) Bob would meet Mary. past future
 (15) Bob had met₂ Mary. past perfect

state の場合と同様の手順を踏み、表 5 の対応を取る。will, would の意味表示には state/event の区別がなく、表 4 のものと同一である。

表 5: CCG と Kamp の素性の対応

項目	語彙項目	統語範疇	参照位置	e の位置	e の性質	完了
1	meet	VB, $S_{base} \setminus NP / NP$	—	—	-STAT	-PERF
2	met ₁	VBD $S \setminus NP / NP$	-PAST	past	-STAT	-PERF
3	met ₂	$S_{pp} \setminus NP / NP$	—	—	-STAT	-PERF
4	had	$(S \setminus NP) / (S_{pp} \setminus NP)$	+PAST	past	—	-PERF

3.3 state/event の意味表示

前提に存在する時間区間同士の関係性の推論を行うために、Allen [8] の区間算術の合成テーブルに示される多くの性質を少ない定義と公理で示すことを提案する。Kamp による意味表示と Allen の記述形式との対応を表 6 に示す。cgg2lambda の記述形式との対応を表 7 に示す。

表 6: Kamp と Allen の対応

Kamp	Allen
$i = i'$	$i = i'$
$i \circ i'$	$\neg(i < i' \vee i > i')$
$i < i'$	$i < i'$
$i \subseteq i'$	$i d i' \vee i = i'$
$end(i)$	i_{end}
$begin(i)$	i_{start}
$end(i) \supseteq i'$	$i m i'$

表 7: Allen と cgg2lambda の対応

Allen	cgg2lambda	Allen	cgg2lambda
=	$equal(t, t')$	>	$before(t', t)$
<	$before(t, t')$	>	$before(t', t)$
m	$meet(t, t')$	mi	$meet(t', t)$
o	$overlap(t, t')$	oi	$overlap(t', t)$
s	$start(t, t')$	si	$start(t', t)$
f	$finish(t, t')$	fi	$finish(t', t)$
d	$during(t, t')$	di	$during(t', t)$
i_{end}	$end(t)$	i_{start}	$begin(t)$

新たに導入した 2 項述語の定義を表 8 に示す。

表 8: 時間区間を引数に取る 2 項述語の定義

$equal(i, i') \stackrel{def}{=} equal(begin(i), begin(i')) \wedge equal(end(i), end(i'))$
$before(i, i') \stackrel{def}{=} before(end(i), begin(i'))$
$meet(i, i') \stackrel{def}{=} equal(end(i), begin(i'))$
$overlap(i, i') \stackrel{def}{=} before(begin(i), begin(i')) \wedge before(begin(i'), end(i)) \wedge before(end(i), end(i'))$
$start(i, i') \stackrel{def}{=} equal(begin(i), begin(i')) \wedge before(end(i), end(i'))$
$finish(i, i') \stackrel{def}{=} before(begin(i'), begin(i)) \wedge equal(end(i), end(i'))$
$during(i, i') \stackrel{def}{=} before(begin(i'), begin(i)) \wedge before(end(i), end(i'))$

表 9 にこの表記を用いた推論のための公理を示す。

表 9: 時間区間に関する公理

$i, begin(i), end(i) : \mathbf{Time}$
$\psi(i, i') : \mathbf{Time} \rightarrow \mathbf{Time} \rightarrow \mathbf{Prop}$
$before(begin(i), end(i))$
$before(i, i') \vee equal(i, i') \vee before(i', i)$
$\neg before(i, i)$
$before(i, i') \wedge before(i', i'') \rightarrow before(i, i'')$
$equal(i, i)$
$equal(i, i') \rightarrow equal(i', i)$
$equal(i, i') \wedge equal(i', i'') \rightarrow equal(i, i'')$
$equal(i, i') \wedge \psi \rightarrow \psi[i'/i]$

以上を用いて、state は表 10、event は表 11 のように意味表示を変換される。

表 10: state の意味表示

live	$\neg(before(dur(e), t) \vee before(t, dur(e)))$
lives	$equal(now, r) \wedge equal(t, r) \wedge \neg(before(dur(e), t) \vee before(t, dur(e)))$
lived ₁	$(equal(now, r) \wedge before(t, r) \wedge \neg(before(dur(e), t) \vee before(t, dur(e)))) \vee (before(r, now) \wedge equal(t, r) \wedge \neg(before(dur(e), t) \vee before(t, dur(e))))$
lived ₂	$\neg(before(dur(e), t) \vee before(t, dur(e)))$
have	$\exists e'. \exists t'. meet(dur(e), dur(e')) \wedge \neg(before(dur(e'), t') \vee before(t', dur(e')))$
has	$\exists e'. \exists t'. \exists r'. equal(now, r') \wedge equal(t', r') \wedge meet(dur(e), dur(e')) \wedge \neg(before(dur(e'), t') \vee before(t', dur(e')))$
had	$(before(r, now) \wedge before(t, r)) \vee (\exists e'. \exists t'. \exists r'. equal(now, r') \wedge before(t', r') \wedge meet(dur(e), dur(e')) \wedge \neg(before(dur(e'), t') \vee before(t', dur(e')))) \vee (\exists e'. \exists t'. \exists r'. before(r', now) \wedge equal(t', r') \wedge meet(dur(e), dur(e')) \wedge \neg(before(dur(e'), t') \vee before(t', dur(e'))))$
will	$equal(now, r) \wedge before(r, t)$
would	$before(r, now) \wedge before(r, t)$

表 11: event の意味表示

meet	$during(dur(e), t) \vee equal(dur(e), t)$
met	$equal(now, r) \wedge before(t, r) \wedge (during(dur(e), t) \vee equal(dur(e), t))$
met	$during(dur(e), t) \vee equal(dur(e), t)$
had	$before(r, now) \wedge before(t, r)$

3.4 state/event 間の関係の推論

以上の定義を用いて、異なる時制をもつ文間に成り立つ推論を検討した。state に関する (4)–(11) の異なる 2 文を前提と結論をとすると、過去完了から過去のみ可能である。event に関する (12)–(15) も同様である。event と state を組み合わせた推論の例として、表 1 の前提 1 と前提 2 から、結論を導くことができる。3 節の分析に基づいて、前提 1、前提 2、結論の意味表示として以下のものが導出可能であり、前提から結論への含意関係が成立することが証明可能である。

前提 1 $\exists e \exists t \exists r. (meet(e, bob, mary) \wedge (now = r) \wedge t < r \wedge dur(e) \subseteq t \wedge t \subseteq 1993)$

前提 2 $\exists e \exists t \exists r. (live(e, bob) \wedge in(e, paris) \wedge (now = r) \wedge t < r \wedge dur(e) \circ t \wedge 1993 \subseteq t \wedge t \subseteq dur(e))$

結論 $\exists e \exists t \exists r. (live(e, bob) \wedge in(e, paris) \wedge (now = r) \wedge t < r \wedge dur(e) \circ t \wedge \exists e' \exists t' \exists r'. (meet(e', bob, mary) \wedge (now = r) \wedge t' < r \wedge dur(e') \subseteq t' \wedge t' \subseteq t \wedge t \subseteq dur(e)))$

4 今後の課題

本稿では、eventuality の分類として event/state を採用したが、States/Activities/Accomplishments/Achievements に分類する方法もある。さらに、eventuality の決定に関して、接続詞・前置詞・主語・目的語が影響を与えるため、文レベル以下では決定されないものがある。今後の分析の試みとして、時間副詞の意味表示の決定と、cgg2lambda での実装を進めていく予定である。

参考文献

- [1] James Pustejovsky, Patrick Hanks, Roser Sauri, Andrew See, Robert Gaizauskas, Andrea Setzer, Dragomir Radev, Beth Sundheim, David Day, Lisa Ferro, et al. The TimeBank corpus. In *Corpus linguistics*, p. 40. Lancaster, UK., 2003.
- [2] Fei Cheng and Yusuke Miyao. Classifying temporal relations by bidirectional lstm over dependency paths. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Vol. 2, pp. 1–6, 2017.
- [3] Tomohiro Sakaguchi, Daisuke Kawahara, and Sadao Kurohashi. comprehensive annotation of various types of temporal information on the time axis.
- [4] Mark J. Steedman. *The Syntactic Process*. The MIT Press, 2000.
- [5] Pascual Martínez-Gómez, Koji Mineshima, Yusuke Miyao, and Daisuke Bekki. cgg2lambda: A Compositional Semantics System. In *Proceedings of ACL 2016 System Demonstrations*, pp. 85–90, 2016.
- [6] Hans Kamp and Uwe Reyle. *From Discourse to Logic: Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Springer, 1993.
- [7] Stephen Clark and James R. Curran. Wide-coverage efficient statistical parsing with CCG and log-linear models. *Computational Linguistics*, Vol. 33, pp. 493–552, 2007.
- [8] James F. Allen. Maintaining Knowledge about Temporal Intervals. *Commun. ACM*, Vol. 26, pp. 832–843, 1983.