

要求記述文章の静的側面の可視化による CSCW の実現

黒田 剛 塚本 明 上田 賀一

茨城大学 工学部 情報工学科

〒 316-8511 茨城県日立市中成沢町 4-12-1

本研究では、要求情報をシステム開発に関わる人達にとって分かりやすく表現し、理解支援・伝達支援を行うことで協調活動のコンピュータ支援の実現を試みた。要求記述を形態素解析/構文解析し、文章の静的構造を捉え、各要素とそれらを結ぶ関係をリッチピクチャとして表した。本研究の特徴として、ドメイン辞書と連携させたことによる様々な情報の付加が可能となったことが挙げられる。ドメイン辞書を活用して、実現したツールに画像を添付する機能や、要素の意味を表示させる機能、リッチピクチャから類推できる要求記述を作成する機能を持たせた。リッチピクチャに様々な情報を加えることで多面的に内容を捉えることが可能となり、共同作業におけるメンバへの伝達効率の改善が可能となった。

A tool for computer supported cooperative works by visualization of static aspect of requirement description

Tsuyoshi Kuroda Akira Tsukamoto Yoshikazu Ueda

Ibaraki University

4-12-1 Naka-Narusawa, Hitachi, Ibaraki, 316-8511 Japan

This report deals with realization of computer supported cooperative works in order to help understanding and communication of requirers and developers. The static structure among requirements description is caught through morphological analysis and syntax analysis. And each element and relation between elements are represented as rich picture. It is characteristic of the report that addition of various information became possible by connection with a domain dictionary. Using a domain dictionary, the following 3 functions are made: to attach an image to rich picture, to indicate meaning from rich picture, to make the requirements description sentence that can be reasoned from rich picture. With the tool developed by this time, members in group work can catch demand contents versatily and may understand each other's intention more.

1 はじめに

近年の情報処理技術の発展や機器の普及、ネットワークの整備に伴い、従来紙に記述されてきた要求記述やメモ書き、口頭で伝えられてきた内容までもが電子化されている。さらに、そこから得られる情報や知識は膨大な量となり、その情報や知識を正確に判断することが困難となっている。

そこで、要求情報の収集・蓄積・提示を計算機によって効率的に自動化し、特に開発に関わる人達にとって分かりやすく表現することで理解支援・伝達支援を行うことが必要となる。しかし、要求記述だけといった単一ビューでは、それを閲覧した人達によって理解が異なることもしばしばである。その問題を克服する方法のひとつとして、要求記述など一覽性の悪いビュー構造を概念図などのより直観的なビューで表現する方法が挙げられる。

そこで本研究では、要求記述の静的側面に注目し、その内部構造から概念図をリッチピクチャ[1]として自動生成するツールを開発した。要求記述を形態素解析/構文解析し、品詞の同定や依存関係を決定することで、文章の静的側面を捉え、形態素からなる各要素とそれらを結ぶ関係をリッチピクチャとして表す。

またドメイン辞書と連携し、作成したリッチピクチャに画像を添付し、要素の意味を表示させることで様々な情報を付加し、多面的に内容を捉えることが可能な高品質のリッチピクチャを作成することで、柔軟な要求分析の支援およびコンピュータ支援共同作業 [2] の実現することを目的とする。

2 基盤知識

本研究において基礎知識となることについて説明する。

2.1 本研究で利用したドメイン辞書

現在利用されているドメイン辞書 [3] は、主に定義や類義語といった辞書的な情報を有している。本研究では、本研究で開発したツールを用いたシステム開発に必要な情報を保持できるようにするためにドメイン辞書の拡張を試みた (図 1)。拡張した点は次の通りである。

- 辞書要素に関する画像データの格納が可能
- 辞書要素に関する要素名の情報の格納が可能

本論文で以降に記述されている「ドメイン辞書」は、この拡張されたドメイン辞書を指すことにする。

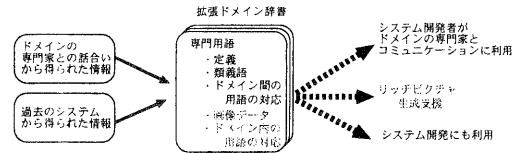


図 1: ドメイン辞書の拡張

2.2 ドメイン辞書の管理

本来、各ドメインは極めて複雑な関係を保有しながら存在している。ドメイン辞書の要素もドメイン毎に複雑な関係はそのままに管理されるのが理想であるが、そのような状態では、管理しづらく機能も限定されてくる。そこでオブジェクト指向におけるオブジェクト間の関係のように、ドメイン間の関係も明確であれば管理面・機能面もより向上できる。そこで本研究では、ドメイン辞書に格納されるドメインを管理しやすくするために、階層構造 (木構造) を用いてドメイン辞書を実現した。本研究で用いた木構造の中の全てのドメインは、それぞれの子ドメインが自分の親ドメインを継承することにより親子関係を持つ。従って親ドメインに属する用語は、その子ドメインから参照することが可能となる。

3 解析・生成

ここでは、本研究で開発したツールの概略を説明すると共に、要求記述を解析する「解析部」、解析した結果を図的表現する「生成部」について説明する。

3.1 概略

本研究で開発したツールは、要求記述を意味的構造へと解析する「解析部」と、意味的構造へと

解析された結果を用いて図的表現する「生成部」から構成されている。

解析部では、要求記述からリッチピクチャの要素と要素間の関係を抽出するために、形態素解析と構文解析を行う。本研究では日本語形態素解析システム JUMAN[4] と日本語構文解析システム KNP[5] を利用する。

また、要求記述から図的表現であるリッチピクチャを生成するまでの流れを図 2 に示す。形態素解析/構文解析より得られる結果を本ツールでは意味的構造として用いる (図 5)。

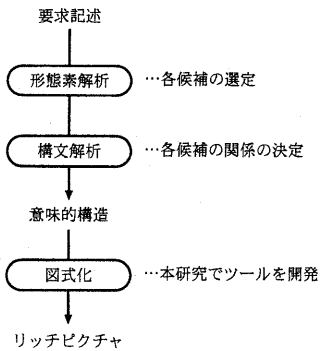


図 2: 要求記述からリッチピクチャを生成する流れ

3.2 解析部

ここでは、本研究で開発したツールの解析部について説明する。

3.2.1 JUMAN による形態素解析

本研究で開発したツールでは、専用 GUI のテキストフィールドに記述された文章をシステムに対する入力とし、はじめに形態素解析を行う。

専用 GUI から JUMAN パーサへの入力までの流れを図 3 に示す。

3.2.2 KNP による構文解析

前述の形態素解析に続いて、KNP を用いた要求記述の構文解析を行う。

KNP による構文解析は、JUMAN の出力結果を利用して行われる。JUMAN による解析結果が

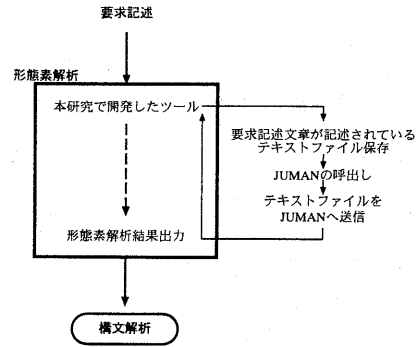


図 3: JUMAN による形態素解析の流れ

ら KNP への入力までの流れを図 4 に示す。

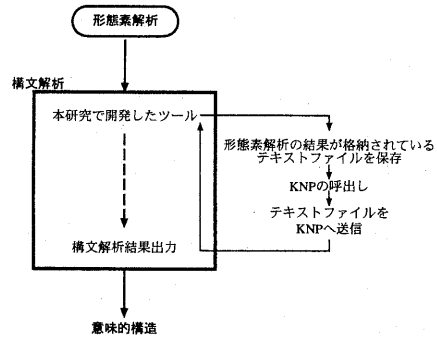


図 4: KNP による構文解析の流れ

3.3 生成部

ここでは、本研究で開発したツールの生成部について説明する。

3.3.1 解析結果の読み込み

JUMAN/KNP で解析された結果である意味的構造から、静的構造を作成する上で必要な情報である「名詞」と「動詞」を抽出する。解析結果では、それぞれの形態素に対して「名詞」「動詞」「助詞」「特殊」等の品詞が割り当てられている。それと同時に「名詞」には「9」、「動詞」には「2」といったユニークな番号が割り振られている。本研

究で開発したツールではこのユニークな番号に着目し、解析のキーとして用いることにした。

3.3.2 名詞と動詞の関係付け

意味的構造にはもうひとつ情報が盛り込まれている。それは、「対象としている形態素がどの形態素と依存関係にあるか」ということである。

意味的構造(図5)は、“5D”や“3D”といった深層格の情報を持っている。本研究において、この深層格の情報を深層度と言うことにする。任意の深層度に対して、それより深いもの(文中ではそれ以前に出現する要素)、すなわち深層度が小さいものは、現在対象としている深層度に依存している。また逆に、対象としている深層度より浅いもの(文中ではそれ以後に出現する要素)、すなわち深層度が大きいものは、現在対象としている深層度には依存しない。視点の変更は、依存元となる深層度よりさらに深い(文中ではそれ以前に出現する要素)、すなわち深層度が小さいものが出現したときに行われる。この作業を繰り返すことで、意味的構造から要素間の関係を決定する。

3.3.3 関係の方向を定めるための助詞分類

意味的構造の解析によって、要素間の関係が明らかになった。ここでは、その関係を色分けし、方向を定めることで、視覚的に直観で分かりやすくするための作業を行う。

方向を定めるキーとして、本ツールでは「助詞」に着目した。JUMANの辞書に登録されている助詞を、以下の4つに分類した。

- 矢印の元であることを示しそうな助詞
は、が、から、より、くらい、ぐらい
- 矢印の先であることを示しそうな助詞
に、を、へ、まで、の、や、やら、だの、か、および、ならびに、または等
- 属性にもなりそうにない助詞
など、ったら、って、さて、でも、すら、だって、ばかり、ばかり、なんか、なんて等
- まだ判別できていない助詞
と、で、にて、なら、も、だけ、のみ、しか、こそ、それとも、それに、けど等

3.3.4 リッチピクチャの生成

本研究で開発したツールでは、単純明解にして理解が容易であるダイアグラム(ノードとアーク)方式を表記法として用いた。アークは要素を、ノードは関係を表している。

これまでの解析結果に基づいて表現すると、例文は図6のようなリッチピクチャに変換される。

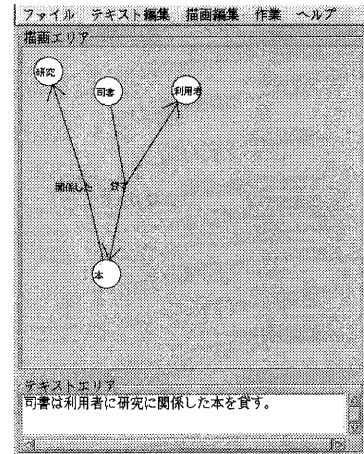


図6: 例文を解析した結果

4 可視化ツールの実現

リッチピクチャの表記法、より良くツールを利用するための機能について説明する。

4.1 表記法

本研究で開発したツールでは、前章で述べた通り、ノードとアークを用いて基本的なりッチピクチャを作成している。ここではさらに複雑な関係に対する表記法について述べる。要素間の関係は3種類存在すると考えられる。それは要素間の利用・参照関係、要素間の構造的な集約関係、要素間の概念的な包含関係である。要素間の利用・参照関係は、要素間の静的な繋がりを意味し、要素間の利用・参照関係を定義するために用いられる関係であるため、要素同士の関係は階層的観点において対等となる。だが、要素間の構造的な集約

```
# S-ID:1
* 5D <文頭></><助詞><体言><係:未格><提題><区切:3-5>
司書 ししょ 司書 名詞 6 普通名詞 1 * 0 * 0 NIL <文頭><漢字><自立>
は は は 助詞 9 副助詞 2 * 0 * 0 NIL <付属>
* 5D <二><助詞><体言><係:二格><区切:0-0>
利用 りよう 利用 名詞 6 サ変名詞 2 * 0 * 0 NIL <漢字><サ変><自立>
者 しゃ 者 接尾辞 14 名詞性名詞接尾辞 2 * 0 * 0 NIL <漢字><自立><←複合>
に に に 助詞 9 格助詞 1 * 0 * 0 NIL <品曖-格助詞><品曖-接続助詞><付属>
* 3D <二><助詞><体言><係:二格><区切:0-0>
.....
```

図 5: 意味構造 (一部)

関係やそして要素間の概念的な包含関係では、要素間の関係は、階層的観点において対等でなくなる。よって本研究では、要素間の参照・利用関係を「対等な関係」、要素間の構造的な集約関係や要素間の概念的な包含関係を「対等でない関係」と述べることにする。

本研究では「対等な関係」、「対等でない関係」を判別するために助詞に着目した。JUMANに登録されている助詞を分け、「対等である」/「対等でない」を色分けした棒線で表現することでリッチピクチャの理解容易性をさらに高めることにした。色が表す意味は表 1 に示す通りである。

表 1: 色分けで表している意味

色	意味
赤	助詞だけだと判別できない関係
青	主格に関する関係
緑	属性にもならない関係
黒	助詞を含まない関係

4.2 機能

本研究で開発されたツールでは、クライアントとベンダ間の意志疎通をさらに促進させるために、さらに4つの機能を実装した。

4.2.1 要素の削除

本研究で開発されたツールで作成されたリッチピクチャは、要求記述全文を解釈するため、当然クライアントとベンダにとって余分だと思われる部分の文章も解釈されることになる。そこで、「要素を削除する」という編集機能を付け加えることで、システム利用者の視点から、あまり有効ではないリッチピクチャの要素が削除され、リッチピクチャとしての質を高めることができる。

4.2.2 関係画像の貼り付け

前節において、本ツールで用いる表記法の基礎について説明した。しかし、前節での表記法だけでは、クライアントとベンダ双方が、対象世界における要素同士の関係は把握できても、対象世界の具体的なイメージが食い違う可能性もある。対象世界に適した具体的なイラストやイメージを貼り付けることは、対象世界の概要の理解容易性を高めるために有効な手段であると考えられる。そこで本システムにおいて、リッチピクチャにおけるそれぞれの要素に適したイラストやイメージを貼り付けることで、クライアントとベンダ間の意志疎通を促進し、双方の対象世界に関する理解容易性を高めようと試みた。

本研究では、対象世界でのそれぞれの要素に適したイラストやイメージを選定するために、ドメイン辞書を利用し、通信することにした。ドメイン辞書は、PostgreSQL[6]で管理されているデー

データベース内に格納されている。本ツールとドメイン辞書の通信概要を図7に示す。

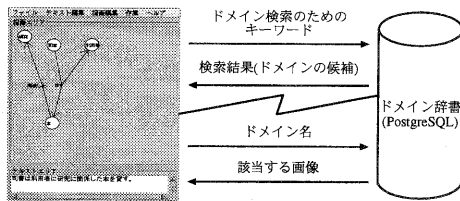


図7: ドメイン辞書との通信概要 (画像問合せ)

関係イラストやイメージを検索し貼り付ける際、本ツールからドメイン辞書が格納されているデータベースへ2回の通信を要求する。要求内容と順番は次の通りである。

(1) ドメイン検索のためのキーワード

リッチピクチャを構成する要素をキーワードに、ドメイン辞書内をOR(論理和)検索するようにデータベースに依頼する。

(2) ドメインの候補群

キーワード(要素)から類推されるドメインの候補群を返答してもらいそれを表示することで、システム利用者に最適なドメインを選択してもらおう基準とする。

先の例文から類推されるドメインの候補を図8に示す。このように、“レンタルドメイン”と“図書館ドメイン”が候補となる。

ドメインの候補	▼ 画像貼付
レンタルドメイン	
図書館ドメイン	

図8: ドメインの候補

(3) ドメイン名

次に図8から該当するドメインを選択する。例文の場合では“図書館ドメイン”を選択し、ドメイン辞書に送信する。

(4) 該当する画像

最後に(3)で選択されたドメインの候補に基づき、ドメイン辞書に登録されている情報とキーワード(要素)を比較し、それに見合ったイラストやイメージが登録されているならばそれらを送り返してもらおう。送り返されたイラストやイメージのデータは任意のディレクトリに展開され、その後、各要素の近隣に先に返答されたイラストやイメージを貼り付ける。

4.2.3 要素の意味問合せ

さらにリッチピクチャの質を高めるために、ここでは要素の持つ意味に注目した。

クライアントとベンダ間において、意志疎通の障害となっているのが“専門用語”である。普段何気なく利用している言葉でも、ドメインが異なると意味が異なる場合もある。そこで、ドメイン辞書に登録されている情報を基に、リッチピクチャに表示されている要素の意味をオンラインで表示することが可能ならば、意志疎通の障害を取り除くことが可能となる。

要素の意味問合せにおける本ツールとドメイン辞書の通信概要を図9に示す。

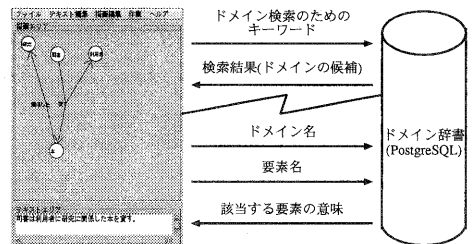


図9: ドメイン辞書との通信概要 (意味問合せ)

また、意味問い合わせにおけるドメイン辞書への通信要求と要求内容・順番は次の通りである。

(1) ドメイン検索のためのキーワード

リッチピクチャを構成する要素をキーワードに、ドメイン辞書内をOR(論理和)検索するようにデータベースに依頼する。

(2) ドメインの候補群

キーワード(要素)から類推されるドメインの候補群を返答してもらいそれを表示することで、システム利用者に最適なドメインを選択してもらおう基準とする。

先の例文から類推されるドメインの候補は、図8に示すように、“レンタルドメイン”と“図書館ドメイン”とである。

(3) ドメイン名

次に図8から該当するドメインである“図書館ドメイン”を選択し、ドメイン辞書に送信する。するとドメイン辞書側で待ち状態になる。

(4) 要素名送信

意味が知りたい要素にフォーカスを当てると、自動的にフォーカスされた要素名がドメイン辞書に送信される。

(5) 該当する要素の意味

最後に (3) で選択されたドメインの候補に基づき、ドメイン辞書に登録されている情報とキーワード(要素)を比較し、それに見合った意味が登録されているならば意味データを送り返してもらい、送り返された意味データはすぐに専用ウィンドウに表示される。

4.2.4 成立しない要素間関係の検出

日本語の特徴のひとつに様々な省略を許可していることが挙げられるが、何気なく単語を省略することによって文章の依存構造は崩れ、要素同士での関係が成り立たなくなる。記述されている文章の長さが長い程、この落とし穴は多く存在する。さらにこのことが原因となってクライアントとベンダ間で共通の理解を得ることが妨げられる要因にもなる。

そこで本研究では、構文解析を済ませた後のデータをさらに解析することで、要素間の関係が不成立であるものをリストとしてまとめた。さらにそのリストを表示することで、孤立している要素を確認し要求記述を改良することで、要求記述の質を向上させることを試みた。図 10は構文解析した後のデータから要素間の関係として成立しないものをリストとしてまとめるまでの流れである。

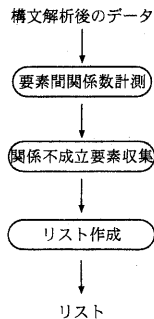


図 10: リスト作成までの流れ

4.2.5 リッチピクチャからの文章生成

前項までで紹介した機能によって、リッチピクチャの質は向上し、クライアントとベンダ間での意志疎通を支援できるものになった。しかし、いくら質が高いリッチピクチャが作成されても、それ

が入力、つまり要求記述の本意とかけ離れたものでは意味を成さない。そこで本システムでは、リッチピクチャからドキュメントを類推し作成する作業を試みた。

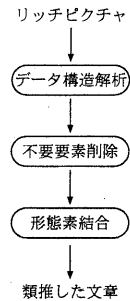


図 11: 類推される文章を生成する手順

図 11はリッチピクチャから文章を生成するまでの流れを示している。本研究で開発したツールでは、入力である要求記述を解析し形態素とする時にフラグを持たせている。そのフラグは、リッチピクチャを生成する時に対象とする要素が不要であると判断されたとき、また対象とする要素を削除した時に外される。これらの作業を経て、必要であると思われる要素を繋ぎ合わせて文章とし、それを表示する。

5 可視化ツールの実行

本研究で開発したツールの実行状態を、図 12に示す。本ツールにより要求記述から自動的にリッチピクチャを生成し、2つのビューを用いることで内容を多面的に捉えることが可能となった。また、柔軟な要求支援が可能となり、コンピュータ支援共同作業の実現が可能となるという本システムの有用性を確認できた。

6 関連研究との比較

本研究と同様に文章内容の可視化を行う研究と比較した。

村山らの研究 [7] は、与えられた文章内容を可視化するために XML の拡張である GDA を用いたタグ付き文章を用い、そこから意味的構造を生

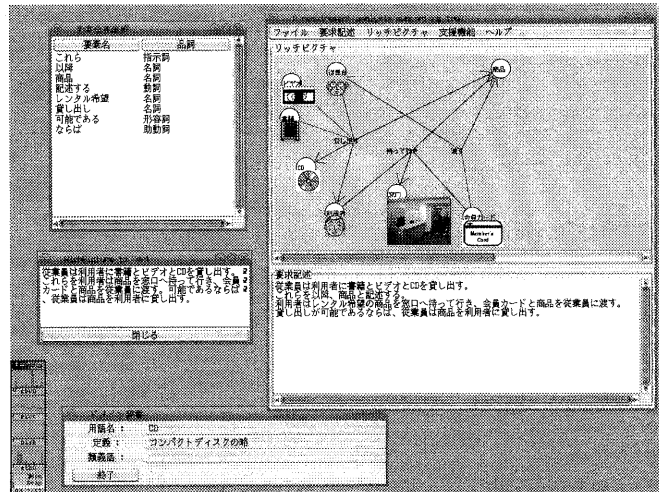


図 12: 本ツールの試行実験

成しさらに静的概念図を生成する方法を提案している。これは静的概念図を生成するだけではなく、GDA を用いたことで今後需要が高まるであろう XML 形式での蓄積にも対応できるというものである。

村山らの研究で利用されるタグ付き文章は自動生成の段階まで到達しておらず、全て人の手によって作成されている。また生成された静的概念図の修正は要素の移動だけであり、GDA を用いたタグ付き文章にまで修正が施せないところは、本研究とは大きく異なる点である。

7 まとめ

本研究では、既存の形態素解析/構文解析ツールを利用し、その結果から要求記述の静的側面をリッチピクチャとして可視化することを可能にした。また、リッチピクチャに様々な情報を付加することで多面的に内容を捉えることが可能となり、グループワークにおけるグループメンバへの伝達効率の改善が可能となった。そして、試行実験を行い、関連研究と比較することで本研究で開発したツールの有用性および問題点を明確にした。

本研究で開発したツールによって、要求記述の内容を多面的に捉えられ、より柔軟な要求分析を

支援することが可能となった。

今後は、要求記述の動的側面を捉える方法を開発すると共に、静的側面を捉えた図、すなわちオブジェクト図の自動生成へ向けて研究を進める予定である。

参考文献

- [1] Peter Checkland, Jim Scholes 著, 妹尾堅一郎 監訳: “ソフト・システムズ方法論”, 有斐閣 (1994)
- [2] 原島 博 監修 電子情報通信学会 編 NTT ヒューマンインタフェース研究所 著: “CSCW とグループウェア”, オーム社 (1993)
- [3] 伊藤 潔, 杵嶋 修三, 田村 恭久 著, 吉田 裕之 編集: “ドメイン分析・モデリング”, 共立出版 (1996)
- [4] Sano Hiroshi, Kawada Ryouichi, Hashimoto Minako: “Morphological Grammar Rules: An Implementation for JUMAN”, International Workshop on Sharable Natural Language Resources, <http://cl.aist-nara.ac.jp/lab/events/SNLR/snlr.html> (1994)
- [5] Sadao Kurohashi, Makoto Nagao: “KN Parser: Japanese Dependency/Case Structure Analyzer”, International Workshop on Sharable Natural Language Resources, <http://cl.aist-nara.ac.jp/lab/events/SNLR/snlr/html> (1994)
- [6] 石井 達夫 著: “PostgreSQL 完全攻略ガイド”, 技術評論社 (1999)
- [7] 村山 正司, 中村 裕一, 太田 友一: “概念図の自動生成による文章内容の可視化-タグ付き文章からの自動変換-” 第 5 回知識情報メディアシンポジウム, Nov (1999)