

依存型意味論による日本語のテンス・アスペクトの分析に向けて

Towards an Analysis of Tense and Aspect in Japanese by Dependent Type Semantics

宇津木 舞香 *1 戸次 大介 *1*2*3
Maika Utsugi Daisuke Bekki

*1 お茶の水女子大学大学院人間文化創成科学研究科 *2 国立情報学研究所
Ochanomizu University, Graduate School of Humanities and Sciences National Institute of Informatics

*3 独立行政法人科学技術振興機構, CREST
CREST, Japan Science and Technology Agency

Natural language sentences include time information. For that reason, an analysis of tense and aspect is essential to grasp the meaning of sentences in natural language. In this study, we propose an analysis of tense and aspect in Japanese language by means of Dependent Type Semantics (DTS). DTS is a framework for natural language semantics based on dependent type theory. In this work, we extend DTS by introducing the concept of time. To that purpose, we give to each lexical item concerning tense and aspect a semantic representation in DTS, and show the calculation of semantic composition and entailment relation.

1. はじめに

形式意味論による自然言語の研究はさかんに行われているが、テンス・アスペクトについてはほぼ全ての文に現れるにも関わらず、これらがもたらす言語現象を網羅するような理論は確立されていない。

テンスは文中の出来事の時刻を参照点を基準として位置づける文法形式を指し、アスペクトは出来事の中の様々な局面を表す文法形式を指す。

(1) 太郎は毎朝 7 時にレッドブルを飲んでいる。

文 (1) では、アスペクトに関わる動詞性接尾語 [テイル] が接続することで「太郎は毎朝 7 時にレッドブルを飲む」という出来事が今まさに進行しているという局面を表している。また、量化表現 [毎朝]、時間副詞 [7 時]、非過去を表す終止形活用語尾 [ル] が接続することでその出来事の時間的な位置を指定している。

本研究では、依存型理論に基づく自然言語の証明論の意味論のひとつである依存型意味論を用いて日本語のテンス・アスペクトの分析を行うために、依存型意味論に時間的な概念を導入した。本論文では、テンス・アスペクトに関わる語彙項目に意味表示を与え、意味合成および演繹関係の計算を示す。

2. 依存型意味論

依存型意味論 (Dependent Type Semantics)[1] は、依存型理論 [3] に基づいた証明論の意味論である。

証明論の意味論であることから、自然言語の意味について証明論的な推論が可能となる。依存型意味論では文 A から文 B が推論可能であるという関係を、文 A の意味表示から文 B の意味表示への演繹可能性として捉える。また、合成的意味論でもあることから、全体の意味は部分の意味から計算するという特徴をもつ。つまり、文全体の意味は、文中に含まれる各語彙項目の意味表示を合成することで構築される。

2.1 依存型理論

依存型理論は項に依存した型を記述することが可能な体系である。重要な構成子として Π と Σ をもつ。 Π は「A であるようなすべての x について $B(x)$ が成り立つ」ことを意味し、自然言語の全称量化に対応する。 Σ は「 $B(x)$ であり、かつ A であるような x が存在する」ことを意味し、自然言語の存在量化に対応する。それぞれ、implication(\rightarrow) と conjunction(\wedge) の一般化として定義されている。

$$(x:A) \rightarrow B(x) \qquad \left[\begin{array}{c} x:A \\ B(x) \end{array} \right]$$

Π 型 Σ 型

Π 、 Σ の導入規則と除去規則を以下に示す。ここで、 Σ の除去規則に現れる π_1 、 π_2 は射影であり、それぞれペアの第一要素、第二要素を取り出す操作として定義されている。

$\frac{(x:A) \rightarrow B : \text{type} \quad \frac{x:A}{\vdots} (i)}{\lambda x.M : (x:A) \rightarrow B} (i)$	$\frac{M : (x:A) \rightarrow B \quad N : A}{MN : B[N/x]} (\Pi E)$
$\frac{M : A \quad N : B[M/x]}{(M, N) : \left[\begin{array}{c} x:A \\ B(x) \end{array} \right]} (\Sigma I)$	$\frac{y : \left[\begin{array}{c} x:A \\ B(x) \end{array} \right]}{\pi_1(y) : A} (\Sigma E)$
	$\frac{y : \left[\begin{array}{c} x:A \\ B(x) \end{array} \right]}{\pi_2(y) : B[\pi_1(y)/x]} (\Sigma E)$

2.2 依存型意味論による意味表示

依存型意味論は、自然言語の意味表示に依存型理論の型を対応づけているため、先行する文脈や句を考慮した意味表示を記述する事が可能となる。たとえば、“Every man raged.” の意味表示は以下のように表される。

$$\lambda c. \left(u : \left[\begin{array}{c} x:\text{entity} \\ \text{man}(x) \end{array} \right] \right) \rightarrow \text{raged}(\pi_1(u))$$

文全体の意味表示は Π 型である。これは “every” が全称量化として扱われているからであり、各語彙項目の意味表示と合成されることで文全体が Π 型となる。ここで u は、型 entity である項 x と、 x に依存した型 $\text{man}(x)$ からなる Σ 型をもつ。 $\text{raged}(\pi_1(u))$ の $\pi_1(u)$ は u の第一要素である entity を指す。また、 c は local context と呼ばれ、先行する文脈がもつ情報を指す。 c を引数として受け取ることで、先行する文脈の情報を全体の意味表示に組み込むことが可能となる。

連絡先: 宇津木舞香, 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戸次研究室, 東京都文京区大塚 2-1-1, utsugi.maika@is.ocha.ac.jp

3. 組合せ範疇文法

組合せ範疇文法 (Combinatory Categorical Grammar, CCG)[4] は語彙化文法のひとつである。語の統語・意味に関する情報が記述された辞書と、規則から構成される。CCGにおいて、統語範疇は基底範疇 $\{N, NP, S, \bar{S}, CONJ\}$ と “/” “\” から以下のように再帰的に定義される。

1. A が基底範疇ならば、 A は統語範疇である。
2. (i) A, B が統語範疇ならば、 A/B は統語範疇である。
(ii) A, B が統語範疇ならば、 $A \setminus B$ は統語範疇である。

ここで統語範疇 A/B をもつ語句は、自身の右側に統語範疇 B をもつ語句が接続した場合に統語範疇 A の語句となる。一方、統語範疇 $A \setminus B$ をもつ語句は、自身の左側に統語範疇 B をもつ語句が接続した場合に統語範疇 A の語句となる。(2i, ii)に現れる “/” “\” の性質は関数適用規則によって定義される。以下に CCG の規則の一部を示す。

順/逆関数適用規則 $\begin{array}{l} > \frac{X/Y : f \quad Y : a}{X : fa} \quad < \frac{Y : a \quad X \setminus Y : f}{X : fa} \end{array}$
等位接続規則 $\langle \Phi \rangle \frac{X : f_1 \dots \quad CONJ : \circ \quad X : f_m}{X : \lambda \vec{x}. (f_1 \vec{x}) \circ \dots \circ (f_m \vec{x})} \quad \text{但し, } 1 < m$
順/逆関数合成規則 $\begin{array}{l} >_B \frac{X/Y : f \quad Y/Z : g}{X/Z : \lambda x. f(gx)} \quad <_B \frac{Y/Z : g \quad X \setminus Y : f}{X \setminus Z : \lambda x. f(gx)} \end{array}$

CCG では各語彙項目は、音韻表示・統語範疇・意味表示の3つ組で表され、戸次 (2010)[9] では意味表示に動的意味論を用いて日本語の分析を行っている。

4. 依存型意味論への時間的概念の導入

$$\text{time} \stackrel{\text{def}}{=} \mathbb{Q} \times \mathbb{Q}$$

本研究では依存型意味論に時間概念を導入するため、新たに time 型を定義した。time 型は有理数の直積であるとし、time 型の項は、開始時刻 pivot と区間長 length からなる2つ組であるとする。従って、投射 π_1 と π_2 を用いる事で、 $t : \text{time}$ について t の開始時刻と区間長を取り出す事ができる。つまり、 $\pi_1(t) \equiv \text{pivot}(t)$ 、 $\pi_2(t) \equiv \text{length}(t)$ となる。time 型に関する関数の一部を表 1 に示す。

$\text{dur} : \text{eventuality} \rightarrow \text{time}$	eventuality 型をもつ項を受け取ったらその項の時間を返す
$\text{end}(t) \equiv \text{pivot}(t) + \text{length}(t)$	t の終了時点
$t < t' \equiv \text{end}(t) < \text{pivot}(t')$	時間 t が時間 t' よりも過去に位置する
$t \sqsubseteq t' \equiv \left[\begin{array}{l} \text{pivot}(t') < \text{pivot}(t) \\ \text{end}(t) < \text{end}(t') \end{array} \right]$	時間 t が時間 t' に含まれる

表 1: 時間に関する関数

5. 分析

依存型意味論に time 型を導入することで、これまで表現することの出来なかった時間情報を意味表示に記述することが可能となる。本研究では、CCG の意味表示部分に time 型によって拡張した依存型意味論を用いて日本語のテンスについての分析を行う。

冒頭でテンスは出来事の時間的位置を指定し、アスペクトは出来事が含むさまざまな局面を表す文法形式であると述べた。本研究では Kamp ら (1993)[2] に従い、出来事 (eventuality) は動作を表す event か状態を表す state のどちらかであるとする。以下では、テンス・アスペクトに関わる表現のうち、助動詞 [タ] 接尾語 [テイル] に焦点をあて、依存型意味論による意味表示の記述について述べる。また、時間副詞 [今日][毎日] を例に演繹関係の計算を示す。

5.1 出来事 (eventuality)

一般に文中の出来事 (eventuality) は述語によって表される。本研究では宇津木ら (2015)[7] に従い述語を以下のように動態述語と状態述語の2種類に分類し、それぞれ動態述語は event を、状態述語は state を表す。

eventuality	品詞分類		例
event	動態述語	動態動詞	ノム, アルク …
		連用形複合動詞	ナグリタオス …
state	状態述語	状態動詞	アル, イル …
		連用形複合動詞	デキスギル …
		形容詞	アカイ, ウツクシイ …
		状詞	クレイダ, シズカダ …

表 2: eventuality 分類表

動態述語と状態述語は意味表示の形式が異なる。(2) は動態動詞 [飲 (む)]、状詞 *¹[静か (だ)] の意味表示である。[飲 (む)] は event、[静か (だ)] は state を表す。動態述語では event の時間と参照点がちょうど重なっているのに対し、状態述語では参照点の上に state がまたがるように存在する。

$$\begin{array}{l} (2) \quad \text{a. 飲 (む)} \vdash S_e \setminus NP_{ga} \setminus NP_o : \lambda y. \lambda x. \lambda t. \lambda c. \left[\begin{array}{l} e: \text{drink}(x, y) \\ t \sqsubseteq \text{dur}^*(e) \end{array} \right] \\ \quad \text{b. 静か (だ)} \vdash S_s \setminus NP_{ga} : \lambda x. \lambda t. \lambda c. \left[\begin{array}{l} s: \text{calm}(x) \\ t \sqsubseteq \text{dur}^*(s) \end{array} \right] \end{array}$$

5.2 テンス [タ]

日本語では、テンスに関わる言語表現として 助動詞 [タ]、終止形活用語尾 [ル]*² が挙げられる。一般に、[タ][ル] はそれぞれ出来事を過去、非過去に位置づけるはたらきをもつ表現として知られている。本論文では、まず助動詞 [タ] に注目する。

(3)(4) に現れる助動詞 [タ] は、動態動詞 [食べる] が表す event と状詞 [サンマだ] が表す state を過去に位置づけている。

(3) 今日は7時に朝ご飯を食べた。

(4) 昨日の晩ご飯はサンマだった。

しかし、助動詞 [タ] は出来事を過去に位置づけるだけに留まらず、様々な意味を持つことが日本語学では知られている。

(5) (野球の実況中継) イチロー走った!

(6) さあ、行った行った!

(5)(6) は出来事を過去に位置づけるはたらきを持たない [タ] の例である。(5) は [走る] という event が発話時と同時に開始されることを意味し、(6) は命令のニュアンスをもたらし。つまり、統語上では同じ振り舞いをする助動詞 [タ] であるが、その用法によってそれぞれ異なる意味表示を与える必要がある。

表 6 は助動詞 [タ] の用法分類である。表 6 には7つのカテゴリが設けられているが、これは助動詞 [タ] には少なくとも7種類の語彙項目が必要であることを意味している。以下、このう

*1 戸次 (2010) では、状詞の概念は、形容動詞と、名詞+判定詞、副詞、連体詞にまたがる分類であるとしている。

*2 終止形活用語尾の表層形は必ずしも [ル] ではないが、ここでは総称として [ル] と呼ぶ事にする。

ち [タ-1][タ-2] の意味表示について説明する。[タ-1] は一般に知られている [出来事を過去に位置づけるタ] であり、[タ-2] は [動作開始時に使用されるタ] である。(3)(4) の [タ] は [タ-1]、(5) の [タ] は [タ-2] である。(7) は [タ-1][タ-2] の語彙項目であり、図 3, 4 はそれぞれ [タ-1][タ-2] を用いた文の意味合成である。

$$(7) \quad \text{a. } \text{タ-1} \vdash S \setminus S : \lambda p. \lambda t. \lambda c. \left[\begin{array}{l} ptc \\ t \prec st \end{array} \right]$$

$$\text{b. } \text{タ-2} \vdash S \setminus S : \lambda p. \lambda t. \lambda c. \left[\begin{array}{l} ptc \\ \text{pivot}(t) = \text{pivot}(st) \end{array} \right]$$

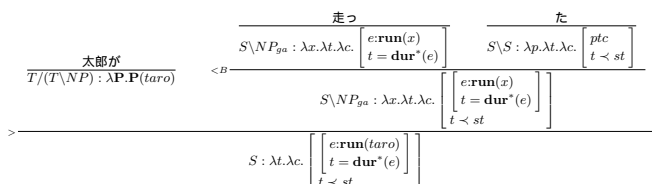


図 3: [タ-1] を用いた文の意味合成

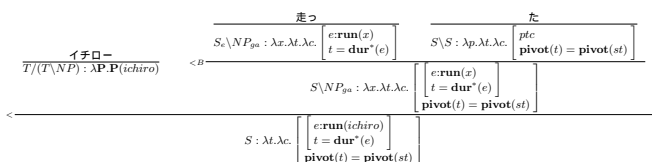


図 4: [タ-2] を用いた文の意味合成

文全体の意味表示には [タ-1][タ-2] の意味が反映され、図 3 では $e:\text{run}(x)$ という event が過去に位置することが記述され、図 4 では $e:\text{run}(x)$ という event が発話時に開始されることが記述されている。助動詞 [タ] に複数の語彙項目を設ける事で、助動詞 [タ] がもたらす言語現象をより正確に表現することが可能となった。

5.3 アスペクト [テイル]

アスペクトに関する語の例として動詞性接尾語 [テイル] について解説する。助動詞 [タ] と同様、[テイル] にも複数の用法が存在することが知られている。吉本 (1993)[5] では [テイル] には、進行、結果、経験、習慣の 4 つの用法があるとされている。(8) は [テイル (進行)] の意味表示であり、図 5 は (8) を用いた「走っている」の意味合成である。述語に [テイル (進行)] が後接する事によって、参照点に $e:\text{run}(x)$ という event がまたがって存在するという意味になり、(2a) に示した状態述語での state の位置づけによく似た形式の意味表示となる。

$$(8) \quad \text{テイル (進行)} \vdash S \setminus S : \lambda p. \lambda t. \lambda c. \left[\begin{array}{l} t' : \text{time} \\ t \sqsubseteq t' \\ pt'c \end{array} \right]$$

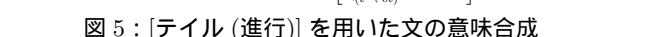


図 5: [テイル (進行)] を用いた文の意味合成

5.4 時間副詞

(9) は時間副詞 [今日][毎日] の意味表示である。[今日] の意味表示には $today$ の表す時間内に述語の表す時間的区間が存在することが記述されている。一方、[毎日] は全称量化として扱われ、II 型の形式をとる。すべての day が表す時間内に述語の表す時間的区間が存在することが記述されている。

$$(9) \quad \text{a. } \text{今日} \vdash (S \setminus NP) / (S \setminus NP) : \lambda p. \lambda x. \lambda t. \lambda c. \left[\begin{array}{l} t \sqsubseteq \text{today} \\ ptc \end{array} \right]$$

$$\text{b. } \text{毎日} \vdash (S \setminus NP) / (S \setminus NP) :$$

$$\lambda p. \lambda x. \lambda t. \lambda c. (t' : \text{time}) \rightarrow \left[\begin{array}{l} t' \sqsubseteq t \\ \text{day}(t') \end{array} \right] \rightarrow \left[\begin{array}{l} t'' : \text{time} \\ t'' \sqsubseteq t' \\ pxt''c \end{array} \right]$$

依存型理論の枠組みでは、文 A から文 B が推論可能であるという関係を意味表示間の演繹可能性として捉える。すなわち、文 A の意味表示から文 B の意味表示への証明が存在すれば、その文間での推論は成立することになる。以下に示す (10a) から (10b) は、[毎日] が表す時間内に [今日] が含まれるという仮定のもとで推論が可能となる。

$$(10) \quad \text{a. } \text{太郎は毎日レッドブルを飲む。}$$

$$\text{b. } \text{太郎は今日レッドブルを飲む。}$$

図 7, 8 (次頁) はそれぞれ (10a) と (10b) の意味合成であり、(10a) から (10b) への演繹関係は以下のように表される。

$$\Gamma, x : (t' : \text{time}) \rightarrow \left[\begin{array}{l} t' \sqsubseteq t \\ \text{day}(t') \end{array} \right] \rightarrow \left[\begin{array}{l} t'' : \text{time} \\ t'' \sqsubseteq t' \\ e:\text{drink}(\text{tarou}, \text{redbull}) \\ t'' = \text{dur}^*(e) \\ t'' \prec st \end{array} \right] \vdash \left[\begin{array}{l} t \sqsubseteq \text{today} \\ e:\text{drink}(\text{tarou}, \text{redbull}) \\ t = \text{dur}^*(e) \\ \neg(t \prec st) \end{array} \right] \quad \text{true}$$

図 9 (次頁) は上記の演繹関係の計算過程である。ここで a は、[毎日] の中に [今日] が含まれるという仮定を指す。図 9 では証明項 $(x(\text{today}))(a, d)$ が導出されているため、(10a) の意味表示から (10b) の意味表示は演繹可能であり、つまり (10a) から (10b) は推論可能であることが説明される。

今回は主に副詞による時間関係に注目したが、意味表示に証明論的意味論を用いる事で、事実性判断や確実性判断、含意関係認識など、さまざまな出来事間の関係を意味表示間の演繹関係として捉える事が可能となる。

6. おわりに

本論文では、日本語のテンス・アスペクト体系の分析に向けて依存型意味論に時間的概念を導入した。複数の用法をもつ語については、意味表示ごとの分類を設ける事で様々な言語現象を予測・説明できると考えている。引き続きテンスに関わる語の用法分類の作成と意味表示の構築を行っていく予定である。

参考文献

- [1] Bekki Daisuke. Representing Anaphora with Dependent Types In Logical Aspectes of Computational Linguistics (8th international conference, LACL2014, Toulouse, France, June 2014 Proceedings), N.Asher and S.Soloviev (Eds), LNCS 8535, pp.14-29, Springer, Heiderburg, 2014.
- [2] Kamp, Hans, and Uwe Reyle. From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory. No. 42. Springer Science & Business Media, 1993.
- [3] Martin-Löf, Per. Intuitionistic Type Theory. vol. 17. Naples: Italy: Bibliopolis, 1984.
- [4] Steedman. Surface Structure and Interpretation. The MIT Press, Cambridge, 1996.
- [5] Yoshimoto, Kei, and Jean De Dieu Mvuanda. Tense and aspect in Japanese and English. Lang, 1998.
- [6] 井上優. 現代日本語の「タ」 主文末の「...タ」の意味について .「た」の言語学. ひつじ書房, 2001.
- [7] 宇津木舞香, 稲田和明, 金子貴美, 戸次大介, 乾健太郎. 形式意味論に基づく出来事間関係認識に向けて リソース構築の展望とテンス「タ」のアノテーション. NLP2015.
- [8] 中村ちどり, 日本語の時間表現, 日本語研究叢書 14, くろしお出版, 2001.
- [9] 戸次大介. 日本語文法の形式理論 活用体系・統語構造・意味合成, 日本語研究叢書 24. くろしお出版, 2010.
- [10] 益岡隆志・田窪行則, 基礎日本語文法・改訂版, くろしお出版, 1992.

カテゴリー	例文
1 過去	「モーツアルトは交響曲を作曲した。」「8時に朝ご飯を食べた。」「父はもう死んだ。」「昔は英語が話せた。」「太郎は自宅にいた。」
2 直前	「お、見えた。」「(赤ちゃんを眺めていたら笑い出した)あ、笑った!」(野球の実況中継)イチロー、走った!」
3 直前の認知	「あ、ここにあった。」「(先ほど買ったCDを確認して)ベートーベンの第九だった。」「なんだ、井上さんでしたか。」
4 過去の認知	「そういえば、今日は給料日だった。」「そういえば、今日の午後にはミーティングがあった。」「そうか、そういう手があったか。」
5 命令	「ちょっと待った!」「さあ、行った行った!」「ほら、買った買った!」
6 あいさつ	「ありがとうございます。」「お疲れさまでした。」「ご苦労様でした。」
7 慣用表現	「この試合、買った!」「勝った!死ねえ!」

表 6 : [タ] 用法分類

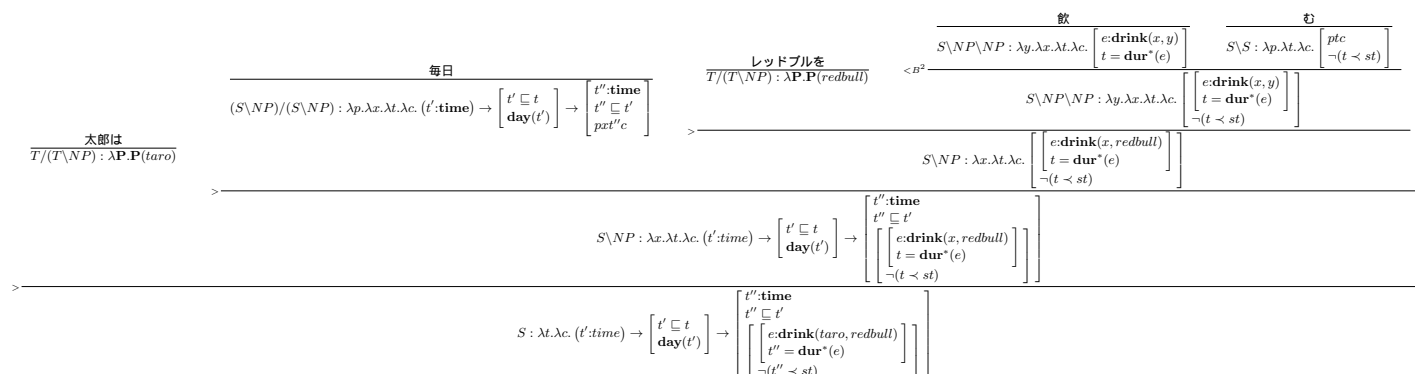


図 7 : 文 A 「太郎は毎日レッドブルを飲む」の意味合成

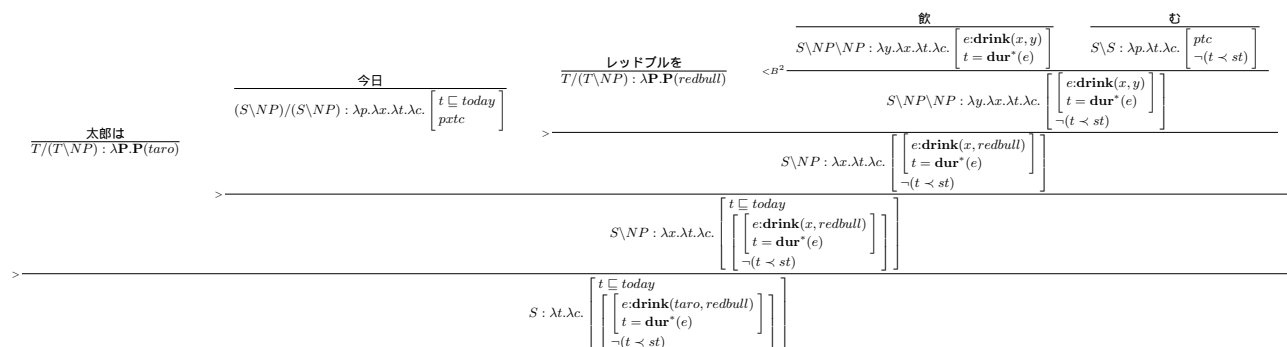


図 8 : 文 B 「太郎は今日レッドブルを飲む」の意味合成

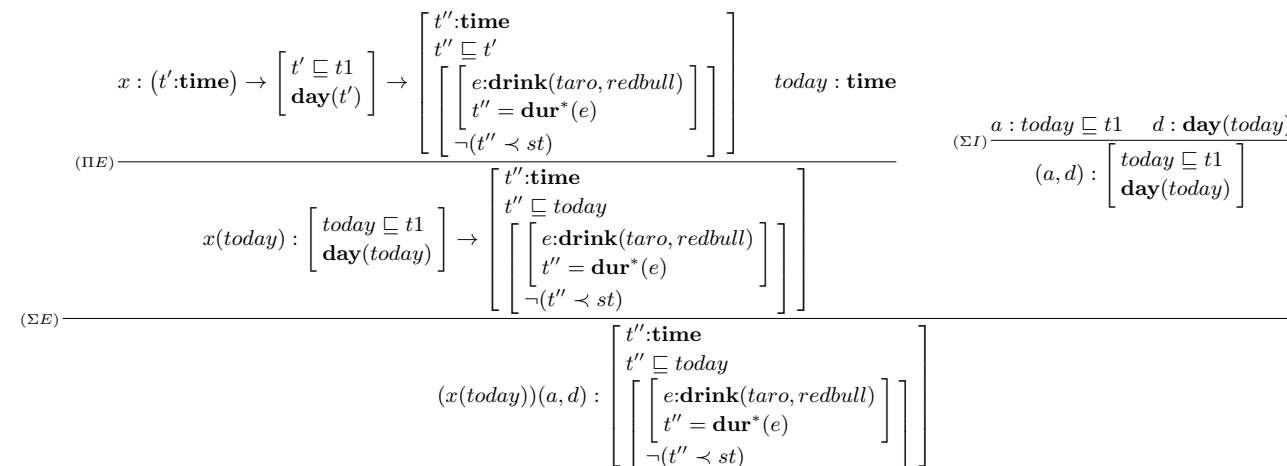


図 9 : 文 A の意味表示から文 B の意味表示への導出過程