

CCGに基づく時間解析の実装に向けて

大西 舞子^{1,a} 峯島 宏次^{1,b} 戸次 大介^{1,c}

お茶の水女子大学¹

{g1520507^a, bekki^c}@is.ocha.ac.jp, minesima.koji@ocha.ac.jp^b

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 背景

Kamp and Reyle [6] によると、eventuality は、様々な時間区間と比較することで位置付けられる。またその関係は、eventuality の種類 (state/event) によって変化する。[6] では、動詞や助動詞の持つテンスとアスペクトの性質が、eventuality を state/event に分類した場合の決定に与える影響を整理している。

本稿では [6] の見解を採用し、例文としてテンスとアスペクトのみで eventuality が決定するものを用いる。[6] によると、文は瞬間 (instants) ではなく、時間区間 (intervals) と捉える。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 として導入すると、**Time** に写像する dur 関数を用いることで時間区間が $\text{dur}(e)$ と表せる。現在の時間を **Time** 型の定項 now で表す。文中で eventuality を捉える location time という **Time** 型を持つ項 t を導入する。文中の eventuality が参照する reference time という **Time** 型を持つ項 r を導入する。これらを用いて、Kamp は e の state の場合と event の場合の t との関係を表す意味表示を $\text{dur}(e) \circ t$ および $\text{dur}(e) \subseteq t$ と表記した。event は、eventuality の状態が連続に変化し、いずれ達成する境界があり、それを越えると達成したという状態が続く。state は、そのような境界が存在せず。一定の状態が続く。

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

例えば、(1) は、Bob が Amsterdam に住んでいる期間が、1993年の期間と重複している箇所がある場合に真である。(2) は、Bob が大統領に会うまでに家を出発し、電車を乗り継ぎ徐々に大統領に会う時点に接近している期間があり、ようやく大統領に出会う時点に到達する時が訪れる。この一連の出来事が1993年の間に存在している時に真である。

ccg2lambda の意味表示は、eventuality を位置付ける条件に $\text{now}, t, r, \text{dur}$ が存在しないため、 t, r の束縛変数を加える。 $\varphi(n, \text{dur}(e), t, r)$ 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) と表すと、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ことにする。それを可能とする語彙項目の意味表示は用意している。

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

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。+/-PAST は、参照時の現在時に対する位置づけを表し、+ は参照時が現在より過去、- は参照時が現在時である。pres/fut/past は、eventuality の location time と参照時の関係を表す。pres は「参照時と等しい」、fut は「参照時より未来にある」、past は「参照時より過去にある」という関係をそれぞれ表す。+/-STAT は、+STAT は state を、-STAT は event を表す。+/-PERF は、+PERF が完了形、-PERF は非完了形を表す。

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

素性と意味表示との対応は表 3 が示すように、 $\text{now}, \text{dur}(e), t, r$ の関係として表現することができる [6]。

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	$\exists t'. t' < \text{now} \wedge r \subseteq t'$	-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	$\exists e'. \text{end}(\text{dur}(e)) \supset \text{dur}(e')$	-PERF	—

以下では、state 文と event 文の意味表示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 のそれぞれの時制に対する例を (4)–(11) に示す。

- (4) Bob lives in Amsterdam. **present**
 (5) Bob will live in Amsterdam. **future**
 (6) Bob lived₁ in Amsterdam. **simple past**
 (7) Bob would live in Amsterdam. **past future**
 (8) Bob have lived₂ in Amsterdam. **present perfect**

- (9) Bob will have lived₂ in Amsterdam. **future perfect**
 (10) Bob had lived₂ in Amsterdam. **past perfect**
 (11) Bob would have lived₂ in Amsterdam. **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 に示す。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。表 4 では、CCG の語彙項目と表 3 の素性との対応が与えられている。これらの素性の意味が文中の時制に関わる語彙、つまり動詞や助動詞の意味に割り振られる。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、それぞれ参照時、e の位置、e の性質が決定されない語彙であることを指す。

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, $S_{\text{base}} \backslash NP$	—	—	+STAT	-PERF
2	lives	VBP $S \backslash NP$	-PAST	pres	+STAT	-PERF
3	lived ₁	VBD $S \backslash NP$	-PAST	past	+STAT	-PERF
			+PAST	pres	+STAT	-PERF
4	lived ₂	$S_{\text{pp}} \backslash NP$	—	—	+STAT	-PERF
5	have	$(S_{\text{base}} \backslash NP) / (S_{\text{pp}} \backslash NP)$	—	—	—	+PERF
6	has	$(S \backslash NP) / (S_{\text{pp}} \backslash NP)$	-PAST	pres	—	+PERF
7	had	$(S \backslash NP) / (S_{\text{pp}} \backslash NP)$	+PAST	past	—	-PERF
			-PAST	past	—	+PERF
			+PAST	pres	—	+PERF
8	will	$(S \backslash NP) / (S_{\text{base}} \backslash NP)$	-PAST	fut	—	-PERF
9	would	$(S \backslash NP) / (S_{\text{base}} \backslash NP)$	+PAST	fut	—	-PERF

(5) の意味表示の構成を例に挙げると、時制に関連する語彙が will と live であり、 $\boxed{\varphi(\text{now}, \text{dur}(e), t, r)}$ に入る意味表示は、それぞれの意味表示の連言である。will の意味表示は、表 4 の項目 8 の -PAST と fut の意味表示の連言で結合したものである。-PAST の意味表示は、表 3 に示す。

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 the president. **future**
 (13) Bob met₁ the president. **simple past**
 (14) Bob would meet the president. **past future**
 (15) Bob had met₂ the president. **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, $S_{base} \setminus NP$	—	—	—STAT	—PERF
2	met ₁	VBP $S \setminus NP$	—PAST	past	—STAT	—PERF
3	met ₂	VBD $S \setminus NP$	—	—	—STAT	—PERF
4	had	$(S \setminus NP) / (S_{pp} \setminus NP)$	+PAST	past	—	—PERF

3.3 state/event の意味表示

時間区間同士の関係性の推論は、前提に存在する時間区間同士であれば、Allen [8] の区間算法の合成テーブル [8] を満たす時間区間に関する公理や定義より示せる。そこで、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。

表 6: Kamp [6] と Allen [8] の対応

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

表 7 は、cgg2lambda のテンプレートの記述形式との対応を示している。

表 7: Allen [8] と cgg2lambda の対応

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

新たに導入した 2 項述語を定義を表 8 に示す。

表 8: 時間区間を引数にする 2 項述語の定義

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

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

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

以上を用いて、state の意味表示を表 10 のように、event の意味表示を表 11 のように変換し、推論を行うことが可能となる。

表 10: state の意味表示

live	$\neg(\text{before}(\text{dur}(e), t) \vee \text{before}(t, \text{dur}(e)))$
lives	$\text{equal}(\text{now}, r) \wedge \text{equal}(t, r) \wedge \neg(\text{before}(\text{dur}(e), t) \vee \text{before}(t, \text{dur}(e)))$
lived₁	$(\text{equal}(\text{now} = r) \wedge \text{before}(t, r) \wedge \neg(\text{before}(\text{dur}(e), t) \vee \text{before}(t, \text{dur}(e)))) \vee (\exists t'. \text{before}(t', \text{now}) \wedge (\text{during}(r, t') \wedge \text{equal}(t, r) \wedge \neg(\text{before}(\text{dur}(e), t) \vee \text{before}(t, \text{dur}(e))))$
lived₂	$\neg(\text{before}(\text{dur}(e), t) \vee \text{before}(t, \text{dur}(e)))$
have	$\exists e'. \text{meet}(\text{dur}(e), \text{dur}(e'))$
has	$\exists e'. \text{equal}(\text{now}, r) \wedge \text{equal}(t, r) \wedge \text{meet}(\text{dur}(e), \text{dur}(e'))$
had	$(\exists t'. \text{before}(t', \text{now}) \wedge \text{during}(r, t') \wedge \text{before}(t, \text{now})) \vee (\exists e'. \text{equal}(\text{now}, r) \wedge \text{before}(t, r) \wedge \text{meet}(\text{dur}(e), \text{dur}(e'))) \vee (\exists t'. \exists e'. \text{before}(t', \text{now}) \wedge \text{during}(r, t') \wedge \text{equal}(t, r) \wedge \text{meet}(\text{dur}(e), \text{dur}(e'))$
will	$\text{equal}(\text{now}, r) \wedge \text{before}(r, t)$
would	$\exists t'. \text{before}(t', \text{now}) \wedge \text{during}(r, t') \wedge \text{before}(r, t)$

表 11: event の意味表示

meet	$\text{during}(\text{dur}(e), t) \vee \text{equal}(\text{dur}(e), t)$
met	$\text{equal}(\text{now}, r) \wedge \text{before}(t, r) \wedge (\text{during}(\text{dur}(e), t) \vee \text{equal}(\text{dur}(e), t))$
met	$\text{during}(\text{dur}(e), t) \vee \text{equal}(\text{dur}(e), t)$
had	$\exists t'. \text{before}(t', \text{now}) \wedge \text{during}(r, t') \wedge \text{before}(t, r)$

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

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

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

4 今後の課題

本稿では、eventuality の分類として event/state という単純な分類を採用したが、Aktionsarten [9] としては、States、Activities、Accomplishments、Achievements のさらなる細かい区分が存在する。さらに、eventuality の決定に関して、接続詞・前置詞句・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、時間副詞の中でも本稿で触れた *when* のみではなく、*before* や *after* などの接続詞、*in* や *at*, *for* などの前置詞、主語と目的語を行い、ccg2lambda での実装を進めていく予定である。

謝辞 本研究は、JST AIP-PRISM JPMJCR18Y1 の支援を受けたものである。

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. ccg2lambda: 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.

[9] Zeno Vendler. *Linguistics in Philosophy*. Cornell University Press, 1967.