

様式 C-19

科学研究費補助金成果報告書

平成 22 年 6 月 8 日現在

研究種目： 基盤研究（A）
研究期間： 2007 ～ 2009
課題番号： 19200010
研究課題名（和文）
自然言語意味概念の共通的記述に基づく次世代 Web 基盤環境構築とその知能化機能
研究課題名（英文）
A Next-generation Web Foundation based on Common Description of Concept Meaning
expressed in Natural Language Texts and its Related Intelligent Functions
研究代表者
石塚 満 (ISHIZUKA MITSURU)
東京大学・大学院情報理工学系研究科・教授
研究者番号： 50114369

研究成果の概要（和文）：

情報流通と情報蓄積・共有，更には共創のグローバルな基盤に成長した Web を，コンピュータがその情報コンテンツの意味内容を把握できるようにして，検索・マイニング・編集・組織化等の意味に立ち入る操作を可能にする次世代知能化 Web に向けた意味計算基盤の研究開発を行った．具体的には，メタデータ記述を中心とする Semantic Web とは異なり，我が国発の技術として我々が中心で国際標準化を行ってきた，自然言語テキストが表す概念意味の共通的記述言語である CDL(Concept Description Language)に関する研究を進めた．CDL はコンピュータ・エスペラント語とも言える固有言語に依存しない共通言語の性質を有し，Web における言語障壁克服にも寄与する技術である．CDL に関する第一の技術課題は，テキストから CDL への変換である．全自動機械翻訳が困難であると同様に，テキストから CDL への全自動変換も困難であるので，人の負担の少ない半自動変換を目指し，語義(word sense)の選択と決定を介して，CDL へ変換するアプローチについて研究した．出来るだけの語義曖昧性解消をコンピュータにより行うことにより，人手介在の部分減らすインタフェースの開発を行った．また，語句の単なるマッチングでなく，意味的なマッチングを可能とし，CDL データの意味的検索法の研究開発を行った．本研究と並行する形で W3C のインキュベータグループで CDL を次世代 Web へに向けた意味計算基盤とする国際標準化の活動も行った訳であるが，英語を主とし英語で事足りる国では CDL のような言語の必要性の認識が薄く，この進展は一旦中断している．

CDL は単語エンティティ間を関係識別子で結んで表すことをベースとしているが，関連研究として分布仮説に基づき，単語エンティティペア間の関係類似性を Web 検索エンジンの結果を利用して計算する手法を考案，開発した．この研究成果は，2009 年 5 月の WWW 国際会議(Web 技術分野で世界で最も権威ある会議)にフルペーパーとして採録され，国際的に高い評価を得た．この原理を活用し，{(Tokyo, Japan), (?, France)} のようなエンティティ対を検索入力として，{? = Paris} を答えるような潜在関係検索エンジン(Latent Relational Search Engine)を開発した．これは特許出願も行った．

更に，Web から抽出した多数の任意エンティティペアから代表的で有意な関係を有するエンティティペアを効率的に抽出する，新しい共クラスタリング法による Open Information Extraction 手法を考案，開発した．この研究成果も特許出願を行い，また 2010 年 4 月の WWW 国際会議にフルペーパーとして採録され，高い評価を得た．

研究成果の概要（英文）：

The Web has grown to an indispensable social platform for global information flow, sharing and even co-creation. We carried out our research on a foundation of semantic computing toward a next-generation Web, which allows computers to understand the semantic meaning of Web information and thereby to work on the retrieval, mining, editing and organizing Web information considering not only the surface level of information but also the semantic level. Specifically, we worked focusing on CDL (Concept Description Language), which is a common language for representing concept meaning expressed in texts. CDL is a technology originated from Japan, and has been undergone the process of international standardization in W3C promoted by our group. Because CDL is not dependent on a particular language and has a universal property for all the natural languages, it can be called “computer Esperanto language” and may also contribute to overcome the language barrier problem in the world. The current first issue on CDL is the conversion from natural language texts to the CDL representation. Like in machine translation, its full automatic conversion is impossible in near future. Thus we have studied and developed a semi-automatic conversion method passing through the process of word sense disambiguation, where a human writer/operator selects the correct word sense when the computer cannot determine the correct one. For another issue regarding CDL, we have developed a semantic retrieval method for CDL data, which achieves a semantic-level matching efficiently.

The base of the CDL representation is to connect two word entities in a sentence with one of predefined relational labels. As its related research, we have devised an efficient method for computing the similarity between two word entities based on the distributional hypothesis through the use of a search engine. This research was highly evaluated internationally such that it was accepted as a full paper at WWW2009 (the most prestigious conference in the Web technology field). Based on this research result, we have developed a new type search engine named latent relational search engine, which accepts two entity pairs with one blank element, e.g., {(Tokyo, Japan), (?, France)}, as a query, and produces an answer, e.g., {? = Paris} in this case. The mechanism of this latent relational search has been applied for patent.

In addition, as an extension of this relational search, we have invented an efficient co-clustering method for mining typical and meaningful entity pairs from lots of entity pair candidates extracted from Web texts; this framework is called open information extraction. This research was also accepted as a full paper at WWW2010, and has been also submitted for a patent application.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	11,100,000	3,330,000	14,430,000
2008年度	9,800,000	2,940,000	12,740,000
2009年度	10,900,000	3,270,000	14,170,000
年度			
年度			
総計	31,800,000	9,540,000	41,340,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

1. 研究開始当初の背景

WWW(Web)は今や情報流通と情報蓄積・共有の重要なグローバルなインフラストラクチャに成長し、今後ともその重要性は増し活用の幅も拡大していく。我が国でも多くのWeb関係の研究が行われているが、W3Cを中心として国際標準化され出来あがったWebの土俵を前提にしたものであり、土俵としての次世代Web基盤構築を目指す研究は非常に乏しく、我が国のWeb関連の国際標準化への貢献は皆無といってよい状態である。Web創始者であるTim Berners-Leeの提唱により、W3Cで次世代Web基盤としてSemantic Webの標準化が進められており、我が国でもSemantic Web関連の研究が行われているが、その標準、基幹部分への貢献（すなわち土俵作りへの貢献）は皆無といえる状況である。国際標準（de factoおよびde jure）と関連した基幹部分でのリーダーシップを取ることなく今後の情報技術での国際的優位性、競争力を得ることは難しい事を考えると、これは由々しき事態である。

本研究の目的は、Semantic Webとは異なる側面を持つ（対極的面と相補的面を持つ）Webテキストコンテンツ概念記述言語CDL(Concept Description Language)の申請者らのグループによる国際標準化活動と連動し、そのツール群の研究開発、CDLのSemantic Webとの融合、及びこれらによるWebテキストコンテンツの（深層意味まで立ち入れないが）共通性のある意味計算（セマンティック・コンピューティング）の基盤を確立することが重要である。これによって、Webテキストコンテンツに関する意味計算を可能とする日本発の技術により、次世代Web基盤構築に貢献する。今後の情報技術（IT）分野におけるWebの重要性を考慮した時、我が国の情報技術の国際的競争力強化には、このような次世代Web基盤部分での国際標準化を含む貢献が不可欠であり、他にないこの役割の一端を本研究は担う。

2. 研究の目的

Web上の膨大な量の情報を活用するためには、コンピュータにもWeb情報の表す意味を把握できるようにすることが重要な方向となる。本研究は、Web情報の中で中心的役割を果たしている自然言語テキストに注

目し、その“意味”をコンピュータにも理解できる共通的な形式で表し、Web上の意味計算(Semantic Computing)基盤を構築することを目標にした。“意味”といっても表層から深層に至るまでの階層があるが、ここで対象とするのは広く合意が得られている表層に近い概念レベルであり、これを表すのに共通的な（英語や日本語といった特定の言語に依存しない）概念記述言語としてCDL(Concept Description Language)を採用している。

CDLは我が国で生み出された独自の枠組みであり、NPO法人ISEC（セマンティックコンピューティング研究開発機構；現在は石塚が理事長を務めている）が中心となり、W3CでセマンティックWebとは異なる側面を持つ次世代Web基盤に向けた国際標準化活動を行っている。本研究では主に自然言語テキストからCDLへの半自動変換、CDLデータの類似検索も含む意味的検索のメカニズムについて研究し、意味計算基盤の確立に貢献する。

CDLは単語エンティティ間を関係識別子で結んで表すことをベースにしているが、関連研究としてWebテキストからの関係抽出とその利用による、新しいWebインテリジェンス機能の開発を行う。具体的には、単語エンティティ対間の関係類似性を計算する手法、この原理を活用して{(Tokyo, Japan), (? , France)}のようなエンティティ対を検索入力として、{? = Paris}を答えるような潜在的関係検索エンジン(Latent Relational Search Engine)の研究開発を行う。

3. 研究の方法

最初に本研究でベースにしているCDLについて説明しておく。

CDL(Concept Description Language)は意味計算(Semantic Computing)基盤および次世代Web基盤へ貢献すべく努力している日本発の技術である。Semantic Webがメタデータの共通記述を行うのに対し、CDLは自然言語テキストが表す概念をコンピュータが把握できる形式で共通的に記述する言語である。（CDLは画像、映像なども含むコンテンツの意味記述への配慮もなされているが、それらの意味も多くの場合に自然言語を介して記述されるので、ここでは自然言語テキストを記述対象とする。）即ち、メタ

データには限定せず、Web で情報/知識伝達の主要なソースデータであるテキストが表す意味概念を記述対象にしている。更に、Semantic Web のオントロジー構築が領域毎で問題があるが、CDL の語彙はユニバーサルに定義され、体系的に構成されている。これら語彙を結びつけて複雑な概念、そして文の表す意味概念を表すのに用いる関係子も、特定の言語に依存することなくユニバーサルに定義されている。これによって、言語の壁を越える情報/知識の交流を可能にする。

英語が支配的であるグローバルな Web の情報流通・共有において、言語の壁を克服し多言語による交流を可能にすることは今後の重要な課題である。しかし、英語を主とする国ではこの課題に対して関心が薄く、解決の技術は生まれてこない。日本のような非英語、非印欧系言語の国が主導して、国際貢献を果すべき領域と言える。

言語の壁の克服という第一に機械翻訳を思い浮かべよう。機械翻訳或いはコンピュータによる言語意味理解は着実に進歩しているものの、人間の能力との間のギャップは依然として解消されず、この解消は(永遠に)不可能ではないかとも言われている。CDL では、自然言語テキストの表す概念を十分近似的に表現する適切なレベルを設定し、機械翻訳のようにテキストからの全自動変換を目指すのではなく、人手による(負荷の軽い形態での)援助を含む形でテキストから CDL へ半自動変換する。意味表現といっても深層から表層まで幾つかのレベルがあるが、深層レベルはどのような要素によって表すかについてのコンセンサスがなく、まだ共通化は困難である。CDL で扱うのはテキストの表層表現から僅かに意味に踏み込んだレベル(概念表現レベルと称する)であり、具体的には各語彙要素の意味役割(semantic role)を明示化したレベルと位置付けられる。このレベルは長年の自然言語処理・理解研究により一定程度の共通的理解ができてきており、今日、標準化を図る意義は大きい。

CDL/CWL に先行して UNL (Universal Networking Language) の開発があり、その基になったのは多言語機械翻訳技術の中間言語(ピボット言語)であった。図 1 はこのような経緯を示している。

UNL は 1996 年にスタートした国連大学の UNL プロジェクトのもとで開発された。主

な目的は Web 上での言語の壁を越える情報の交流を可能にすることであり、元になっているのは多言語機械翻訳における中間言語方式(ピボット方式)の中間言語である。

CDL は単にテキストだけでなく、今後の Web において他も含む広いメディア一般が表す概念を記述する汎用で基本的な枠組み、及び Web を中心とした今後の意味計算基盤として設計された。CDL.nl (CDL natural language version) は UNL の成果を CDL の枠組みの上に移し再設計したものであり、自然言語テキストの意味する概念を汎用的に

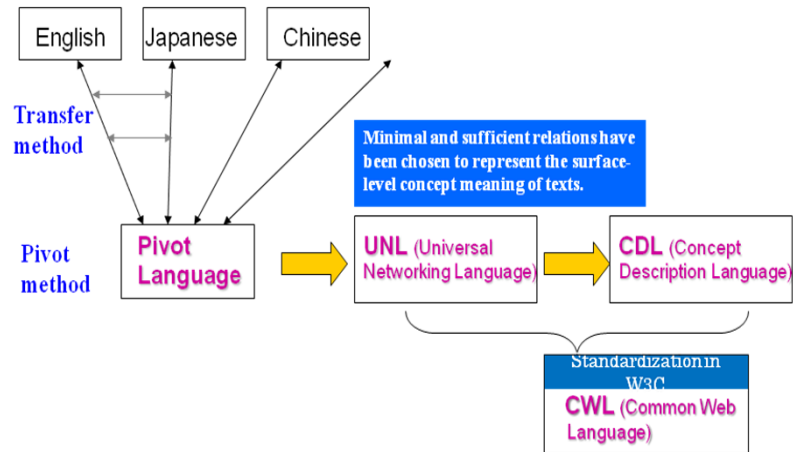


図 1 CDL / CWL 開発の経緯

記述する言語である。(以下では CDL.nl を単に CDL と記す。)

CDL は次世代 Web 基盤として世界に認知され且つ世界標準として貢献する必要性から、2006 年から W3C(World Wide Web Consortium) で標準化活動を開始するのの際し、より分かりやすく受け入れられやすい名称にするために名付けたのが Common Web Language(CWL)である。実は CWL はアンブレラ仕様であり、CWL.unl (UNL 記述に対応)、CWL.cdl (CDL 記述に対応)、CWL.rdf(RDF 表現を採用)を包含している。UNL を利用する活動は世界数カ国で行われており、UNL とその関連機能の意義も高いので CWL の一部に含めている。CWL.rdf は RDF で CDL 記述を表現したものであり、記述自体は長いものになってしまうが、RDF 関連ツール類の利用が可能になるメリットがある。

W3C での標準化はまだ準備段階のインキュベータグループ活動になっており、2006 年 10 月から 2008 年 5 月までの第 1 期で CWL の仕様を定めて公開した。その後の第 2 期で関連ツール類を含めて評価を行った。標準化に際しては Semantic Web 標準との違いを強調するのではなく、メタデータも記述対象の一部と考え、むしろ整合をとるようにし

ている。図2はこのインキュベータグループ憲章(charter)のWeb ページを示している。

UNL はハイパーノード(ノード中にネットワークを含むことが出来る)を含む意味ネットワークの形式で、自然言語テキストが表す意味概念を曖昧性なく表す。ノードは概念を表し、アークは概念間の関係を表す。概念には注釈が付けられる。

UNL の語彙を構成するのが UW (Universal Word) である。UW は概念を表す“単語”であり、他の UW と関係子によってリンクされ複合概念、そして文の意味を表現する。関係子は文を構成する各々の単語 (UW) の主に意味役割を指定することになる。作者が意図した主観的な意味は属性で表される。この主観性(subjectivity)やモデリティも属性として記述することで、UNL は文のニュアンスも含めて表現可能である。

UW は世界中の人々が認識できる概念を正確に表現し、且つそのような概念が誰でも同じ記号をもって定義できるようになっている。意味素には便宜的に英語の単語表現を借用して用いるが(人工的記号にすると人が見て理解出来なくなってしまう)、意味素の表現する概念の範囲を限定し、表現したい概念の範囲を正確に表現するために、UNL の関係子に基づく他の UW からの束縛(意味的包含関係と意味的共起関係)によって規定している。これにより多義性をもつ単語も、多義を分解した一義の語義(word sense)として定義していることが大きな特長である。

例えば”spring”は語義毎に以下のように規定されている。

[名詞概念]

- 1) spring(icl>tool) : バネ, スプリング
- 2) spring(icl>season) : 春
- 3) spring(icl>fountain) : 泉, 湧水

[動詞概念]

- 4) spring(agt>thing, obj>wood) : 反り返らせる
- 5) spring(agt>thing, obj>mine) : 爆発させる

- 6) spring(agt>thing, obj>person, src>prison) : 出獄させる, 釈放させる
 - 7) spring(agt>thing, gol>place) : 跳ぶ, 跳ねる
 - 8) spring(agt>thing, gol>thing) : 飛び越す
 - 9) spring(obj>liquid) : 湧き出る
- 簡単に記号の説明をすると, ”icl”は(includedで) subclass 関係を表し, ”agt”, “obj”, “src”, “gol”はそれぞれ動作主, 対象, 始状態, 終状態の関係子を表している。

CDL 自体は自然言語テキストだけでなく、他にも含む広いメディア一般が表す概念を表

The mission of the Common Web Language Evaluation and Installation Incubator Group, part of the Incubator Activity, is to substantiate the CWL (Common Web Language) in actual web environment using the pilot model of the CWL platform. The CWL is a graphic language of semantic network with hyper node and is used to describe contents and meta-data of web pages in three different type of form such as UNL, CDL and RDF. The CWL platform allows people to input CWL using natural languages and display information written in CWL in natural languages. Using this CWL platform, the CWL will be evaluated from multilingualism, semantic computing and semantic web points of view. Based on these evaluation and feedback, the CWL and its platform will be bearable in actual use in the web.

End date	20 May 2009
Confidentiality	Proceedings are public
Initial Chairs	Hiroshi Uchida (ISeC)
Initiating Members	<ul style="list-style-type: none"> • Institute of Semantic Computing (ISeC) • National Institute of Advanced Industrial Science and Technology (AIST) • Keio University • JustSystems Corporation
Usual Meeting Schedule	Teleconferences: No teleconferences Face-to-face: 2 per year

Scope

The CWL is a graphic language of semantic network with hyper node, a node represents a concept, an arc represents a relation between nodes and a node can be annotated by attributes. This CWL can be expressed in three forms such as UNL, CDL and RDF. The same information in CWL can be described in each form but in different manner. The CWL.unl is a language in UNL form, the CWL.cdl is a language in CDL form and the CWL.rdf is a language in RDF form. Information in the web is basically expressed in natural languages. UNL is for multilingual activities. CDL is for semantic computing activity. RDF is for semantic web activities. Various information will be expressed in three types of representation, and applications based on those representations will be developed, and information will be utilized.

Three different types of representations of CWL allow different way of treatment for the same information described. CWL.unl is unstructured text for

図2 W3C インキュベータグループの憲章(charter)部分

現する汎用で基本的な枠組みとして設計されている。簡素な3つの組表現 (<実体1, 関係, 実体2>, <主語, 述語, 目的語>あるいは<実体, 属性, 属性値>などを表す)を基礎とし、ネットワーク表現だと実体を表すノードと関係を表すアークから成り、ノードは(更には一般にはアークも)ネットワークを含むことが出来る(複合実体となる)のでハイパーノードとなる。関係構造とハイパーノードに対応する入れ子構造の表現を基本にしている。実体と関係には詳細特性記述のために任意個の属性-属性値(attribute-attribute value)を付加できる(関係に準じるのものであり、もし属性値が複合実体となるなら関係とすることができる)。

CDL.nl は UNL の成果を CDL の枠組みの上で再設計した、自然言語テキストの意味する概念を汎用的に表現する概念記述言語である。UNL の成果を活用し、自然言語表現をカ

パーする実用レベルの基本語彙オントロジーと関係子を備えていることも大きな特徴である。Semantic Web におけるメタデータの記述に留まらず、Web の主要情報ソースである自然言語テキストを対象にして共通的な意味概念記述を可能とする。そして次世代 Web 基盤としての国際標準化を視野に入れ、十分な汎用性、国際性を持つことを前提にして設計されている。図 3 は CDL.nl 表現と対応するネットワーク表現を例示している。ここでは that 以下の構文がハイパーノードになっている。

自然言語の意味に関しては表層から深層まで幾つかの階層があるが、CDL で記述対象とする意味概念は、表層表現から僅かに意味に踏み込んだ一般性がある概念レベルとしている。意味表現という深層意味をとらえて、正しく理解できないだろう、共通的な記述法は無理ではないかといった質問をよく受けるが、深層意味を記述の対象にしている訳ではない。

文章は複数文から構成されるが、文章の複合実体は複数の文の複合実体をノードとして含み、それらのノード間は文間の接続構造を表す接続関係か、指示表現や代用関係等を表す参照関係で結ばれる。

基本語彙(記述に必要な関係子語彙セットを含む)は UNL の成果を受け継ぐ形で構成されているが、これが充実していることが大きな特徴であることから、ここで全体を示すことは出来ないものの、その一端を示すことにする。

CDL の語彙を分ける最上位オントロジーは、

- Entity,
- Relation,
- Attribute

である。Entity の下位オントロジーには Thing (事物) と CompositeEntity (複合実体) がくる。Thing の下位に以下のカテゴリが置かれる。

- NominalThing (抽象物, 具体物, 機能,

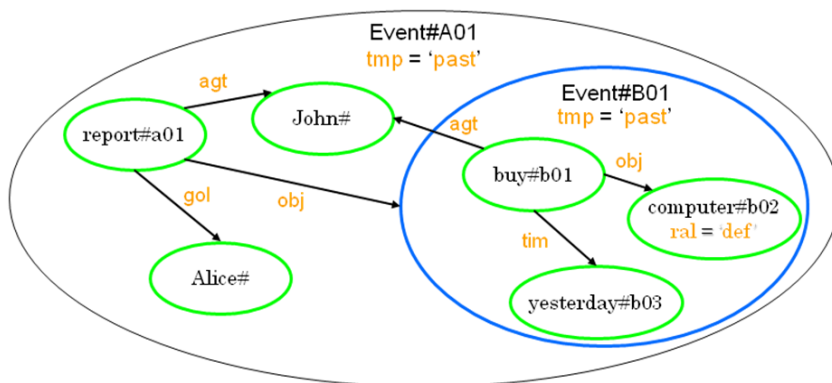
有意志体, 場所, 代名詞を含む)

- NominalModifier (限定詞と形容詞)
- VerbalThing (Do, Occur, Be 関係の動詞, 述語)
- VerbalModifier (数量と副詞)
- ValueOfAttribute (属性値)

これらカテゴリの下に現在約 20 万語が UNL の UW の形で登録されている。

Relation (関係) については、自然言語で表される概念をどのような観点からとらえるかを示す意味で重要であるので、表 1 にそ

<John reported to Alice that he bought a computer yesterday.>



```

{#A01 Event tmp='past';
  {#B01 Event tmp='past';
    <#b01:buy;>
    <#b02:computer ral='def';>
    <#b03:yesterday;>
    [#b01 agt #John]
    [#b01 obj #b02]
    [#b01 tim #b03]
  }
  <#John:John;>
  <#Alice:Alice;>
  <#a01:report;>
  [#a01 agt #John]
  [#a01 gol #Alice]
  [#a01 obj #B01]
}

```

図 3 テキスト(上段), グラフ記述(中段)と CDL.nl 記述(下段)の例

の関係子語彙を提示する。

意味役割ラベルの例は他でも見られ、例えば英語については PropBank で用いられている意味役割ラベルがある。しかしこれは英語動詞毎に役割を変えて規定する部分があり、言語独立のユニバーサル性はない。また、意味役割だけでは文の要素である実体概念間の関係を記述することは出来ず、論理関係、概念関係(上位や具体化など)、接続関係、参照関係も必要になり、表 1 のように CDL はこれらの関係子も含んでいる。CDL の関係

子が本当に必要最小限かを証明することは難しいが、多言語機械翻訳の中間言語、UNLによる記述を通じて実践的に十分であると実証されてきている。

ノード(実体概念)に付加される属性及び属性値は文の命題部分だけではなく、モダリティ部分も十分な精度で表現できるようになっている。他にこのように充実、整備された表現機能を有するシステムは存在しない。

表1 CDLの関係子語彙

Time (with respect to writer) (時制)	past present future
Aspect (view on aspect of event) (事象の相)	begin complete continue custom end experience progress repeat state
View of emphasis, focus and topic (強調, フォーカス, 話題)	contrast emphasis entry qfocus theme title topic
Attitudes (modality) (態度)	affirmative confirmation exclamation imperative interrogative invitation politeness respect vocative
Feeling and judgements (感情と判断)	ability get-benefit give-benefit conclusion consequence sufficient grant grant-not although discontented expectation wish insistence intention want will need obligation obligation-not should unavoidable certain inevitable may possible probable rare regret unreal admire blame contempt regret surprised troublesome
View of reference (参照ビュー)	generic def indef not ordinal
Logicity (論理的性質)	transitive symmetric identifiable disjoint
Modifying attribute on aspect (相の修正)	just soon yet not
Attribute for convention (記号表現規約など)	passive pl angle_bracket brace double_parenthesis double_quote parenthesis single_quote square_bracket

以上で分かるように、CDLは記述形式を定めているだけでなく、様々な自然言語テキスト表層表現の意味を、高い近似度の概念レベルで表現するのに必要な関係子語彙と属性を規定して提供している。これによってコンピュータは、曖昧性なく意味概念を把握できることになる。より深い意味理解や推論機能等は外部プログラムに委ねられる。すなわち、自然言語テキストは一旦CDL.nlの概念記述に変換され、これを共通の基盤として幅広い意味計算(セマンティックコンピューティング)の展開が可能になる。

CDL自体は個別の言語に依存せず、汎用的に自然言語の意味概念表現に使えるもので

あるが、共通である関係子語彙(機能語)を除く語彙オントロジーを各国語対応に用意することにより、各国語対応バージョンを作成できる枠組みとなっている(日本語版のCDL.jp, 中国語版のCDL.chなど)。各国語の語彙は基本とするUNLのUW(Universal Word)に対応付けることにより、CDL.nlは言語の壁を超える情報/知の交流の基盤ともなる。

本研究ではこのCDLをベースにして、次世代Webの意味計算基盤として広く世界に受け入れられるために必要な機能の研究開発を行った。同時に、CDLの基礎である関係記述に関連して、単語エンティティ間の潜在的な意味の関係に関する研究も進めた。

4. 研究成果

(4-1)テキストからCDLへの半自動変換

CDLの現状での一つの課題は、その記述生成法である。自然言語テキストからの全自動変換は理想ではあるが、正しい機械翻訳と同様に困難であり、何らかの人手による援助が必要となる。この人の労力は出来るだけ負担が軽いものである必要がある。

UNLシステム内で英文テキストからUNL記述に変換するのに基本的にルールベースの翻訳機構を用いているが、このルールセットは通常用途のものに加えて、単語語義曖昧性を解消し一義になった語彙を対象とする翻訳ルールセットも備えている。後者のルールセットを用いることが出来れば、テキストからCDLへの変換の精度を大幅に高めることができる。そこで、まず多義で語義曖昧性を有する英語単語に対して、

人手による語義選択、確定を行うユーザインタフェースを開発した。全ての多義を持つ単語に対して人手による語義選択を要請するのは煩雑過ぎ、前後の文脈から語彙を決定できる場合も多いので(例えば名詞のbankはお金に関する文脈中では銀行に相当する語彙(UWの表現でbank(inl>organization))になる)、自動的に単語語義曖昧性解消の可能な部分を増やすようにしている。図4はこのシステム構成並びに半自動変換インタフェースの様子を示している。日本語かな漢字入力変換でも完全自動化は困難だが、今日は誰でもが"かな"か

ら変換の漢字候補中から人手選択を行うことにより、人に無理のない形でかな漢字入力変換が実現されている。このアプローチの目指すところは、かな漢字入力変換のように軽負担で人手が関与する形態でのテキストか

(Support Vector Machine 等の手法) を使って識別を行うことが難しい問題がある。

(4-2) CDLデータの意味的検索

CDL記述したデータはコンピュータも意味が把握出来ることにより、コンピュータも動

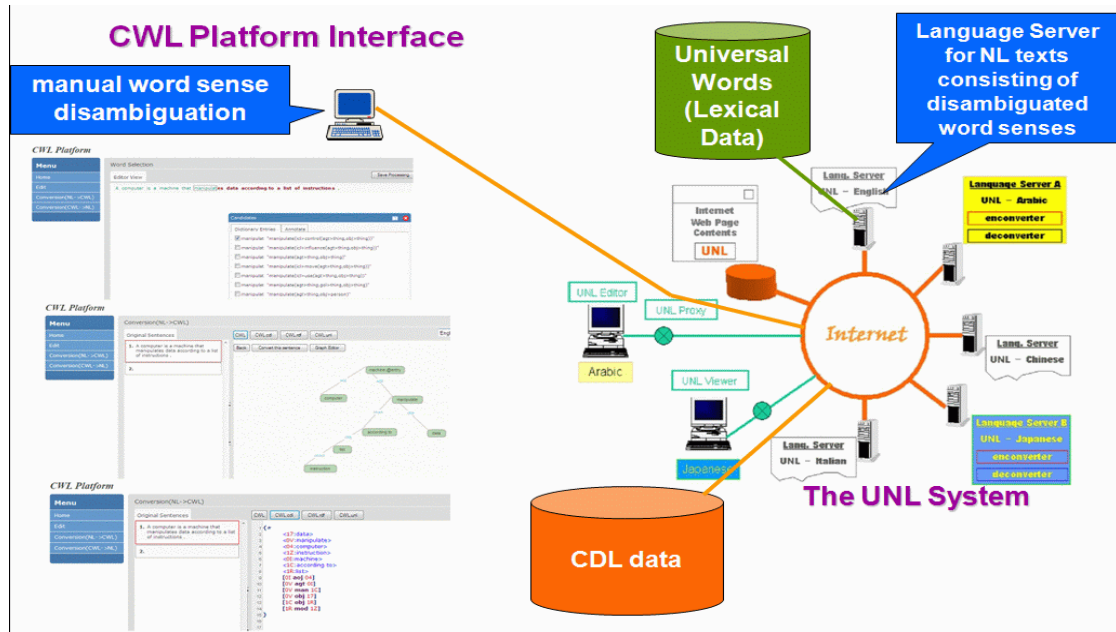
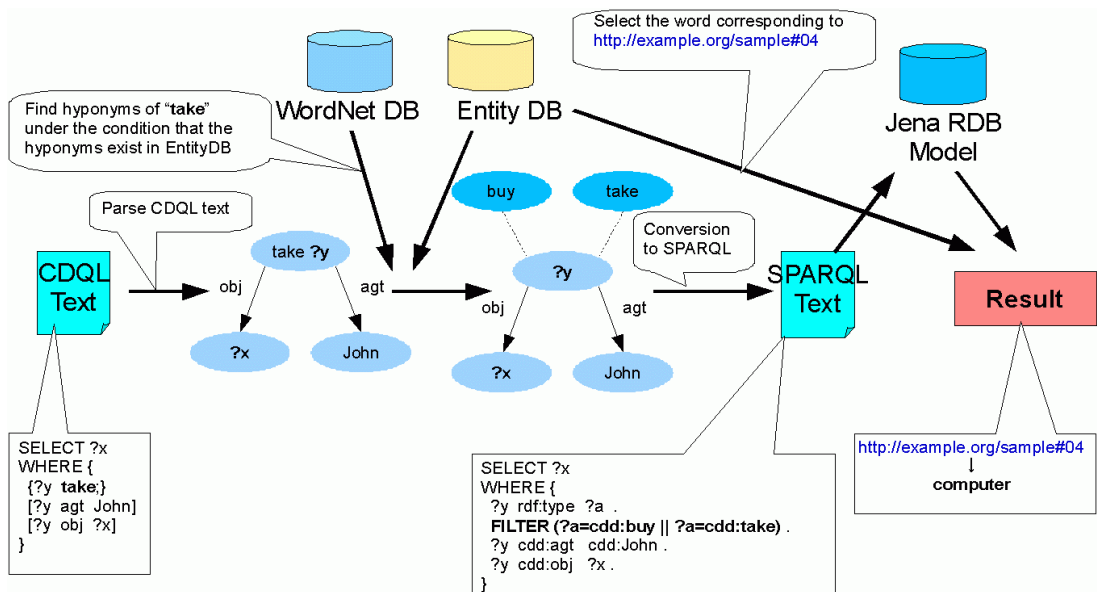


図4 単語語義選択によるテキストから CDL への半自動変換 (右部は半自動変換のグラフィカルインターフェース画面)

らの CDL への変換である。

係り受け解析とまた十分ではない意味役割解析のツールを利用して、文中のエンティティ間に存在する CDL 関係を自動抽出する研究も行った。しかしこれは十分な性能になるまでにはまだ研究が必要である。特に CDL で記述されたデータ (コーパス) がまだ十分な量でないことから、出現頻度が低い CDL 関係に関して機械学習

員する広範な意味計算の基盤になるのだが、最も直接的な効用は意味的検索ということになる。現在のキーワード照合を基本とする検索を超えて、まず実体 (entities) 間の関係を指定した検索が可能になる。これは問い合わせもグラフで表すことにより、一種のグラフ照合によって実現できることになるが、まず完全な照合がとれるパターンを検索し、それが得られないときは単語を下位語、上位語、



Query: What did John take?

図5 CDLデータの意味的検索システムの構成

関係語に置き換えて探索することにより、意味的類似性に基づく検索が実現される。図5はこのような考えで実現したCDLデータの検索システムのプロトタイプを示している。ここでの問い合わせ言語はSQL-likeなCDQL

(Concept Description Query Language)であり、処理系の内部ではRDFの問い合わせ言語として標準化されているSPARQLを用いている。SPARQLはRDFデータを関係データベース(Jena RDB)で管理しており、大規模データにも対応できる機能を備えていることから、利用した。

意味的類似性に基づく検索の次の段階では、「著者は本を書いた人」といった一般知識があれば、「石塚満は“知識の表現”の本を書いた。」というテキストとCDL記述があれば、「“知識の表現”の本の著者は誰か？」という問い合わせに対し、“石塚満”と答えることが出来るようになる」と期待できる。

(4-3)エンティティペア間の関係類似性計算法

関係抽出や関係情報検索でエンティティ間の関係類似度の計測法は重要な要素技術である。属性類似性(attributional similarity)は、各エンティティ(単語)に付随する属性の類似性によって計測される。例えば、carとautomobileは多くの属性(例えば、車輪を有する、輸送に使われるなど)を共有することから高い類似性を有し、類義語(synonyms)と判定される。これに対し、ここでの関係類似性(relational similarity)は2つのエンティティペア間に存在する意味的関係の対応に注目する。例えば、(ostrich, bird)と(lion, cat)の2つの単語ペアを考えると、“Ostrich is a large bird”と“Lion is a large cat”のような関係が存在することにより、これら2単語ペア間の関係類似性は高いと判定される。ここでは、大量に存在するWebテキスト情報をWeb検索エンジンを通して利用し、このような関係類似度を計測する効果的な方法を考案、開発している。これにより単に一般辞書に記された事項間だけでなく、(Google, YouTube)と(Adobe Systems, Macromedia)は買収関係の点で類似であり、(Eric Schmidt, Google)と(Steve Ballmer, Microsoft)はそれぞれのCEOであるという関係の点で類似性が高いといったような判定も可能になる。関係類似性は人工知能分野では従来から類推(analogy)として研究されてきたが、Webのような大量情報を利用する点で、本研究の新規性がある。

ここで類似性計算のベースとするのは、共起する単語ペア周辺に現れる語彙系列パターンが類似していれば関係類似性は高いとする、分布仮説(distributional hypothesis)である。問題は、1)如何に判定に効果的ななるそのデータを如何に効率的に処理して

類似度を計算するかとなる。後者に関する関連研究としては、P. TurneyによるLSA

(Latent Semantic Analysis: 2次元行列のSVD(Singular Value Decomposition)を用いる)による方法があるが、我々はそれとは異なる大量データを扱う際に効率的な手法を考案した。

1)に関しては、概ね以下の手順で抽出を行う。単語ペア(X, Y)に対して、検索エンジンでX&YのAND検索を行い、XとYが共起する検索結果のスニペットよりXとYの間及び前後の長さL語以内の周辺語彙パターンをG語以内のギャップを許容して抽出する。

(細かい点としては、数字はすべてDとして同一に扱う、Inc., Ltd., Corp., Mr., Ms., Prof., Dr. 付き固有名詞は付随の固有名詞と同一視する、didn'tなどの省略形はdid notに展開するなどあり。)詳細はここでは省略するが、prefixspanアルゴリズムの変形版により、出現頻度がN以上となる周辺語彙パターンを効率的に選び出し、特徴要素とする。

2)に関しては、以上のようにして抽出される語彙系列パターンの特徴要素は非常に多数(数万個)にもなること、特徴要素間には同様な意味的關係を表す複数の表層的变化が多数存在すること、また特徴要素間には相関が存在することを考慮し、以下の方法で単語ペア間の類似性を計算する。まず、同様な意味的關係を表す語彙系列パターンを創案したクラスタリングによりまとめる。例えば、買収関係については“X buys Y.”、“X acquires Y.”、“X completed the acquisition of Y.”などは同様な意味的關係を表している。この新規のクラスタリングはまず、語彙系列パターンを出現度数によりソーティングする。次いで、出現度数が多いパターン順にそれが出現する単語ペアの分布形の近さ(即ち分布仮説に基づき)で比較し、既に形成されたクラスタとの距離を計り、スレッシュド以下であったら既存クラスタへまとめる処理を行う。これは貪欲(greedy)な線形時間処理であり、数万の要素に対しても非常に高速処理が可能となる。クラスタ化され圧縮された特徴要素間でも無相関ではなく相関が存在するので、単語ペア間の類似性は特徴間の相関を考慮するマハラノビス距離によって算出する。LSAによる方法では特徴要素数が数万になると禁止的な計算時間になるが、本計算法では数分の計算時間となる。

本手法の適用例として、表2の関係を有する計100単語ペアについて各々類似性を計算し、類似度に基づいてクラスタリングにより関係を判別する実験を行った。この時、周辺語彙パターン抽出のパラメータはL=5, G=2であり、各単語ペアについて1,000個のスニペット、計100,000スニペットより約47万のパターンが抽出されるが、このうち2度以

上出現する約 15 万パターン (約 31%) を初期特徴パターンとして使用する. 上記特徴パターンのクラスタリングを行った結果の要素数が大きい順に上位 10 個のクラスタを示

との関係は「Paris」と「France」との関係に類似するからである(東京が日本の首都, パリもフランスの首都などの関係があるからである). 潜在関係検索のアイディアはアナロジ

表 2 関係類似性計算評価用のデータセット

Relation Type	Total contexts	Examples (20 in all for each relation type)
ACQUIRER-ACQUIREE	91439	(Google, YouTube), (Adobe Systems, Macromedia), (Yahoo, Inktomi)
PERSON-BIRTHPLACE	72836	(Franz Kafka, Prague), (Charlie Chaplin, London), (Marie Antoinette, Vienna)
CEO-COMPANY	82682	(Terry Semel, Yahoo), (Eric Schmidt, Google), (Steve Jobs, Apple)
COMPANY-HEADQUARTERS	100887	(Microsoft, Redmond), (Yahoo, Sunnyvale), (Google, Mountain View)
PERSON-FIELD	99660	(Albert Einstein, Physics), (Roger Federer, Tennis), (Shane Warne, Cricket)

表 3 上位 10 クラスタの最頻語彙系列パターン

cluster 1 (2868)	X acquires Y	X has acquired Y	X's Y acquisition	X, acquisition, Y	Y goes X
cluster 2 (2711)	Y legend X was	X's championship Y	Y star X was	X autographed Y ball	Y star X robbed
cluster 3 (2615)	Y champion X	world Y champion X	X teaches Y	X's greatest Y	Y players like X
cluster 4 (2008)	X to buy Y	X and Y confirmed	X buy Y is	Y purchase to boost X	X is buying Y
cluster 5 (2002)	Y founder X	Y founder and CEO X	X, founder of Y	X says Y	X talks up Y
cluster 6 (1364)	X revolutionized Y	X professor of Y	in Y since X	ago, X revolutionized Y	X's contribution to Y
cluster 7 (845)	X and modern Y	genius: X and modern Y	Y in DDDD, X was	on Y by X	X's lectures on Y
cluster 8 (280)	X headquarters in Y	X offices in Y	past X offices in Y	the X conference in Y	X headquarters in Y on
cluster 9 (144)	X's childhood in Y	X's birth in Y	Y born X	Y born X introduced the	sobbing X left Y to
cluster 10 (49)	X headquarters in Y	X's Y headquarters	Y - based X	X works with the Y	Y office of X

したのが表 3 である.

性能評価のために表 2 のデータセットについて, 5 重交差検定により k 最近傍識別法 (ここでは k=10) による関係識別実験を行い (4/5 のデータセットでクラスタリングによる圧縮特徴パターン空間を学習し, 残りの 1/51 のデータにより性能を測定する), 既存手法より高い性能が達成できることを実証した.

(4-4) 潜在関係検索エンジンの開発

Web などのテキストコーパスの中に, エンティティ間の関係が多数潜在的に記述されている. 従来のキーワードベース検索エンジンは単語ペア間の関係を利用して検索することが出来なく, 潜在的に存在する多くの関係情報を積極的に利用できない. 本研究では, 単語間の潜在的な関係を積極的に利用し, 類似関係による潜在関係検索エンジンを開発した (2009 年 12 月に特許出願済).

潜在関係検索とは, 単語ペア間の潜在的な関係を利用することにより, 入力単語ペアと類似する単語ペアを検索する新しい検索パラダイムである. 潜在関係検索エンジンの概要は図 6 に示している. クエリー「(Tokyo, Japan), (?, France)」が入力されたときに, 「Paris」を最初にランキングされた結果リストを返す. その理由は「Tokyo」と「Japan」

ー・シソーラスの研究や我々グループによる関係類似度計算法の研究で検討されてきた. 本研究では(4-3)の関係類似性計算法の研究に基づき, 検索速度も考慮に入れた最初の潜在関係検索エンジンを実現している. 先行して, 既存のキーワード・ベース検索エンジンを利用して潜在関係検索を実現したが, 単語間における関係を十分に抽出できず, 精度や平均逆順位(MRR)がまだ低い. また, クエリー処理時にキーワード・ベース検索エンジンに数十個のクエリーを投げているので速度が遅い. 本研究では関係抽出の手法を使い, 潜在関係検索エンジン用のインデックスを作成し, 高精度の関係類似度測定の研究成果を応用し, 高速かつ精度の高い潜在関係検索の実現を図っている.

潜在関係検索を実現するためには, エンティティ(単語)ペアと, ペアの二つのエンティティ間の関係を抽出することと, 抽出されたペアの関係類似度を測定する必要がある.

(i) エンティティ・ペア抽出と関係の表現

本システムはWWWなどのテキストコーパスからエンティティ・ペアとペア中の単語間関係を抽出し, 検索用インデックス・データベースを作成する. まず, テキスト・ドキュメントから文を抽出し, 得られた文を単語に分節し, 品詞タグを付ける(そのとき, named entityも同時に認識する). 例えば, 「東京は日本の首都である」という文は, 「東京/固有名詞; は/助詞; 日本/固有名詞; の/助詞; 首都/名詞; である/動詞」の列に分けられる. 分

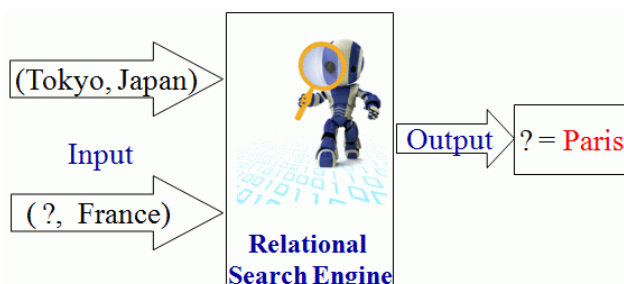


図 6 潜在的関係検索エンジンの概要

かられた列中の名詞，固有名詞を抽出し，ペアを形成する(現在の実装では固有名詞しか抽出していない). 上記の文からは，(東京，日本)のペアが抽出される. 実質に関係を持つペアは普通出現頻度が高く，偶然に共起するペアはあまり頻度が高くない. そこで，ある頻度以上のペアだけを検索対象とする(現在の実装では頻度5以上). また，文中の距離が遠ければあまり関係を持たない可能性が高いので，距離がある閾値 D_{max} 以上のペアは検索対象としなく関係を抽出しない.

単語ペアを抽出すると同時に，ペアの関係を表す語彙パターンも抽出する. 語彙パターンはペアの出現位置の周辺語彙列を取り，列n-gramとして抽出する. 例えば，上記の文では，(東京，日本)ペアの語彙パターン集合は{ X は Y の首都， X は Y の首都である， X は

Y *首都}を含む(「*」はワイルドカードで0個以上の語彙を意味する). 単語ペア S_j と語彙パターン pi が一緒に出現する頻度を $h(S_j; pi)$ とする. そのとき，単語ペア S_j の特徴ベクトルの要素は i の値を $h(S_j; pi)$ として定義する. また，語彙パターン pi の特徴ベクトルの要素 j を $h(S_j; pi)$ とする. 抽出された単語ペアと語彙パターンやその頻度をデータベースに保存する.

(ii) 関係類似度計算とクエリー処理

入力されたクエリー(A, B), (C, ?)に対して，まず，ペア(A, B)をデータベースから取得し，ペアの関係を表す語彙パターンを取り出す. ペアの関係を表すパターンを含む他の(C, X)ペアを検索結果の候補ペアとする(つまり，候補ペアの第一要素がCで，入力ペアと一個以上の語彙パターンを共有する). (A, B)ペアの特徴ベクトル(要素 i が $h(A; B; pi)$ のベクトル)と(C, X)ペアの特徴ベクトルをデータベースから取り出し，ペアの関係類似度を特徴ベクトルのコサイン類似度により計算する. (A, B)ペアと(C, X)ペアの正確にマッチする語彙パターンは少ないので，上記の語彙パターンの特徴ベクトルを使い，類似するパターンをクラスタリングする(例えば，“X is CEO of Y”，“Y’s CEO X”，“Y chief executive X”を同じクラスタにまとめる). 同じクラス

Query:

Word pair 1:	<input type="text" value="steve jobs"/>	<input type="text" value="apple"/>
Word pair 2:	<input type="text" value="?"/>	<input type="text" value="microsoft"/>
<input type="button" value="Search"/>		

steve jobs is to apple as:

- **['ballmer', 'steve ballmer']** is to **microsoft** (Score = 295) [Show evidence](#)
- **['bill gates']** is to **microsoft** (Score = 52) [Show evidence](#)
- **['danny thorpe']** is to **microsoft** (Score = 27) [Show evidence](#)

図7 関係検索クエリーと検索結果画面

Query:

Word pair 1:	<input type="text" value="steve jobs"/>	<input type="text" value="apple"/>
Word pair 2:	<input type="text" value="?"/>	<input type="text" value="microsoft"/>
<input type="button" value="Search"/>		

steve jobs is to apple as:

- **['ballmer', 'steve ballmer']** is to **microsoft** (Score = 295) [Hide evidence](#)
 - Steve Jobs is the CEO of Apple, which he co-founded in 1976. <http://www.apple.com/pr/bios/jobs.html>
 - Steve Jobs is the CEO of Apple, which he co-founded in 1976. <http://jobsearchtech.about.com/od/historyoftechindustry/a/SteveJobs.htm>
 - But nothing as mundane would prompt Steve Jobs, Chief Executive, Apple, to predict that cities will be built around it. <http://www.rediff.com/netguide/2001/dec/03ginger.htm>
 - Bangalore: Steve Jobs, Chief Executive, Apple has admitted that a mechanism exists within the iPhone that enables the company to unilaterally remove software from the iPhone. http://www.siliconindia.com/shownews/Steve_Jobs_confirms_application_-_nid-45489.html
 - In a abbreviated broadcast, SEGuru reports on the announcement that Steve Ballmer, CEO of Microsoft will deliver the keynote at SMX West 2010 in Santa Clara. <http://www2.webmasterradio.fm/seo-rockstars/2009/microsoft-ceo-steve-ballmer-to-keynote-smx-west/>
 - Steve Ballmer is the CEO of Microsoft. <http://qna.rediff.com/questions-and-answers/who-is-the-ceo-of-microsoft/8431179/answers>
 - Steve Ballmer is the CEO of Microsoft and one of the richest men in the world -LRB- \$ 25 Billion -RRB- so I guess he's allowed to do pretty much what he wants including looking like a fat sweaty dancing monkey. <http://www.sonicbomb.com/modules.php?name=News&file=article&sid=32&mode=thread&order=0&thold=0>
 - The news broke after a meeting in Seattle on Saturday at which Steve Ballmer, chief executive of Microsoft, sought to persuade Jerry Yang, co-founder of Yahoo, to yield to the software group by raising his offer from \$ 31 a share to \$ 33. http://business.timesonline.co.uk/toI/business/markets/mergers_and_acquisitions/article3872866.ece

図8 潜在関係検索結果の根拠となったセンテンスを示す画面

タに入るパターンは同じパターンとして扱い、ペア間のコサイン類似度を計算する。

また、エンティティの複数表現形を吸収するために、似ているエンティティをクラスタリングする必要がある(例えば, “United States”, “U.S.”, “US” は同じクラスタに入るようにする)。そのために、エンティティとのペアをなすエンティティを次元とするエンティティの特徴ベクトルを定義する。エンティティの類似度はエンティティの特徴ベクトルのコサイン類似度として定義する。類似度の高いエンティティを1つのクラスタにまとめ、出力結果で同じクラスタのエンティティであれば、同じグループとして出力する。

現在Wikipediaテキストから名詞単語ペアの上記のような検索用インデックスを作成している。従って、カバー率は必ずしも多くなく、このインデックスに含まれない単語ペアの検索の場合はon-the-flyで検索エンジンを用いての潜在関係検索を行うモードになり、時間がかかることになる。

図7は「(steve jobs, apple), (?, microsoft)」の関係検索クエリーに対して、適合度指標付きで3つの結果が返される画面を示している。この “show evidence” をクリックすると、その根拠になっているセンテンスが示され、図8はその例である。

(4-5)エンティティ間関係の共クラスタリングによる抽出法

これは(4-3)の関係類似性計算法の拡張として得た研究成果であり、Web から抽出した多数の任意エンティティペアから有意な関係を有するエンティティペアの集合を、新しい共クラスタリング法により抽出する方法である。このような問題設定は Open Information Extraction(Open IE)と呼ばれており、エンティティペア間の関係についての Open IE の効率的で有効な方法の創案となっている。Open IE に近い関係抽出の方法に半教師付き(semi-supervised)アプローチもあり、これは少数の種(seeds)を与えて、この集合を順次拡張していくのだが、これは初期集合(即ち種)への依存性が強いという問題がある。これに対し、ここでの共クラスタリングによる方法では、可能な関係種別とその数は事前に与えられず、自己学習的に存在する関係を見出すことが大きな特徴となっている。

この概念的原理は関係双対性(relational duality)にあると言える。即ち、ある関係は外延的定義(extensional definition)と内包的定義(intentional definition)によって別々に規定することが出来る。前者はある関係を持つ事例を網羅的に列挙することにより規定する方法であるのに対し、後者はその関係を定めるのに必要な性質、条件を規定する方法である。エンティティペア間の関係について考えると、ある関係のエンティティペアを列挙する

のが外延的定義となり、それらの周辺語彙系列パターンによって規定するのが内包的定義となる。例えば、買収関係について述べると、(Google, YouTube), (Microsoft, Powerset), (Adobe Systems, Macromedia), (Yahoo, Inktomi)..... などが外延定義となり、X bought Y, X acquired Y, Y is acquired by X, Y is purchased by X, X completed its acquisition of Y, などが内包的定義に相当することになる。

方法としては具体的なエンティティペアを縦軸、語彙系列パターンを横軸にした2次元マトリックスを作り、各々の共起回数を各マトリックス要素にすることを考える。この場合、縦軸のエンティティペアの関係類似性は横軸の分布形の類似性から判断でき、また、横軸の語彙系列パターンの関係類似性は縦軸の分布形の類似性から判断できることになる。即ち、分布仮説が相互に適用可能であり、これを我々は関係双対性と呼んだ。これにより、縦軸、横軸を相互にクラスタリングする共クラスタリングが実現される。これは意味的な同一性を保ちつつ、語彙系列パターンや名称のばらつきを吸収し、特徴空間を圧縮して安定な分類、識別を達成する上で有効となる。

具体的には(4-3)の関係類似性計算法で横軸(語彙系列パターン)のみについて適用していた greedy で効率的なクラスタリングを、横軸、縦軸交互に適用する共クラスタリング法として構成する。(なお、語彙系列パターンだけでなく、横軸には語の品詞(POS タグ)系列パターンも含め扱うようにすることにより、パターンの汎化効果が現れることになり、そのようにしている。)共クラスタリングの結果として得られる圧縮されたマトリックスの縦軸は、類似の関係を有するエンティティペア・クラスタであり、横軸は関係識別の観点で類似の特徴をまとめたクラスタとなる。本研究では形成されたクラスタについて、代表例を見出して関係種別のレベルを付与する機能も実現している。また、クラスタリングにおける適切なスレッシュホールド設定の目安を与える方法も与えている。

この共クラスタリング法の効果を実証するため、表2と同じ5種の間関係を有するエンティティペア(5×100=500個)と検索エンジン出カスニペットから得られる語彙系列パターンのデータセットを用いた関係類似性による分類問題、人間関係ネットワーク Web から得られる関係を持つ人物ペア(50,000ペア)から代表的で有意な関係種別の自動抽出問題(Open IE 問題)の実験を行い、既存手法を凌ぐ性能を実証している。

本技術は他のパターン分類、マイニング問題にも有効であり、2010年4月に特許申請を行った。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 13 件)

- [1] 榊 剛史, 松尾豊, 内山幸樹, 石塚 満: Web 上の情報を用いた関連語のシソーラス構築について, 自然言語処理, Vol.14, No.2, pp.3-31 (2007.4)
- [2] Yutaka Matsuo, Junichiro Mori, Masahiro Hamasaki, Takuichi Nishimura, Hideaki Takeda, Koiti Hasida and Mitsuru Ishizuka: POLYPHONET: An Advanced Social Network Extraction System from the Web, Journal of Web Semantics, Volume 5, Issue 4, pp.262-278 (2007)
- [3] Helmut Prendinger, Paul Piwek and Mitsuru Ishizuka: A Novel Method for Automatically Generating Multi-modal Dialogue from Text, Int'l Journal of Semantic Computing, Vol.1, No.3, pp. 319-334 (2007)
- [4] 金 英子, 松尾 豊, 石塚 満: Web 上の情報を用いた弱い社会的関係のネットワーク抽出法, 電子情報通信学会論文 D, Vol.J91-D, No.3, pp.709-722 (2008.3)
- [5] Yingzi Jin, Mitsuru Ishizuka and Yutaka Matsuo: Extracting Inter-firm Networks from the World Wide Web using a General-purpose Search Engine, Online Information Review, Vol.32, Issue 2, pp.196-210 (2008.5)
- [6] 唐門 準, 松尾豊, 石塚 満: リンクに基づく分類のためのネットワーク構造を用いた属性生成, 情報処理学会論文誌, Vo.49, No.6, pp.2212-2223 (2008.6)
- [7] Mostafa Al Masum Shaikh, Helmut Prendinger and Mitsuru Ishizuka: Sentiment Assessment of Text by Analyzing Linguistic Features and Contextual Valence Assignment, Applied Artificial Intelligence, Vol.22, Issue 6, pp.558-601 (2008.7)
- [8] Werner Breitfuss, Helmut Prendinger and Mitsuru Ishizuka: Automatic Generation of Gaze and Gestures for Dialogues between Embodied Conversational Agents, Int'l Jour. of Semantic Computing, Vol.2, No.1, pp.71-90 (2008)
- [9] 古川忠延, 松尾 豊, 大向一輝, 内山幸樹, 石塚 満: ブログ上での話題伝播に注目した重要語判別, 知能と情報 (日本知能情報ファジィ学会誌), Vol.21, No.4, pp.557-566 (2009.8)

[10] 石塚 満, 内田裕士, 横井俊夫: 自然言語テキスト意味概念の共通的記述による次世代 Web 基盤, 知能と情報 (日本知能情報ファジィ学会誌), Vol.21, No.4, pp.519-526 (2009.8)

[11] Arturo Nakasone, Helmut Prendinger and Mitsuru Ishizuka: ISRST: Generating Interesting Multimedia Stories on the Web, Int'l Jour. of Applied Artificial Intelligence, Vol.23, No.7, pp 633-679 (2009.9)

[12] Danushka Bollegala, Naoaki Okazaki and Mitsuru Ishizuka: A bottom up Approach to Sentence Ordering for Multi-document Summarization, Information Processing and Management, Vol.46, Issue 1, pp.89-109 (2010.1)

[13] Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka: EmoHeart: Conveying Emotions in Second Life Based on Affect Sensing from Text, In: Advances in Human-Computer Interaction (open access journal from Hindawi Pub. Co.), Special Issue on Emotion-Aware Natural Interaction, Vol.2010, Article ID 209801, 13 pages (2010)

[学会発表(国際会議分)] (計 66 件)

[1] Dat P.T. Nguyen, Yutaka Matsuo and Mitsuru Ishizuka: Subtree Mining for Relation Extraction from Wikipedia, Companion Volume of Proc. of the main conf. of Human Language Technologies 2007: The Conf. of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2007), pp.125-128, Rochester, New York (2007.4)

[2] Danushka Bollegala, Yutaka Matsuo and Mitsuru Ishizuka: An Integrated Approach to Measuring Semantic Similarity between Words Using Information Available on the Web, Companion Volume of Proc. of the main conf. of Human Language Technologies 2007: The Conf. of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2007), pp.340-347, Rochester, New York (2007.4)

[3] Danushka Bollegala, Yutaka Matsuo and Mitsuru Ishizuka: An Integrated Approach to Measuring Semantic Similarity between Words Using Information Available on the Web, Companion Volume of Proc. of the main

- conf. of Human Language Technologies 2007: The Conf. of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2007), pp.340-347, Rochester, New York (2007.4)
- [4] Yingzi Jin, Yutaka Matsuo and Mitsuru Ishizuka: Extracting Social Networks among Various Entities on the Web, *The Semantic Web: Research and Applications* (E. Franconi, M. Kifer and W. May (eds.)) (Proc. 4th European Semantic Web Conf., ESWC2007, Innsbruck, Austria), Springer LNCS 4519, pp.251-266 (2007.6)
- [5] Arturo Nakasone and Mitsuru Ishizuka: ISRST: An Interest based Storytelling Model using Rhetorical Relations, Proc. 2nd Int'l Conf. of E-Learning and Games (Edutainment 2007), pp.324-335, Hong Kong (2007.6)
- [6] Yingzi Jin, Yutaka Matsuo and Mitsuru Ishizuka: Extracting Inter-Firm Networks from World Wide Web, Proc. Int'l Workshop on Web Mining for E-commerce and E-services (WMEE2007) (as a part of IEEE CEC/EEE2007), pp.635-642, Tokyo (2007.7)
- [7] Tadanobu Furukawa, Yutaka Matsuo, Ikki Ohmukai, Koki Uchiyama and Mitsuru Ishizuka: Social Networks and Reading Behavior in Blogosphere, Proc. 22nd Conf. on Artificial Intelligence (AAAI-07), pp. 1353-1358 (2007.7)
- [8] Nguyen Dat, Yutaka Matsuo and Mitsuru Ishizuka: Relation Extraction from Wikipedia Using Subtree Mining, Proc. 22nd Conf. on Artificial Intelligence (AAAI-07), pp.1414-1420 (2007.7)
- [9] Hua Wang, Jie Yang, Mark Chinell and Mitsuru Ishizuka: Character Agents in E-Learning Interface Using Multimodal Real-Time Interaction, *Human-Computer Interaction* (J. A. Jacko (Ed.)) (Proc. 12th HCI International 2007, Beijing, China), Part III, Springer LNCS 4552, pp.225-231 (2007.7)
- [10] Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka: Recognition of Affect Conveyed by Text Messaging in Online Communication, *Online Communities and Social Computing* (D. Schuler (Ed.)) (Proc. Second Int'l Conf. OCSC 2007, held as a part of HCI International 2007, Beijing, China), Springer LNCS 4564, pp.141-150 (2007.7)
- [11] Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka: Textual Affect Sensing for Social and Expressive Online Communication, *Affective Computing and Intelligent Interaction* (A. Pavia, P. Prada and R. W. Picard (Eds.)), (Proc. Int'l Conf. ACII2007, Lisbon, Portugal), Springer LNCS 4738, pp.218-229 (2007.9)
- [12] Mostafa Al Masum Shaikh, Helmut Prendinger and Mitsuru Ishizuka: Roles of Emotions: A Linguistic Interpretation of an Emotion Model for Affect Sensing from Texts (poster paper), *Affective Computing and Intelligent Interaction* (A. Pavia, P. Prada and R. W. Picard (Eds.)), (Proc. Int'l Conf. ACII2007, Lisbon, Portugal), Springer LNCS 4738, pp.737-738 (2007.9)
- [13] Mostafa Al Masum Shaikh, Helmut Prendinger and Mitsuru Ishizuka: Assessing Sentiment of Text by Semantic Dependency and Contextual Valence Analysis, *Affective Computing and Intelligent Interaction* (A. Pavia, P. Prada and R. W. Picard (Eds.)), (Proc. Int'l Conf. ACII2007, Lisbon, Portugal), Springer LNCS 4738, pp.191-202 (2007.9)
- [14] Jun Karamon, Yutaka Matsuo, Hikaru Yamamoto and Mitsuru Ishizuka: Generating Social Network Features for Link-Based Classification, *Knowledge Discovery in Databases: PKDD 2007* (J. N. Kok, J. Koronacki, R. L. de Mantaras, S. Matwin, D. Mladenic, A. Skowron (Eds.)) (Proc. 11th European Conf. on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland), Springer LNAI 4702, pp.127-139 (2007.9)
- [15] Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka: An Expressive Avatar for Instant Messaging Endowed with Emotional Intelligence (poster paper), *Intelligent Virtual Agents* (C. Pelachaud, J-C. Martin, E. Andre and G. Chollet (Eds.)) (Proc. 7th Int'l Conf. IVA 2007, Paris, France), Springer LNAI 4722, pp.395-396 (2007.9)
- [16] Paul Piwek, Hugo Hernault, Helmut Prendinger and Mitsuru Ishizuka: T2D: Generating Dialogues Between Virtual Agents Automatically from Text, *Intelligent Virtual Agents* (C. Pelachaud, J-C. Martin, E. Andre and G. Chollet (Eds.)) (Proc. 7th Int'l Conf. IVA 2007, Paris, France), Springer LNAI 4722, pp.161-174 (2007.9)

- [17] Helmut Prendinger, Paul Piwek and Mitsuru Ishizuka: Automatic Generation of Multi-modal Dialogue from Text based on Discourse Structure Analysis, Proc. 1st IEEE Int'l Conf. on Semantic Computing (ICSC-07), pp.27-34, Irvine, CA, USA (2007.9)
- [18] Jie Yang, Yutaka Matsuo and Mitsuru Ishizuka: An Augmented Tagging Scheme with Triple Tagging and Collective Filtering, Proc. 2007 IEEE/WIC/ACM Int'l Conf. on Web Intelligence (WI07), pp.35-38, Silicon Valley, CA, USA (2007.11)
- [19] Mostafa Al Masum Shaikh, Helmut Prendinger and Mitsuru Ishizuka: Emotion Sensitive News Agent: An Approach Towards User Centric Emotion Sensing from the News, Proc. 2007 IEEE/WIC/ACM Int'l Conf. on Web Intelligence (WI07), pp.614-620, Silicon Valley, CA, USA (2007.11)
- [20] Werner Breitfuss, Helmut Prendinger and Mitsuru Ishizuka: Automated Generation of Non-verbal Behavior for Virtual Embodied Characters (Short paper), Proc. 9th Int'l Conf. on Multimodal Interfaces (ICMI-07), ACM Press, pp.319-322, Nagoya, Japan (2007.11)
- [21] Mostafa Al Masum Shaikh, Helmut Prendinger and Mitsuru Ishizuka: An Analytical Approach to Assess Sentiment of Text, Proc. (CD-ROM) Int'l Conf. on Computer and Information Technology (ICCIT2007), session C1, 6 pages, Dhaka, Bangladesh (2007.12)
- [22] Mostafa Al Masum Shaikh, Helmut Prendinger and Mitsuru Ishizuka: Linguistic Interpretation of Emotions for Affect Sensing from Text, (poster paper), Proc. (CD-ROM) 3rd Int'l Joint Conf. on Natural Language Processing (IJCNLP2008), Vol.II, pp.895-900, Hyderabad, India (2008.1)
- [23] Naoaki Okazaki, Mitsuru Ishizuka and Jun'ichi Tsujii: A Discrimination Approach to Japanese Abbreviation Extraction, (poster paper), Proc. (CD-ROM) 3rd Int'l Joint Conf. on Natural Language Processing (IJCNLP2008), Vol.II, pp.889-894, Hyderabad, India (2008.1)
- [24] Danushka Bollegala, Yutaka Matsuo and Mitsuru Ishizuka: A Co-occurrence Graph-based Approach for Personal Name Alias Extraction from Anchor Texts, (poster paper), Proc. (CD-ROM) 3rd Int'l Joint Conf. on Natural Language Processing (IJCNLP2008), Vol.II, pp.865-870, Hyderabad, India (2008.1)
- [25] Kenji Hirohata, Naoaki Okazaki, Sophia Ananiadou and Mitsuru Ishizuka: Identification of Sections in Scientific Abstracts using Conditional Random Fields, Proc. (CD-ROM) 3rd Int'l Joint Conf. on Natural Language Processing (IJCNLP2008), Vol.I, pp.381-388, Hyderabad, India (2008.1)
- [26] Tadanobu Furukawa, Yutaka Matsuo, Ikki Ohmukai, Koki Uchiyama and Mitsuru Ishizuka: Extracting Topics and Innovators Using Topic Diffusion Process in Weblogs, Proc. Int'l Conf. on Weblogs and Social Media (ICWSM2008), pp.182-183, Seattle (2008.3)
- [27] Werner Breitfuss, Helmut Prendinger and Mitsuru Ishizuka: Automatic Generation of Gaze and Gestures for Dialogues Between Embodied Conversational Agents: System Description and Study on Gaze Behavior, Proc. AISB 2008 Symp. on Multimodal Output Generation (MOG'08), (in conj. with the AISB Convention on Communication, Interaction and Social Intelligence), pp.18-25, Aberdeen, Scotland (2008.4)
- [28] YingZi Jin, Yutaka Matsuo and Mitsuru Ishizuka: Ranking Entities on the Web using Social Network Mining and Ranking Learning, Online Proc. WWW2008 Workshop on Social Web Search and Mining (SWSM2008), 4 pages, Beijing, China (2008.4)
- [29] Danushka Bollegala, Taiki Honma, Yutaka Matsuo and Mitsuru Ishizuka: Identification of Personal Name Aliases on the Web., Online Proc. WWW2008 Workshop on Social Web Search and Mining (SWSM2008), 8 pages, Beijing, China (2008.4)
- [30] Danushka Bollegala, Taiki Honma, Yutaka Matsuo and Mitsuru Ishizuka: Mining for Personal Name Aliases on the Web (poster paper), Proc. (CD-ROM) 17th Int'l World Wide Web Conf. (WWW2008), pp.1107-1108, Beijing, China (2008.4)
- [31] Mitsuru Ishizuka: A Common Concept Description of Natural Language Texts as the Foundation of Semantic Computing on the Web (ASC2008 Keynote Talk), Proc. IEEE Int'l Conf. on Sensor Networks, Ubiquitous, and Trustworthy

- Computing, p.385, Taichung, Taiwan (2008.6)
- [32] A. S. M. Mahbub Morshed and Mitsuru Ishizuka: ADROIT: Automatic Discourse Relation Organizer of Internet-based Text (student abstract), Proc. 23rd AAAI Conf. on Artificial Intelligence (AAAI2008), pp.1824-1825, Chicago, Illinois (2008.7)
- [33] Fahim Kawsar, Mostafa Al Masum Shaikh and Mitsuru Ishizuka: Text Beautifier: An Affective-Text Tool to Tailor Written Text (student abstract), Proc. 23rd AAAI Conf. on Artificial Intelligence (AAAI2008), pp.1806-1807, Chicago, Illinois (2008.7)
- [34] Jun Karamon, Yutaka Matsuo and Mitsuru Ishizuka: Generating Useful Network-based Features for Analyzing Social Networks, Proc. 23rd AAAI Conf. on Artificial Intelligence (AAAI2008), pp.1162-1168, Chicago, Illinois (2008.7)
- [35] Danushka Bollegala, Yutaka Matsuo and Mitsuru Ishizuka: WWW sits the SAT: Measuring Relational Similarity from the Web, Proc. 18th European Conf. on Artificial Intelligence (ECAI2008), pp.333-337, Patras, Greece (2008.7)
- [36] Yulan Yan, Yutaka Matsuo, Mitsuru Ishizuka and Toshio Yokoi: Annotating an Extension Layer of Semantic Structure for Natural Language Text, Proc. (CD-ROM) 2nd IEEE Int'l Conf. on Semantic Computing, pp.174-181, Santa Clara, CA, USA, (2008.8)
- [37] Danushka Bollegala, Taiki Honma, Yutaka Matsuo and Mitsuru Ishizuka: Automatically Extracting Personal Name Aliases from the Web, Advances in Natural Language Processing (Proc. 6th Int'l Conf., GoTAL 2008, Gothenburg, Sweden), Springer, LNCS 5221, pp. 77-88 (2008.8)
- [38] Werner Breitfuss, Helmut Prendinger and Mitsuru Ishizuka: Automatic Generation of Conversational Behavior for Multiple Embodied Virtual Characters: The Roles and Models behind Our System (poster), Intelligent Virtual Agents (Helmut Prendinger, James Lester and Mitsuru Ishizuka (eds.)) (Proc. 8th Int'l Conf. IVA 2008, Tokyo), Springer, LNAI 5208, pp.472-473 (2008.9)
- [39] Hugo Hernault, Paul Piwek, Helmut Prendinger and Mitsuru Ishizuka: Generating Dialogues for Virtual Agents Using Nested Textual Coherence Relations, Intelligent Virtual Agents (Helmut Prendinger, James Lester and Mitsuru Ishizuka (eds.)) (Proc. 8th Int'l Conf. IVA 2008, Tokyo), Springer, LNAI 5208, pp.139-145 (2008.9)
- [40] Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka: User Study of AffectIM: an Emotionally Intelligent Instant Messaging System, Intelligent Virtual Agents (Helmut Prendinger, James Lester and Mitsuru Ishizuka (eds.)) (Proc. 8th Int'l Conf. IVA 2008, Tokyo), Springer, LNAI 5208, pp.29-36 (2008.9)
- [41] Paul Piwek, Helmut Prendinger, Hugo Hernault and Mitsuru Ishizuka: Generating Questions: An Inclusive Characterization and a Dialogue-based Application, Online Proc. Workshop on the Question Generation Shared Task and Evaluation Challenge, Arlington, VA, USA (2008.9)
- [42] Mostafa Al Masum Shaikh, Helmut Prendinger, Keikichi Hirose and Mitsuru Ishizuka: An Approach for Ambient Communication by Detecting Real-world Activities from Environmental Sound Cues, Proc. (CD-ROM) IADIS Int'l Conf. on WWW/Internet 2008, pp.504-507, Freiburg, Germany (2008.10)
- [43] Sebastian Ullrich, Helmut Prendinger and Mitsuru Ishizuka: MPML3D: Agent Authoring Language for Virtual Worlds, Proc. Int'l Conf. on Advances in Computer Entertainment Technology (ACE'08), ACM Press, pp.134-137, Yokohama, Japan (2008.12)
- [44] Yulan Yan, Yutaka Matsuo, Mitsuru Ishizuka and Toshio Yokoi: Relation Classification for Semantic Structure Annotation of Text, Proc. (CD-ROM) 2008 IEEE/WIC/ACM Int'l Conf. on Web Intelligence (WI'08), pp.377-380, Sydney, Australia (2008.12)
- [45] Danushka Bollegala, Yutaka Matsuo and Mitsuru Ishizuka: Measuring the Similarity between Implicit Semantic Relations using Web Search Engines, Proc. (CD-ROM) 2009 Second ACM Int'l Conf. on Web Search and Data Mining (WSDM'09), pp.104-113, Barcelona, Spain (2009.02)
- [46] Keigo Watanabe, Danushka Bollegala, Yutaka Matsuo and Mitsuru Ishizuka: A Two-Step Approach to Extracting Attributes for People on the Web, Online Proc. WWW2009 2nd Web People Search

- Evaluation Workshop (WEPS2009), 6 pages, Madrid, Spain (2009.4)
- [47] Yulan Yan, Yutaka Matsuo and Mitsuru Ishizuka: An Integrated Approach for Relation Extraction from Wikipedia Texts, Online Proc. WWW2009 Workshop on Content Analysis in the WEB2.0 (CAW2.0 2009), 7 pages, Madrid, Spain (2009.4)
- [48] Danushka Bollegala, Yutaka Matsuo and Mitsuru Ishizuka: Measuring the Similarity between Implicit Semantic Relations from the Web, Proc. 18th Int'l World Wide Web Conf. (WWW2009), pp.651-660, Madrid, Spain (2009.4)
- [49] Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka: Compositinativity Principle in Recognition of Fine-Grained Emotions from Texts (poster paper), Proc. 3rd Int'l Conf. on Weblogs and Social Media (ICWSM09), pp.278-281, San Jose, California, (2009.5)
- [50] Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka: EmoHeart: Automation of Expressive Communication of Emotions in Second Life, in Online Communities and Social Computing - OCSC2009 (Proc. HCI Int'l 2009, San Diego, USA), Springer, LNCS 5721, pp.584-592 (2009.7)
- [51] Hiroshi Dohi and Mitsuru Ishizuka: A Life-Like Agent Interface System with Second Life Avatars on the OpenSimulator Server, in Online Communities and Social Computing - OCSC2009 (Proc. HCI Int'l 2009, San Diego, USA), Springer, LNCS 5721, pp.182-190 (2009.7)
- [52] Werner Breiffuss, Helmut Prendinger and Mitsuru Ishizuka: Automatic Generation of Non-verbal Behavior for Agents in Virtual Worlds: A System for Supporting Multimodal Conversations of Bots and Avatars, in Online Communities and Social Computing - OCSC2009 (Proc. HCI Int'l 2009, San Diego, USA), Springer, LNCS 5721, pp.153-161 (2009.7)
- [53] Danushka Bollegala, Yutaka Matsuo and Mitsuru Ishizuka: A Relational Model of Semantic Similarity between Words using Automatically Extracted Lexical Pattern Clusters from the Web, Proc. 2009 Conf. on Empirical Methods in Natural Language Processing (EMNLP2009), pp.803-812, Singapore (2009.8)
- [54] Yulan Yan, Naoaki Okazaki, Yutaka Matsuo, Zhenglu Yang and Mitsuru Ishizuka: Unsupervised Relation Extraction by Mining Wikipedia Texts Using Information from the Web, Proc. of Joint Conf. of 47th Annual Meeting of the Association for Computational Linguistics and 4th Int'l Joint Conf. on Natural Language Processing of the Asian Federation of Natural Language Processing (ACL-IJCNLP2009), pp.1021-1029, Singapore (2009.8)
- [55] Alena Neviarouskaya, D. Tsetserukou, H. Prendinger, N. Kawakami, S. Tachi, and Mitsuru Ishizuka: Emerging System for Affectively Charged Interpersonal Communication, Proc. of ICROS-SICE Int'l Joint Conf. 2009 (ICCAS-SICE 2009), Fukuoka, Japan, pp. 3376-3381 (2009.8)
- [56] Mostafa Al Masum Shaikh, Antonio Rui Ferreira Rebordao, Keikichi Hirose and Mitsuru Ishizuka: Emotional Speech Synthesis by Sensing Affective Information from Text, Proc. Int'l Conf. on Affective Computing and Intelligent Interaction (ACII 2009), Amsterdam, The Netherlands, pp.466-471 (2009.9)
- [57] Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka: SentiFul: Generating a Reliable Lexicon for Sentiment Analysis, Proc. Int'l Conf. on Affective Computing and Intelligent Interaction (ACII 2009), Amsterdam, The Netherlands, pp.363-368 (2009.9)
- [58] Werner Breiffuss, Ielka van der Sluis, Saturnino Luz, Helmut Prendinger and Mitsuru Ishizuka: Evaluating an Algorithm for the Generation of Multimodal Referring Expression in a Virtual World: A Pilot Study, Intelligent Virtual Agents (Proc. 9th Int'l Conf. IVA 2009, Amsterdam The Netherlands), Springer, LNAI 5773, pp.181-187 (2009.9)
- [59] Mostafa Al Masum Shaikh, Helmut Prendinger, Keikichi Hirose and Mitsuru Ishizuka: Easy Living in the Virtual World: a Novel Approach to Integrate Real World Activities to Virtual Worlds, Proc. 2009 IEEE/WIC/ACM Int'l Conf. on Intelligent Agent Technology (IAT'09), Milan, Italy, pp.466-473 (2009.9)
- [60] Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka: Semantically Distinct Verb Classes Involved in Sentiment Analysis, Proc. IADIS Int'l Conf. on Applied Computing 2009, Rome, Italy, pp.27-34 (2009.11)

- [61] Hiroshi Dohi and Mitsuru Ishizuka: An Interactive Presentation System in a 3D Virtual World using an OpenSimulator Server, Proc. Image Electronics and Visual Computing Workshop 2010, 3A-1, Nice, France (2010.3)
- [62] Yulan Yan, Haibo Li, Yutaka Matsuo and Mitsuru Ishizuka: Multi-view Bootstrapping for Relation Extraction by Exploiting Web Features and Linguistic Features, Computational Linguistics and Intelligent Text Processing (Proc. 11th Int'l Conf. CICLing 2010, Iasi, Rumania), Springer, LNCS 6008, pp.525-536 (2010.3)
- [63] Hugo Hernault, Danushka Bollegala and Mitsuru Ishizuka: A Sequential Model for Discourse Segmentation, Computational Linguistics and Intelligent Text Processing (Proc. 11th Int'l Conf. CICLing 2010, Iasi, Rumania), Springer, LNCS 6008, pp.315-326 (2010.3)
- [64] Tadanobu Furukawa, Seishi Okamoto, Yutaka Matsuo and Mitsuru Ishizuka: Prediction of Social Bookmarking based on a Behavior Transition Model, Proc. 2010 ACM Symposium on Applied Computing (ACM SAC 2010), Sierre, Switzerland, pp.1741-1747 (2010.3)
- [65] Yulan Yan, Haibo Li, Yutaka Matsuo, Zhenglu Yang and Mitsuru Ishizuka: Multi-View Clustering with Web and Linguistic Features for Relation Extraction, Proc. 12th Asia-Pacific Web Conference (APWeb 2010), Busan, Korea, pp.140-146 (2010.4)
- [66] Danushka Bollegala, Yutaka Matsuo and Mitsuru Ishizuka: Relational Duality: Unsupervised Extraction of Semantic Relations between Entities on the Web, Proc. 19th Int'l World Wide Web Conf. (WWW 2010), Raleigh, North Carolina, USA, pp.151-160 (2010.4)

[図書] (計 3 件)

- [1] Yutaka Matsuo, Junichiro Mori and Mitsuru Ishizuka: Social Network Mining from the Web, in "Data Mining Patterns -- New Methods and Applications" (P. Poncelet, M. Teisseire and F. Masegla (Eds.)), Chapter VII, pp.149-175, Information Science Reference (2007)
- [2] Mostafa Al Masum Shaikh, Helmut Prendinger and Mitsuru Ishizuka: A Linguistic Interpretation of the OCC

Emotion Model for Affect Sensing from Text, in "Affective Information Processing" (Jianhua Tao and Tieniu Tan (eds.)), Chapter 4, pp.45-73, Springer (2008.11)

- [3] Yingzi Jin, Yutaka Matsuo and Mitsuru Ishizuka: Ranking Companies Based on Multiple Social Networks Mined from the Web, in: "E-commerce" (Kyeong Kang (ed.)), INTECH, Chapter 6, pp.75-98 (2010)

[産業財産権]

○出願状況 (計 2 件)

名称: 検索方法及びシステム

発明者: ボッレーガラ・ダヌシカ, 石塚 満,
グエン・トアン・ドック

権利者: 東京大学

種類: 特許出願

番号: 2009-275762

出願年月日: 2009 年 12 月 3 日

国内外の別: 国内

名称: テキストコーパスにおける 2 つのエン
ティティ間の関係抽出方法及び装置

発明者: ボッレーガラ・ダヌシカ, 石塚 満

権利者: 東京大学

種類: 特許出願

番号: 2010-096551

出願年月日: 2010 年 4 月 20 日

国内外の別: 国内

○取得状況 (計 0 件)

[その他]

ホームページ等

<http://www.miv.t.u-tokyo.ac.jp/>

6. 研究組織

(1) 研究代表者

石塚 満 (ISHIZUKA MITSURU)

研究者番号: 50114369

(2) 研究分担者

なし

(3) 連携研究者

なし