

大規模世界知識を用いた仮説推論による談話解析の課題と対策

井之上 直也[†] 乾 健太郎[†] Ekaterina Ovchinnikova[‡] Jerry R. Hobbs[‡]

[†]東北大学 大学院情報科学研究科
{naoya-i, inui}@ecei.tohoku.ac.jp

[‡]University of Southern California / Information Sciences Institute
{katya, hobbs}@isi.edu

1 はじめに

ある一定の意味をもつ文のまとまりを**談話**という。談話の作り手の多くは、読み手にとって自明と考えられる情報をしばしば省略する。例 (1) の談話を考えてみよう。

(1) *Steve will resign from Microsoft next week. His products have attracted many people over the years.*

この談話では、「Steve が Microsoft で働いている」「His が Steve を指している」「Steve が products を製作した」といった、読み手にとって自明と考えられる情報が省略されている。本稿では、このような情報を機械により自動的に顕在化する処理、**談話解析**に取り組む。談話解析技術の実現は、情報抽出や質問応答など自然言語処理の様々な応用技術において重要だと考えられる。

談話解析を実現するためには、例 (1) から分かるように、照応関係の理解、名詞間の関係理解など、様々な種類の推論を組み合わせて談話の内容と整合する情報を出力する必要がある [23]。この処理を自然に表現できる枠組みとして、観測に対する最良の説明を求める論理的推論、**仮説推論**がある。詳しくは後述するが、仮説推論を用いることで、談話解析に必要な推論を背景知識として宣言的に記述することができ、様々な推論の組み合わせによる情報の顕在化の処理は、説明生成の問題に落とし込むことができる。仮説推論には様々な枠組みが存在するが [11, 3, 22, etc.], 我々は Hobbs ら [11] の重み付き仮説推論を用いる。

仮説推論による談話解析を実世界の問題に適用するには、推論に必要な背景知識を十分な規模で取り揃える必要がある。我々は、近年の知識獲得技術の発展により利用可能となった語彙知識を用いて、実スケールで動く談話解析の実現を目的として以下の3つの課題に取り組む:

I) 仮説推論の計算効率の改善: 仮説推論は組み合わせ最適化問題であるため、大規模なデータの上でも実用的な時間で動く推論エンジンが必要である。

II) 実問題で顕在化する課題の発見と対策: 談話解析を実スケールで動かした前例がほとんどないため、実スケールで動かすことで初めて顕著になる問題を明らかにする必要がある。

III) 解析結果の評価方法の検討: 談話解析は情報を“生成する”問題である。直接的な評価が難しいため、評価の方法を検討する必要がある。

本稿では、このうち **I** と **II** の課題に対して解決策を与え、**III** に対しては人手による予備評価実験のほか、RTE[4] による評価を行った結果を報告する。本稿では、まず談話解析の関連研究を概観し、重み付き仮説推論の形式的な説明を与える (2 節)。次に、我々が目標とする

談話解析を仮説推論により実現する方法を説明し (3.1 節)、仮説推論を整数線形計画法 (ILP) により高速化する手法 (3.2 節)、実世界のデータへの適用の際の問題点とその解決策を述べる (3.3 節)。最後に、談話解析モデルを評価した結果について報告する (4 節)。

2 背景

2.1 談話解析の先行研究

文章に明示されていない情報を顕在化する最近の試みとして、Machine Reading (MR) プロジェクト [7] がある。MR のゴールは、世界知識の獲得から談話解析までを教師なし学習 (または自己教師あり学習) で実現することにより、機械の言語理解のための end-to-end solution を開発することである。これまでに MR の各参加チームによりいくつかの情報抽出器が開発されてきたが [18, 8, etc.], いずれも関係の抽出に特化したモデルとなっており、我々のような汎用的な推論の枠組みにはなっていない。

重み付き仮説推論を談話解析の推論機構として利用する最近の試みとしては、文献 [17, 1] がある。Ovchinnikova ら [17] は重み付き仮説推論を用いた談話解析モデルを RTE[4] により評価し、RTE-2 において6割程度の正解率が得られたことを報告している。しかしながら、文献 [17] で用いられている推論エンジン Mini-TACITUS[16] は、大規模データに対して最適解を得ることがほとんどできないことが報告されている [12]。Blythe ら [1] は、Markov Logic Networks (MLNs)[19] を用いた重み付き仮説推論の実装手法を提案しているが、仮説推論を演繹推論に変換するために公理の数が膨大になり [22]、大規模な知識を用いた推論には向きだと考えられる。これらの研究に比べて、我々は実スケールで動く高速な推論エンジンを用いることにより、より現実的な設定のもとで談話解析に取り組んでいるといえる。

2.2 重み付き仮説推論

仮説推論とは、観測に対する最良の説明を求める推論である。形式的には、次のように定義される。

Given: 背景知識 B , 観測 O 。ただし B と O は一階述語論理式の集合。

Find: 仮説 H 。ただし H は $H \cup B \models O, H \cup B \not\models \perp$ を満たす一階述語論理式の集合。

本稿では、上の条件を満たす (複数の) H を**候補仮説**と呼び、そのうち最良の仮説を**解仮説**と呼ぶ。候補仮説の良さを評価する方法は様々だが、我々は重み付き仮説推論の枠組み [11] を用いる。重み付き仮説推論では、観測 O と仮説 H は存在限量されたりテラルの連言として

表現され、各リテラルは正の実数のコスト^{*1}を持つ。背景知識 B は、 $P_1^{w_1} \wedge \dots \wedge P_n^{w_n} \Rightarrow Q_1 \wedge \dots \wedge Q_m$ の形式を持ち、すべての変数が存在限量された一階述語論理式(公理)の集合である。前件のリテラルは、正の実数の重み^{*2}が割り当てられている。コストと重みの直感的な解釈は、リテラルが示す命題の仮説らしさ (implausibility) であり、値が小さいほど蓋然的な仮説となる。

重み付き仮説推論では、次の二つの操作を観測と候補仮説に逐次適用することにより、候補仮説を生成する: (i) **後ろ向き推論**は、背景知識と観測から、新しい仮説を導き、コストを伝播させる操作である。例えば、 $O = \exists x(q(x)^{s_{10}})$, $B = \{\forall x(p(x)^{1.2} \Rightarrow q(x))\}$ が与えられた場合、次の二つの候補仮説を生成する: $H_1 = \exists x(q(x)^{s_{10}})$, $H_2 = \exists x(q(x)^{s_0} \wedge p(x)^{s_{12}})$ 。新しい仮説のコストは、適用対象のリテラルのコストと背景知識の重みの掛け算により決定され(ここでは $1.2 \cdot \$10 = \12)、適用対象のリテラルのコストが 0 となる。(ii) **単一化**は、同一の述語を持つ 2 つのリテラルを合成し、互いの項が等価であるという仮説を置く操作である。例えば、 $O = \exists x, y(p(x)^{s_{10}} \wedge p(y)^{s_{20}} \wedge q(y)^{s_{10}})$ が与えられたとき、 $H = \exists x, y(p(x)^{s_{10}} \wedge q(y)^{s_{10}} \wedge x = y)$ を候補仮説として生成する。この際、2 つのリテラルが持つコストのうち小さい方だけを残す。

解説は、候補仮説の集合の中で最小のコストを持つ候補仮説、すなわち最も蓋然的な仮説と定義される。候補仮説のコストは、候補仮説に含まれるリテラル(要素仮説)のコストの和 $C(H) = \sum_{h \in H} cost(h)$ と定義される(ここで、 $cost(h)$ はリテラル h のコスト)。重み付き仮説推論の仮説評価指標の利点として、仮説の特殊性を評価できることが挙げられる。仮説推論では、後ろ向き推論が適用されるごとに仮説の特殊性が高くなるが(例えば H_2 は H_1 より特殊である)、一般には仮説が特殊になるほど仮説を支持する証拠が少なくなるため、信頼度の高い範囲で推論を停止できることが望ましい。談話解析では、候補仮説が際限なく考えられるため、仮説の特殊性を適切に決めることは特に重要である。重み付き仮説推論では、仮説コストの伝播と単一化の操作により、自然にこれが実現できている(例えば H_1 と H_2 の比較)。

3 仮説推論による談話解析

本節では、談話解析に仮説推論を用いる動機と解析手法の全体像を概観したあと、解析に必要な推論パターン(3.1 節)、1 節で述べた 2 つの課題 I, II に対する取り組みについて説明する(3.2-3.3 節)。ここでは、例(1)を題材として説明を進める。

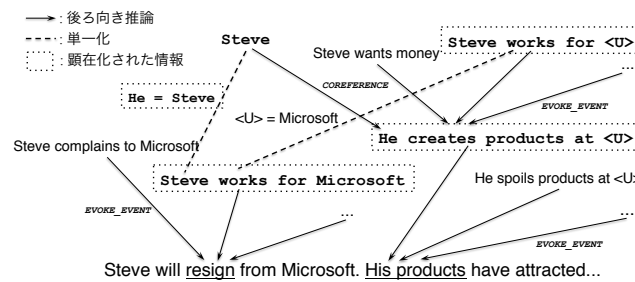


図 1: 例(1)を仮説推論により談話解析した例

^{*1}リテラル P がコスト c を持つことを P^{sc} と表記する。
^{*2}リテラル P が重み w を持つことを P^w と表記する。

(1) *Steve will resign from Microsoft next week. His products have attracted many people over the years.*

まず、顕在化したい三つの情報を再確認しよう: (i) *Steve works for Microsoft*, (ii) *His = Steve*, (iii) *Steve creates products*。ここで、もとの談話と顕在化したい情報の関係を観察してみると、顕在化には少なくとも次の二種類の処理が関わっていることが推測できる:

- **A. 候補仮説生成:** (i) 事象から別の事象を連想する (*resign* から *work for*), (ii) 代名詞の指示先を推測する (*his* から *Steve*), (iii) 名詞句間の関係を連想する (*his* と *products* の *creates* 関係)
- **B. 仮説選択:** A の推論から生成可能な情報の集合のうち、互いに整合性が取れ、かつ談話とも整合性がある情報の集合を選ぶ。例えば、*his* と *products* の間には複数の関係が考えられるが (*spoils, hates* など)、整合性の観点から *creates* を選択する。

これらの処理を計算機で実現する際に注意すべきことは、個々の推論結果が互いに依存しており、手続き的に解析順序を決めるのが難しいことである。

そこで本稿では、それぞれの推論を統合的に処理するために、仮説推論を用いる。談話解析に必要な推論をいくつかのプリミティブな推論のパターン(推論パターン)に分解し、それぞれのパターンの組み合わせで情報の顕在化を行う。仮説推論において、処理 A は各種推論パターンに必要な語彙知識を背景知識とした、談話に対する候補仮説生成作業とみなすことができ、処理 B は最良の説明(仮説の組み合わせ)を選択する問題と考えることができる。仮説推論を談話解析に用いる利点は、処理の順序をあらかじめ決めない宣言的なモデル構築ができ、各種推論の相互依存関係を自然にモデル化できることにある。例(1)を仮説推論により解析する様子を図 1 に示す。図 1 では、談話から様々なイベントを潜在要素仮説として生成し (*Steve complains to Microsoft, Steve wants money* など)、コストが最小の候補仮説を求めることにより前述の 3 つの情報が推論されている。本稿では仮説推論における意味表現として、Davidsonian 形式 [5] の論理式を用いる。

3.1 推論パターン

本稿で用いる 5 種類の推論パターン^{*3}と仮説推論における表現方法を表 1 に示す。このうち、COREFERENCE 以外のパターンは公理の後ろ向き推論で実現し、COREFERENCE は論理変数の単一化で表現する。ここでは詳細は献 [23] に委ね、COREFERENCE と既存の共参照解析の研究との関連について述べる。COREFERENCE は、2 つの mention m_1, m_2 (事象とモノの両方を含む) の共参照関係を、仮説推論における論理変数(または定数)間の単一化として表現する。Hobbs らの枠組みでは論理変数の単一化をコスト 0 としているが、ここに m_1, m_2 の共参照関係の尤もらしさの逆数を設定することで、クラストリングベースの共参照解析や ILP による結合推論型共参照解析 [15, 6, etc.] におおむね対応すると考えることができる。将来的には、変数の単一化のコスト関数を学習する学習器を設計し、既存の照応解析の枠組みとの比較を行う予定である。また、背景知識の重みについても学習アルゴリズムの設計を進めている。

^{*2}<http://webdocs.cs.ualberta.ca/~lindek/>

^{*3}文献 [23] の推論パターンに加え、RTE-2 の開発データから含意関係認識に有用と考えられる推論を吟味した。

表 1: 推論部品と仮説推論における表現

名前	機能	仮説推論における公理の例	公理の数	語彙資源*
EVOKE_EVENT	事象 E_1 から E_1 と関係 R にある別の事象 E_2 を連想	$work(e_1, x) \wedge happens_before(e_1, e_2) \Rightarrow resign(e_2, x)$	53,400	WN, FN
EVOKE_ENTITY	モノ X から X と関係 R にある別のモノ Y を連想	$room(x) \wedge part_of(x, y) \Rightarrow house(y)$ $brunei(x) \wedge member_of(x, y) \Rightarrow asean(y)$		WN, FN
EVOKE_ARG_REL	モノ X から X を項に取る事象 E を連想, または事象 E から事象 E が項としてとりうるモノ X を連想	$declare(e_1, x_1, x_2) \Rightarrow victory(x_2)$, $police(x_1) \Rightarrow arrest(e, x_1, x_2)$		FN
EVOKE_REL_NN	関係 R_n にある 2 つの名詞 N_1, N_2 の間に, 関係 R があることを連想	$work(e_1, x) \wedge for(e_1, y) \Rightarrow nn(x, y)**$, $own(e_1, x, y) \Rightarrow of(y, x)**$	8	MAN
COREFERENCE	2 つの mention m_1, m_2 が共参照関係になりうることを推測***	$japan(x) \Rightarrow country(x)$, $shoot(e, x_1, x_2) \Rightarrow hit(e, x_1, x_2)$	320,616	WN, FN, MAN

* WN: WordNet 3.0[9], FN: FrameNet 1.5[20], MAN: 人手により作成

** 2 つの要素が複合名詞または前置詞 of で結ばれる関係にあることを表す述語; *** mention はモノと事象の両方を含む

3.2 ILP による仮説選択の高速化

仮説推論における最良の説明の探索は, 候補仮説の要素仮説となりうるリテラル (潜在要素仮説) の集合から, コストが最小になるような潜在要素仮説の組み合わせを発見する問題とみなすことができる。組み合わせ最適化問題の単純な解法の計算量は $O(2^n)$ (n は変数の数) となるため, 大規模なデータを処理するためには探索アルゴリズムの工夫が必要である。また, 仮説推論は論理を用いた推論であるため, 要素仮説間の論理的制約 (排他性など) のもとで最適化が行える枠組みが必要である。

そこで本節では, 整数線形計画法 (ILP) に基づく仮説選択の手法を提案する (課題 I の解決)。提案手法では, 背景知識 B , 観測 O が与えられた際, はじめにすべての潜在要素仮説を列挙し, 潜在要素仮説集合 P を作成する。次に, すべての可能な候補仮説を表現するため, P から ILP 変数とそれらの変数間の制約を生成する。紙面の都合上, 本節では 5 つの ILP 変数のうち 2 つの変数 $h_p \in \{0, 1\}, r_p \in \{0, 1\}$ (p は潜在要素仮説) を説明するので, 詳細については文献 [12] を参照されたい。 h_p は, p が候補仮説に含まれているか ($h_p = 1$) 否か ($h_p = 0$) を表し, r_p は p が説明されているか ($r_p = 1$) 否か ($r_p = 0$) を表す。これらの ILP 変数を用いて, 最小のコストを持つ候補仮説を見つけるための ILP の目的関数を次のように定義する:

$$\min. \quad cost(H) = \sum_{p \in \{p | p \in P, h_p=1, r_p=0\}} cost(p) \quad (1)$$

つまり, 目的関数は, 候補仮説に含まれており ($h_p = 1$), かつ説明されていない ($r_p = 0$) リテラルのコストの和である (2.2 節の仮説コストの定義と一致する)。

次に, ILP の変数が表す候補仮説の空間を有効なものに限定するため, 線形制約を導入する。例えば, リテラル p がコストを支払わない状態 ($r_p = 1$) になるためには, p を説明するリテラル q が仮説に含まれている必要がある ($h_q = 1$) という制約を, $r_p \leq h_q$ という線形制約式で表す。そのほかの制約, 一階述語論理への拡張 (リテラル・変数 (または定数) 間の単一化) については, 文献 [13] を参照されたい。以下に, 我々の枠組みの特長をまとめる:

- 命題論理上の仮説推論 [21, 14, etc.] でなく, 一階述語論理上の仮説推論を高速化する
- 論理的制約は線形制約として自然に書けるため, 拡張性が高い (3.3 節を参照)
- 素朴な方法では多項式オーダーとなる変数間の等価性の推移律制約を効率的に表現, 計算できる [13]
- MLNs を用いた仮説推論 [1, 22, etc.] よりも効率が良いと期待できる (MLNs による仮説推論では, 組

み合わせ的に増える説明の排他性制約を書く必要がある)

- Operations Research の分野で開発された最新の ILP ソルバーによる高速な推論が期待できる

3.3 意味的互換性の問題と対策

3.1 節で述べたように, 単一化のメカニズムは共参照関係の認定と深く結びついており, 談話解析において重要な役割を果たす。しかしながら, 現在の枠組みは単一化操作に一切の制約 (または罰則) がないため, 最小コストを持つ仮説の中では単一化可能なリテラルはすべて単一化された状態になる。その結果, ある変数に対して二つ以上の矛盾する性質を仮説してしまう状況が発生する (例えば $O = dog(x) \wedge run(e_1, x) \wedge cat(y) \wedge run(e_2, y)$)。この問題は, 仮説推論を用いた談話解析を実スケールで行う際に極めて重要な問題だが, これまでの先行研究では実スケールでのモデル検証ができておらず (2 節を参照), この問題に十分な注意が払われてこなかった。

そこで本稿では, 任意の論理変数 (または定数) が排他的な概念を同時に持たないように, 2 種類の制約を導入する (課題 II の解決)。第一の制約として, 潜在仮説集合において, 同一の変数に対して非互換的な意味クラスや性質を持つリテラルの集合に対して, 排他性に関する制約を導入する。関数 $disj(x)$ を, 変数 x を項に持つリテラルの集合 $Q \subseteq P$ の中で, 排他的な概念を表すリテラルの集合を列挙する関数とする (例えば $\{\{white(x), black(x), red(x)\}, \{cat(x), dog(x)\}\}$)。この関数を用いて, 潜在仮説集合に存在する論理変数 (または定数) の集合の各要素 x について, $(\bigoplus_{d \in disj(x)} h_d) \vee (\bigwedge_{d \in disj(x)} \neg h_d)$ という制約を導入する。これは線形制約 $\sum_{d \in disj(x)} h_d \leq 1$ として表現できる。本稿では, 関数 $disj(x)$ は WordNet[9] 内の兄弟概念以外の概念対と, WordNet における対義語が排他的となるよう設定した。第二の制約として, 意味的互換性の矛盾を導く論理変数 (または定数) x, y の単一化についても, 排他性制約を導入する。例えば, $desk(x) \wedge lion(y)$ が候補仮説に含まれている場合, x と y の単一化を禁止する。これも第一の制約と同様に線形制約として表現できるが, 本稿では紙面の都合上記述を省略する。意味的互換性の問題を解決するほかの方法としては, 3.1 節で触れた単一化コストを使うことも考えられる。

4 評価実験

本節では, 大きく 2 つの評価実験を行う。まずひとつめに, 課題 I への取り組みを評価するため, 3.1 節で示した背景知識を用い, RTE-2 の開発セットを解析対

表 2: 開発セットに対する処理時間と談話解析の結果

D	PEH	VAR	CON	ALL	TIME	DPXP
1	89	3,378	12,075	99.8%	0.33	1.91
2	479	12,534	22,094	98.4%	1.07	4.93
3	1,171	32,941	58,612	95.8%	2.63	7.42

PEH: 潜在要素仮説集合の平均要素数, VAR: ILP 変数の数の平均, CON: ILP 制約の数の平均, ALL: 2 分以内に推論を終えた解析対象の割合 (800 問中), TIME: ALL の平均処理時間 (秒), DPXP: Text から新たに推論されたりテラルの数の平均

象として処理時間の評価を行う。ふたつめに、談話解析により顕在化した情報が有用であるかを評価する (課題 II, III への取り組み)。しかしながら、顕在化の性能を直接評価できるデータセットがないため、顕在化した情報の一部を直接人手により評価する (*intrinsic* な評価)。また、*extrinsic* な評価として、談話に潜む情報の顕在化が RTE の性能向上に役立つという報告 [10] にもとづき、RTE-2 テストセットによる評価を試験的に行う。テキストから意味表現への変換には、Boxer[2] を用いた。

4.1 実験結果

まず、ILP^{*4}による高速化法の性能評価実験について報告する。本実験では手法のスケラビリティを調べるため、後ろ向き推論の深さ制限 D を 1 から 3 まで変化させ、処理時間の違いを調べた。2 分を過ぎても推論が終わらない場合は、その際に保持している解を返し、処理を終了させた。実験の結果を表 2 に示す。表 2 の結果より、 D が増加してもほとんどすべての問題に対して最適解を求めることができ、その処理時間も現実的であることがわかった。また、表 2 より、談話解析により新しい情報が顕在化されていることが確かめられた。顕在化された情報のうち 30 事例を人手で分析したところ、23 事例は談話の内容と整合した内容であった。例えば、1 番の問題の Text “*As a result of these weaknesses, computer systems and the operations that rely on...*” からは「*computer* と *systems* が *part_of* の関係にある」という情報が顕在化できていることがわかった。次に、本稿の談話解析モデルを含意関係認識 (RTE-2 テストセット) の正解率で試験的に評価したところ^{*5}、60.4% (Bag of Words: 59.4%) の正解率を得た。人手による評価では一定の精度で潜在的情報を顕在化できていることは確認できたが、その有用性を既存の RTE データセットで証明するには至っていない。一つの理由として、既存の RTE データが極めて広範な問題を含んでおり、談話解析の効果を測定する目的には必ずしも向いていないことが挙げられる。今後は、評価方法自体の見直し、検討を含め、内生的、外生的な評価をともに考える必要がある。

5 おわりに

談話に潜む情報を仮説推論により顕在化する談話解析モデルを提案した。特に本稿では、仮説推論による談話解析モデルを実スケールで動かすに当たり問題となる 2 つの課題に注目し、解決策を示した。ひとつめの課題は、仮説推論の計算効率の改善であった。これに対し、我々は ILP による高速化法を提案した。ふたつめの課題は、実スケールで顕著な問題となる意味的互換性の問題であった。我々はこの問題に対し、矛盾する仮説を導

く要素仮説の組み合わせを制約するという解決策を示した。談話解析モデルの評価実験では、ILP による定式化が実データの上でも現実的な時間で動くことを確認し、RTE による評価の結果から、談話解析により抽出された情報が有用であることがわかった。今後の課題としては、意味的互換性の問題に取り組むにあたって、論理変数間の単一化にコストを課すことを検討しているほか、背景知識の重みを学習する手法を開発中である。また、談話解析モデルの評価については、推論された情報を直接人手により評価する内生的評価を行う予定である。

謝辞 本研究は、文部科学省科研費課題 22・9719 および課題 23240018 の一環として行われた。

参考文献

- [1] J. Blythe, J. R. Hobbs, P. Domingos, R. J. Kate, and R. J. Mooney. Implementing Weighted Abduction in Markov Logic. In *Proc. of IWCS*, pp. 55–64, Oxford, UK, 2011.
- [2] J. Bos. Wide-Coverage Semantic Analysis with Boxer. In J. Bos and R. Delmonte, editors, *Proc. of STEP, Research in Computational Semantics*, pp. 277–286. College Publications, 2008.
- [3] E. Charniak and R. P. Goldman. A Probabilistic Model of Plan Recognition. In *Proc. of AAAI*, pp. 160–165, 1991.
- [4] I. Dagan, B. Dolan, B. Magnini, and D. Roth. Recognizing textual entailment: Rational, evaluation and approaches - Erratum. *Natural Language Engineering*, Vol. 16, No. 1, p. 105, 2010.
- [5] D. Davidson. The Logical Form of Action Sentences. In N. Rescher, editor, *The Logic of Decision and Action*, pp. 81–120. University of Pittsburgh Press, 1967.
- [6] P. Denis and J. Baldrige. Joint Determination of Anaphoricity and Coreference Resolution using Integer Programming. In *Proc. of HLT-NAACL*, pp. 236–243, 2007.
- [7] O. Etzioni, M. Banko, and M. J. Cafarella. Machine reading. In *Proc. of AAAI*, 2006.
- [8] O. Etzioni, A. Fader, J. Christensen, S. Soderland, and M. T. Center. Open Information Extraction: The Second Generation. In *Proc. of IJCAI*, pp. 3–10, 2011.
- [9] C. Fellbaum, editor. *WordNet: an electronic lexical database*. MIT Press, 1998.
- [10] A. Hickl and J. Bensley. A Discourse Commitment-Based Framework for Recognizing Textual Entailment. In *Proc. of ACL-PASCAL Workshop on Textual Entailment and Paraphrasing*, pp. 171–176, 2007.
- [11] J. R. Hobbs, M. Stickel, P. Martin, and D. Edwards. Interpretation as abduction. *Artificial Intelligence*, Vol. 63, pp. 69–142, 1993.
- [12] N. Inoue and K. Inui. ILP-Based Reasoning for Weighted Abduction. In *Proc. of AAAI Workshop on Plan, Activity and Intent Recognition*, 2011.
- [13] N. Inoue and K. Inui. An ILP Formulation of Abductive Inference for Discourse Interpretation. Technical Report 3, 2011-09-09.
- [14] M. Ishizuka and Y. Matsuo. SL Method for Computing a Near-optimal Solution using Linear and Non-linear Programming in Cost-based Hypothetical Reasoning. In *PRCAI*, pp. 611–625, 1998.
- [15] M. Klenner and E. Ailloud. Enhancing coreference clustering. In *Proc. of the Second Workshop on Anaphora Resolution*, pp. 31–40, 2008.
- [16] R. Mulkar, J. Hobbs, and E. Hovy. Learning from Reading Syntactically Complex Biology Texts. In *Proc. of the 8th International Symposium on Logical Formalizations of Commonsense Reasoning*, Palo Alto, 2007.
- [17] E. Ovchinnikova, N. Montazeri, T. Alexandrov, J. R. Hobbs, M. McCord, and R. Mulkar-Mehta. Abductive Reasoning with a Large Knowledge Base for Discourse Processing, 2011.
- [18] A. Penas and E. Hovy. Filling knowledge gaps in text for machine reading. In *Proc. of COLING*, pp. 979–987, 2010.
- [19] M. Richardson and P. Domingos. Markov logic networks. *Machine Learning*, pp. 107–136, 2006.
- [20] J. Ruppenhofer, M. Ellsworth, M.R. Petrucci, C.R. Johnson, and J. Scheffczyk. FrameNet II: Extended Theory and Practice. Technical report, Berkeley, USA, 2010.
- [21] E. Santos. Polynomial solvability of cost-based abduction. *Artificial Intelligence*, Vol. 86, pp. 157–170, 1996.
- [22] P. Singla and P. Domingos. Abductive Markov Logic for Plan Recognition, 2011.
- [23] 杉浦純, 井之上直也, 乾健太郎. 説明生成に基づく談話構造解析の課題分析. 言語処理学会第 18 回年次大会発表論文集 (to appear), 2012.

^{*4}ILP ソルバー: <http://www.gurobi.com/>

^{*5}文献 [17] に倣い、Text 側から顕在化された情報が Hypothesis に対する最良の説明のコストをどれだけ下げたかを比較し、差が大きい場合に含意関係と認識する手法を取った。