

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 10 日現在

機関番号：12611

研究種目：若手研究(A)

研究期間：2010～2013

課題番号：22680013

研究課題名(和文)メタラムダ計算とモナドによる自然言語の意味・文脈情報の統合

研究課題名(英文)A Unified Framework for Natural Language Semantics and Pragmatics by Meta-Lambda Calculus and Linguistic Monads

研究代表者

戸次 大介 (BEKKI, Daisuke)

お茶の水女子大学・大学院人間文化創成科学研究科・准教授

研究者番号：90431783

交付決定額(研究期間全体)：(直接経費) 5,000,000円、(間接経費) 1,500,000円

研究成果の概要(和文)：本研究の目的は、メタラムダ計算とモナドを用いて、意味・談話情報を統一的に記述することである。具体的な研究項目は、1) メタラムダ計算とモナドの基礎理論、2) モナドによる意味・談話情報の記述と経験的検証、の二つからなる。1) については、メタラムダ計算の型付き言語としての形式的性質(圏論的意味論、代入操作の健全性、変換の健全性、等)を示し、2)については、メタラムダ計算においては「非決定性モナド」「大域変数モナド」「継続モナド」「慣習の含意のモナド」の四つのモナドを用いることによって、意味・談話の境界に位置する複数の言語現象を統一的に表現できることを示した。

研究成果の概要(英文)：The aim of this project is to provide a uniform framework for the representations of natural language semantics and pragmatics, by means of Meta-Lambda Calculus (MLC) and (linguistic) monads. We have established a formal foundation of MLC and the notion of linguistic monads, especially its categorical semantics, the soundness theorems in the base- and the meta-level substitutions, and the soundness theorems for alpha/beta/eta-conversions in both the base- and the meta-levels.

We have also provided some linguistic applications of MLC and their empirical evaluation. We proposed a unified system for describing phenomena that lie in the interface of semantic and pragmatic by means of linguistic monads such as "non-deterministic monad", "global variable monad (=state monad)", "continuation monad", and "conventional implicature monad".

研究分野：総合領域

科研費の分科・細目：知能情報学

キーワード：自然言語処理 数理言語学 型理論 圏論

1. 研究開始当初の背景

Steedman (2000)による組合せ範疇文法 (CCG)の一連の研究が自然言語の統語論・意味論にもたらした変化の一つは、範疇文法 (CG)が、もはや初期のモンタギュー文法のような「英語のフラグメント」でもなければ、「論理学者のみが使うトイ・グラマー」でもなく、現実のテキスト(real text)を解析するのに十分なポテンシャルを持つ枠組みであることが示されたことにある(Hockenmaier (2003), Boset.al. (2004))。CCG と古典的 CG の差は、主に関数合成規則 (Functional Composition)の有無によってもたらされるが、この差が CG と CCG の弱生成能力を 2 型と 1.5 型に分ち、ひいては現実のテキストに対する記述能力を大きく差別化している。

一方、日本語においては、包括的な形式文法理論 (井上の変形成文法、Gunji (1987) の JPSG 等)が片手で数えるほどしか存在しない状態が続いており、形式文法理論に基づく自然言語処理はあまり有望視されていなかった。筆者は日本語における形式文法理論の不在の原因が、日本語には活用現象があり、特に用言語尾において階層構造と連接構造の対立が生じる点にあると指摘し、その対立が CCG の関数合成規則によって解消されるものとして、CCG による日本語の形式文法理論を提示した (戸次 2010「日本語文法の形式理論」)。この日本語 CCG 文法は、次に挙げるような、一般的には互いに相容れないとされる特長を併せ持つものである。

網羅性：日本語の動詞・形容詞・状詞の語幹・活用語尾から、助動詞、接尾語、体言、態、複文構造、発話行為に至るまで、言語現象を包括的かつ例外なく記述している。

形式的厳密性：記号論理学・型理論・プログラミング意味論の伝統から受け継いだ数学的に厳密な体系に基づいている。

統合性：活用体系のような形態論から、統語論、意味論に至るまで、言語の複数のレベルの情報を統合的に記述している。

CCG は他の形式文法理論 (HPSG, LFG, TAG 等)と比較して、近年の統語解析器において解析エラー要因の六割を占める「等位接続構文」の解析において利点を持ち、言語処理コミュニティにおける潜在的な需要が高まりつつある。また、「日本語文法の形式理論」の CCG 文法に基づいた日本語パーザの開発も進みつつあり、これまで乖離の甚だしかった文科系言語学と工科系言語処理の共通のプラットフォームとなることが期待されている。

2. 研究の目的

英語・日本語、その他の諸言語において、現実的な文法としての CCG の記述能力が示されつつある中で、理論的には次の段階、すなわち CCG を用いた文脈・談話情報の記述方法が課題となっていた。一般的に、文脈・談話情報を扱う言語理論は各論的であり、それら全体の背後にあるシステムは不明瞭である。

それに対して本研究課題では、CCG において意味表示の記述に型付きラムダ計算を用いることに着目し、同じく型付きラムダ計算を用いるプログラミング意味論の手法であるモナド(monad)を用いることで、文脈・談話情報を統一的な枠組みで記述することを試みた。具体的には、1) モナドの理論的整備、および 2) モナドを用いた自然言語の文脈・談話情報の統一的記述と経験的検証を目的とする。

3. 研究の方法

モナドは一般的には Haskell のような関数型言語の拡張記法として知られているが、元々は圏論の概念であり、代数の自由構成の定式化の一つである。すなわち、代数の忘却関手が必ず左随伴であるために存在する随伴関手と自身を合成して得られる自己関手がモナドであり、そのモナドからは元の代数が再構成できる。(cf. MacLane (1997) p.140) 一方、Lambek (1980)以来、型付きラムダ計算の意味論を圏によって与える手法が一般的となり、Moggi (1989)では圏による意味論をモナドによって拡張することで、それまで純粋な関数型言語では表現し得なかった計算的要素 (例外処理、非決定性、入出力、大域変数、環境など) を関数型言語の表示の意味論のなかに埋め込むことに成功した。

このブレイクスルーは、意味記述に型付きラムダ計算を用いる自然言語の意味論にも自然に適用することができる (Shan (2001), Bekki and Asai (2010))。自然言語において、関数型言語における計算的要素に対応するのは、文脈・談話情報が関わる諸現象である。Shan (2001)は、Wadler (1993)によるモナドの定式化を自然言語に応用し、フォーカス、内包性、量化といった現象を記述する方法を提案した。しかし、自然言語の意味論では、関数型プログラミング言語の意味論とは型付きラムダ計算の使用法が異なる (関数型言語では、自然言語の意味論で多用される「高階引数を取る定数記号」が現れない)ため、Wadler の方法では変換できない意味表示が存在する。その問題を解決するために、筆者は Bekki (2009)において「メタラムダ計算」という体系を新たに定義した。メタラムダ計算では、メタ変数、メタラムダ抽象 (ゼータ演算子) の間に以下の関係が成り立つ。

$$(\lambda x. \lambda x.X)(x) = x.x$$

上式が成り立つのは、メタ変数 X が、自身に含まれる自由変数 x の情報を保持していることによる。これを用いて、圏論のモナドとより強い対応を持つモナド (Internal Monad と呼ぶ) が定式化され、さらに Internal Monad をパラメータとするラムダ項の変換規則が定義される。この変換規則は、Wadler (1993) と比べて型理論的により単純であり、より頑健である。

この方式により、様々な種類の「文脈・談話情報」を、それぞれにモナドを定義し、それによる意味表示の変換規則を用いることで表すことができる。

研究の手順としては、1) メタラムダ計算とモナドの基礎理論、およびそれを用いた 2) モナドによる文脈・談話情報の記述と経験的検証、の順で研究を進めた。1) の段階では、メタラムダ計算の型付き言語としての性質 (Progress, Preservation, Normalization 等) Curry-Howard-Lambek 同型において対応する論理や圏の探求、健全性・完全性の問題、およびモナドの圏論的位置付けを研究した。2) の段階では、現時点では「非決定性モナド」「大域変数モナド」「限定継続モナド」「慣習の含意のモナド」の四つのモナドを想定し、それぞれを用いた実際の言語記述と、それに伴う理論言語学の問題の解決を行った。

4. 研究成果

4.1. 理論的研究成果

メタラムダ計算では、通常のラムダ式(ベースレベルの式)に加えてメタレベルのラムダ式も扱え、その意味付けは米田関手によって圏論上で与えられている。もともとはプログラミング言語における impure な要素はモナドとして表現可能である、という事実に基づいて構築されたものだが、その定義をプログラミング言語の視点から見直すと、(2レベルの式が扱えるため) ステージ言語の1つと思える。さらに2つのレベルの式を同一視すれば通常のラムダ式になるので erasure property が成り立つという利点があり、マルチステージ言語に対する基礎理論にもなり得る。

本研究では、代入構文が独立に与えられていないメタラムダ計算に対し、その圏論的意味付けを与えた。また、ベースレベル、メタレベルの λ , μ 変換の健全性も示した。これらの成果は主に (Bekki and Masuko (to appear)) にまとめた。

4.2. 言語学的研究成果

本研究の言語学への貢献として、モナドに

よって扱える文脈・談話情報を、以下のような言語現象にまで拡大したことが挙げられる。

たとえば「フォーカスと only の意味論」[日本語のト・ヤ・カによる名詞句等位接続]の問題において、モナドによる定式化がブレークスルーをもたらしていることを示した。前者は継続モナドにより、後者は非決定性モナドと同型の、複数形モナドによる。

これらの研究により、モナドによる言語理論の一般性が示されつつある。

4.3. 残る課題と今後の展望

様々なモナドを用いた言語現象の分析は統一的ではあるが、それらのうち言語現象の側からみてモナドを本質的に必要とするのは継続モナドのみである。しかし継続モナドについても、(限定)継続の枠組みを持った型理論があれば、モナドを介した定式化は必ずしも必要ではなく、型理論に CPS 変換などの手法によって直接的に継続が導入されていけば十分であるとも言える。

一方、高階の型理論の導入は、意味論の形式的な研究にとって不可避であることも判明しつつある。したがって、高階型理論に継続を導入した体系が、自然言語の意味論にとって必要な理論体系であり、本研究の次の段階において追究されるべきものである。そのような体系が、モナドによって解決してきた諸問題を含めた統一的分析となりうるか、という問題意識のもとで、本研究の成果を今後も発展的に継承していきたい。

5. 主な発表論文等

【雑誌論文(査読有り)](計25件)

1. Meta-Lambda Calculus and Linguistic Monads
Bekki, Daisuke. Masuko, Moe. (to appear). In "Formal Approaches to Semantics and Pragmatics", McCready, E. (Ed), Studies in Linguistics and Philosophy, Vol no. 95, ISBN: 978-94-017-8812-0, Springer.
2. Building a Japanese Corpus of Temporal-Causal-Discourse Structures Based on SDRT for Extracting Causal Relations
Kaneko, Kimi. Bekki, Daisuke. (2014). In Proceedings of the EAACL2014 Workshop on Computational Approaches to Causality in Language (CAtoCL), 26th April, Gothenburg, Sweden.
3. Neighborhood-Sheaf Semantics を用いた一階述語条件論理の意味論

山本華子, 戸次大介. (2014). 第 16 回プログラミングおよびプログラミング言語ワークショップ(PPL2014)論文集, 阿蘇の司 ビラパークホテル (熊本県), 2014/3/5-7.

4. 限定継続を用いたフォーカスの分析と実装

叢悠悠, 戸次大介. (2014). 第 16 回プログラミングおよびプログラミング言語ワークショップ(PPL2014)論文集, 阿蘇の司 ビラパークホテル (熊本県), 2014/3/5-7.

5. A Type-theoretic Approach to Double Negation Elimination in Anaphora

Bekki, Daisuke. (2013). In Proceedings of the Tenth International Workshop on Logic and Engineering of Natural Language Semantics (LENLS10), pp.56-68, ISBN:978-4-915905-57-5 C3004 (JSAI), JSAI International Symposia on AI 2013, Tokyo, Japan.

6. Constructive Generalized Quantifiers Revisited

Tanaka, Ribeka. Nakano, Yuki. Bekki, Daisuke. (2013). In Proceedings of the Tenth International Workshop on Logic and Engineering of Natural Language Semantics (LENLS10), pp.69-78, ISBN:978-4-915905-57-5 C3004 (JSAI), JSAI International Symposia on AI 2013, Tokyo, Japan.

7. First-order conditional logic and neighborhood-sheaf semantics for analysis of conditional sentences

Yamamoto, Hanako. Bekki, Daisuke. (2013). In Proceedings of the Tenth International Workshop on Logic and Engineering of Natural Language Semantics (LENLS10), pp.96-106, ISBN:978-4-915905-57-5 C3004 (JSAI), JSAI International Symposia on AI 2013, Tokyo, Japan.

8. Building Japanese Textual Entailment Specialized Data Sets for Inference of Basic Sentence Relations

Kaneko, Kimi. Miyao, Yusuke. Bekki, Daisuke. (2013). In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL2013), 4-9 August, Sofia, Bulgaria.

9. Subtyping in Logical Polysemy and Copredication

Bekki, Daisuke. Asher, Nicholas. (2013). New Frontiers in Artificial Intelligence (JSAI-isAI 2012 Workshops, LENLS, JURISIN, MiMI, Miyazaki, Japan, November and December 2012, Revised Selected Papers),

Yoichi Motomura, Alastair Butler, Daisuke Bekki (Eds.), LNAI 7856, pp.17-24, Springer, Heidelberg.

10. On the semantic relation between noun phrases and quantity expressions in Japanese

Hayahishita, J.-R.. Bekki, Daisuke. (2013). New Frontiers in Artificial Intelligence (JSAI-isAI 2012 Workshops, LENLS, JURISIN, MiMI, Miyazaki, Japan, November and December 2012, Revised Selected Papers), Yoichi Motomura, Alastair Butler, Daisuke Bekki (Eds.), LNAI 7856, pp.41-52, Springer, Heidelberg.

11. Subtyping in Logical Polysemy and Copredication

Bekki, Daisuke. Asher, Nicholas. (2012). In Proceedings of the Ninth International Workshop on Logic and Engineering of Natural Language Semantics (LENLS9), pp.98-105, ISBN: 978-4-915905-51-3 C3004(JSAI), JSAI International Symposia on AI 2012, Miyazaki, Japan.

12. On the semantic relation between noun phrases and quantity expressions in Japanese

Hayahishita, J.-R.. Bekki, Daisuke. (2012). In Proceedings of the Ninth International Workshop on Logic and Engineering of Natural Language Semantics (LENLS9), pp.198-209, ISBN: 978-4-915905-51-3 C3004(JSAI), JSAI International Symposia on AI 2012, Miyazaki, Japan.

13. Conjoined nominal expressions in Japanese - Interpretation through monad - (revised version)

Hayashishita, J.-R.. Bekki, Daisuke. (2012). In New Frontiers in Artificial Intelligence (JSAI-isAI 2011 Workshops, Takamatsu, Japan, December 2011, Selected Papers from LENLS8), Manabu Okumura, Daisuke Bekki, Ken Satoh (Eds.), LNAI 7258, pp.54-67, Springer, Heidelberg.

14. Extractability as Deduction Theorem in Subdirectional Combinatory Logic (extended version)

Ozaki, Hiroko. Bekki, Daisuke. (2012). In Logical Aspects of Computational Linguistics (7th international conference, LACL2012, Nantes, France, July 2012 Proceedings), LNAI 7351, pp.186-200, Springer.

15. Toward the formulation of presupposition by Illative Combinatory Logic
Ishishita, Yuri. Bekki, Daisuke. (2012). In Logical Aspects of Computational Linguistics (7th international conference, LACL2012, Nantes, France, July 2012 Proceedings), LNAI 7351, pp.74-85, Springer.
16. 言語情報の確実性アノテーションのための韓国語の様相表現の分類
齊藤学. 崔栄殊. 戸次大介. 片岡喜代子. 川添愛. (2012). 中華日本研究第 3 号, pp.17-40.
17. 一階述語条件論理 VCb(CI)とそのタブローシステム
尾崎有梨. 戸次大介. (2012). 第 14 回プログラミングおよびプログラミング言語ワークショップ(PPL2012)論文集, 南紀白浜 むさし(和歌山県), 2012/3/8-10.
18. Illative Combinatory Logic を用いた前提の定式化に向けて
石下裕里. 戸次大介. (2012). 第 14 回プログラミングおよびプログラミング言語ワークショップ(PPL2012)論文集, 南紀白浜 むさし(和歌山県), 2012/3/8-10.
19. Extractability as Deduction Theorem in Subdirectional Combinatory Logic
Ozaki, Hiroko. Bekki, Daisuke. (2011). In Proceedings of the Eighth International Workshop on Logic and Engineering of Natural Language Semantics (LENLS8), pp.80-93, ISBN: 978-4-915905-45-2 C3004(JSAI), JSAI International Symposia on AI 2011, Sunport Hall Takamatsu, Kagawa, Japan.
20. Conjoined nominal expressions in Japanese: Interpretation through monad
Hayahishita, J.-R.. Bekki, Daisuke. (2011). In Proceedings of the Eighth International Workshop on Logic and Engineering of Natural Language Semantics (LENLS8), pp.139-152, ISBN: 978-4-915905-45-2 C3004(JSAI), JSAI International Symposia on AI 2011, Sunport Hall Takamatsu, Kagawa, Japan.
21. Combinatory Categorical Grammar as a Substructural Logic - Preliminary Remarks - (revised version)
Bekki, Daisuke. (2011). New Frontiers in Artificial Intelligence (JSAI-isAI 2010 Workshops, Tokyo, Japan, November 2010, Selected Papers from LENLS7), Takashi Onoda, Daisuke Bekki, Eric McCreedy (Eds.), LNAI 6797, pp.16-29, Springer, Heidelberg.
22. Conditional Logic Cb and its Tableau System
Ozaki, Yuri. Bekki, Daisuke. (2011). In Logical Aspects of Computational Linguistics (6th international conference, LACL2011, Montpellier, France, June/July 2011 Proceedings), LNAI 6736, pp.190-204, Springer.
23. メタラムダ計算の圏論的意味論
増子萌. 戸次大介. (2011). 第 13 回プログラミングおよびプログラミング言語ワークショップ(PPL2011)論文集, pp.60-74, 定山溪ビューホテル(北海道), 2011/3/9-11.
24. 条件論理 Cb とそのタブローシステム
尾崎有梨. 戸次大介. (2011). 第 13 回プログラミングおよびプログラミング言語ワークショップ(PPL2011)論文集, pp.33-47, 定山溪ビューホテル(北海道), 2011/3/9-11.
25. 言語情報の確実性アノテーションのための様相表現の分類
川添愛. 齊藤学. 片岡喜代子. 崔栄殊. 戸次大介. (2010). 九州大学言語学論集, 第 31 号, pp.109-129, 九州大学大学院人文科学研究 院言語学研究室編.

【学会発表】(計 31 件)

- 組み合わせ範疇文法 (CCG) による日本語の敬語表現の分析に向けて
渡辺成美. 戸次大介. (2014). 言語処理学会第 20 回年次大会発表論文集(CD-ROM), B3-4, 北海道大学, 2014/3/19.
- MCN コーパスにおける形式名詞「はず」「わけ」「つもり」のアノテーション
宇津木舞香. 佐藤未歩. 青木花純. 田中リベカ. 川添愛. 戸次大介. (2014). 言語処理学会第 20 回年次大会発表論文集(CD-ROM), B7-1, 北海道大学, 2014/3/20.
- MCN コーパス: 「ノダ」にみるガイドライン作成の手法
田中リベカ. 川添愛. 戸次大介. (2014). 言語処理学会第 20 回年次大会発表論文集(CD-ROM), B7-2, 北海道大学, 2014/3/20.
- SDRT に基づく因果関係認識日本語評価データ構築手法の提案
金子貴美. 戸次大介. (2014). 言語処理学会第 20 回年次大会発表論文集(CD-ROM), B7-4, 北海道大学, 2014/3/20.
- 限定継続を用いた inverse scope の分

析と実装
叢悠悠, 浅井健一, 戸次大介. (2014). 言語処理学会第20回年次大会発表論文集 (CD-ROM), E7-2, 北海道大学, 2014/3/20.

6. 依存型意味論を用いた日本語含意関係コーパスの分析に向けて

中村絢子, 金子貴美, 戸次大介. (2014). 言語処理学会第20回年次大会発表論文集 (CD-ROM), D5-4, 北海道大学, 2014/3/20.

7. SDRTに基づく談話関係アノテーションの試み

金子貴美, 戸次大介. (2013). 第214回自然言語処理研究会, 屋久島環境文化村センター, 2013/11/14-15.

8. 組み合わせ範疇文法(CCG)による日本語の敬語表現の分析に向けて

渡辺成美, 戸次大介. (2013). 第214回自然言語処理研究会, 屋久島環境文化村センター, 2013/11/14-15.

9. 限定継続を用いたフォーカスの分析と実装に向けて

叢悠悠, 浅井健一, 戸次大介. (2013). 第214回自然言語処理研究会, 屋久島環境文化村センター, 2013/11/14-15.

10. 基本文関係に分解した日本語含意関係認識アノテーション

金子貴美, 戸次大介, 宮尾祐介. (2013). 第27回人工知能学会全国大会論文集 (CD-ROM), 4B1-5, 富山国際会議場, 2013/6/7.

11. CCG パーザを用いた未知語の統語範疇自動推定

田中リベカ, 宮尾祐介, 戸次大介. (2013). 第27回人工知能学会全国大会論文集 (CD-ROM), 4B1-6, 富山国際会議場, 2013/6/7.

12. Extractability as Deduction Theorem in Subdirectional Combinatory Logic

Ozaki, Hiroko, Bekki, Daisuke. (2012). The 14th Programming and Programming Language Workshop (PPL2012), Nanki-Shirahama Musashi, Wakayama, Japan, 2012/3/8-10.

13. "Focus Movement" by Delimited Continuations

Bekki, Daisuke, Asai, Kenichi. (2011). ACM SIGPLAN Continuation Workshop 2011, co-located with ICFP 2011, Tokyo, Japan Saturday, September 24, 2011.

14. Conjoined nominal expressions in Japanese

Hayahishita, J.-R., Bekki, Daisuke. (2011). The 21st Japanese/Korean

Linguistics Conference, The 21st Japanese/Korean Linguistics Conference, Poster session, Oct. 20-22, 2011, Seoul National University.

【図書】(計4件)

1. New Frontiers in Artificial Intelligence (JSAL-isAI 2012 Workshops, Miyazaki, Japan, Selected Papers from LENLS9, JURISIN 2012, ALSIP2012 and MiMI)

Motomura, Yoichi, Butler, Alastair, Bekki, Daisuke. (Eds.) (2013). Lecture Notes in Artificial Intelligence 7856, Springer

2. 数理論理学

戸次大介. (2012). 東京大学出版会

3. New Frontiers in Artificial Intelligence (JSAL-isAI 2011 Workshops, LENLS, JURISIN, ALSIP, MiMI: Takamatsu, Japan, December 2011, Revised Selected Papers)

Okumura, Manabu, Bekki, Daisuke, Satoh, Ken. (Eds.) (2012). Lecture Notes in Artificial Intelligence 7258, Springer (290 pages).

4. New Frontiers in Artificial Intelligence (JSAL-isAI 2010 Workshops, Tokyo, Japan, November 2010, Selected Papers)

Onoda, Takashi, Bekki, Daisuke, McCreedy, Eric. (Eds.) (2011). Lecture Notes in Artificial Intelligence 6797, Springer.

【その他】

ホームページ等

<http://www.is.ocha.ac.jp/~bekki/>

6. 研究組織

(1) 研究代表者

戸次大介 (BEKKI Daisuke)

お茶の水女子大学・大学院人間文化創成科学研究科

研究者番号: 90431783

(2) 研究分担者 (なし)

(3) 連携研究者 (なし)