

ラダリング型検索サービスのための 対話エンジンの設計・開発

北村 美穂子 下畑 さより 介弘 達哉 池野 篤司 坂本 仁 折原 幾夫 村田 稔樹

沖電気工業株式会社

E-mail: {kitamura655,shimohata245,sukehiro564,ikeno546,sakamoto535,oriyara351,murata656}@oki.com

あらまし 近年インターネットの普及と価値観の多様性の拡大により、様々なサービスやコンテンツが提供される時代になっているが、多種多様化しているサービスやコンテンツは様々な言葉や形式で表現されているため、従来のキーワード型の検索サービスだけでは自分のニーズに合ったものを見つけないことができない。

我々は、対話システムにおいて、「対話の中で徐々に掘り下げた質問を繰り返すことにより、ユーザの真のニーズや価値観を引き出す」ラダリング手法を用いることによりシステムがユーザに質問を投げかけ、ユーザが単独では表現できなかったキーワードや表現を引き出し、多種多様でかつ大量のサービスやコンテンツの中からそれとマッチするものを探し出す「ラダリング型検索サービスシステム」を構築した。本稿は、対話エンジン部を中心に本システムの全体概要を説明する。

キーワード 対話システム, 対話戦略, 検索サービス, ラダリング

Design and Development of Dialogue System for “Laddering” Search Service

Mihoko KITAMURA Sayori SHIMOHATA Tatsuya SUKEHIRO Atsushi IKENO

Masashi SAKAMOTO Ikuo ORIHARA and Toshiki MURATA

Ok Electric Industry Co., Ltd.

E-mail: {kitamura655,shimohata245,sukehiro564,ikeno546,sakamoto535,oriyara351,murata656}@oki.com

Abstract It has become common to search needed services and contents using the Internet, but it is difficult to find exactly what one is looking for through keywords as each service is described in just too many ways.

We developed "Laddering" Search Service System that matches users with the search targets by communicating with the users through interviews. The system consists chiefly of the laddering dialog engine employing "laddering method" which enables the engine to ask questions to users and extract keywords and expressions that users probably would not have been able to express on their own. This paper describes the outline Laddering Search Service System, focusing the laddering dialog engine.

Keyword dialogue system, dialogue strategy, search service, laddering

1. はじめに

近年インターネットの普及と価値観の多様性の拡大により、様々なサービスやコンテンツが提供される時代になってきている。特に、ネット上ではそのインフラコストが小さくなってきていることから、ロングテール市場と呼ばれるようなニッチなニーズに合わせたサービス・コンテンツの提供がますます増えてきている。

一方、キーワードによるマッチングの限界も見え始めている。あまりにサービス・コンテンツが細分化したために、ユーザはどのようなキーワードで情報を探せばよいかかわらなくなっているのである。例えば、ショッピングにおいて、新しい冷蔵庫を買いたいと思っても、ネット上には数多くの冷

蔵庫の情報が存在するにも関わらず、最近の冷蔵庫がどのような特長を持っているかを知らなければ、単に検索サイトで勧められる冷蔵庫しか検索することができない。つまり、せっかくサービスやコンテンツが多種多様化してきたにもかかわらず、その大多数が情報の海に埋もれたままになって、真に欲するユーザのもとに届けられていない状況になってきている。

ユーザが適切な検索キーワードを知らなくても、検索できるようにするためには、ユーザとの対話によって検索に必要な情報を引き出す対話型の検索システムが有効である。但し、ユーザ自らが発する質問や回答は曖昧であることが多い。そこで、対話によって不足情報をユーザから引き出す対話システム

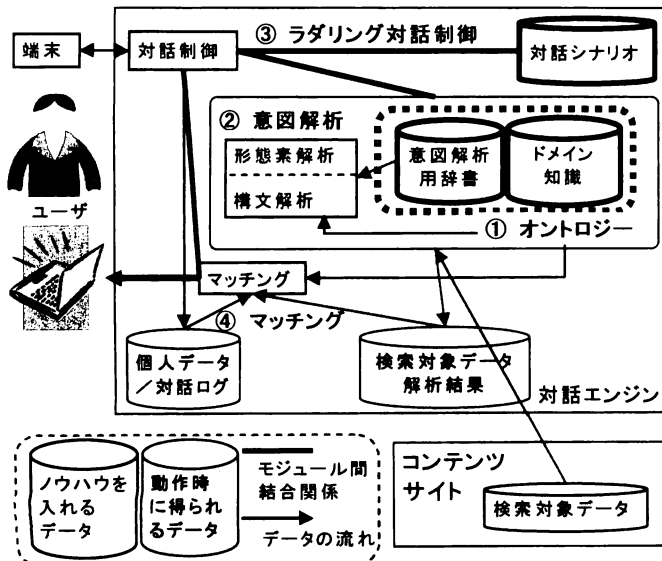


図1 ラダリング型検索サービスシステムの概観

の研究が行われている。テクニカルサポートを扱った対話システムは、単に不足情報を補う質問をするのではなく、ユーザが求める情報に対話を誘導する[1,2]。

テクニカルサポートの場合では、ユーザはシステムの誘導に従って客観的な現象を説明することで求める情報を得ることができるが、一般的な検索の場合では、ユーザ自身求める情報が曖昧なため、何を説明すべきかわからないことが多い。我家に合った冷蔵庫を購入したいが、求める製品の仕様がわからないという前述の例や、ある事柄に対して漠然とした意見があるもののそれを明確に説明できない、という例である。このような問題を解決するために、対話システムとの対話を通じて、ユーザ自らの意志や希望を明らかにするという研究がある[3,4,5]。大塚、丸本ら[3,4]はリフレイミングという対話手法を用いて、対話型意見収集システムを実装している。

我々は、カウンセラーがユーザの真のニーズや価値観を引き出すために用いる対話手法の1つである「ラダリング手法」に基づいて、ユーザとシステムが対話することにより、対話からユーザの真の要求を引き出し、その要求事項にマッチしたデータを検索対象データから探し出す「ラダリング型検索サービス」を設計、開発した。ユーザの真の要求の取得が難しいとされる転職先紹介サービスを実現するシステムを構築し、約800人のユーザによる実証実験を行い、多くの知見を得た。

本稿は、ラダリング型検索サービスシステムについて対話エンジンを中心に説明する。次章では、シ

ステムの全体概要について説明し、3章では対話エンジンの核となるラダリング対話制御について説明する。4章では転職先紹介サービスにおける実装システムについて紹介し、最終章でまとめる。

2. ラダリング型検索サービスの概要

「ラダリング型検索サービス」とは、対話の中で徐々に掘り下げた質問を繰り返すことにより、ニーズや価値観を引き出す（以下、「ラダリング」という）手法を活用した対話型の検索サービスである。利用者とシステムが対話シナリオに基づいて対話を繰り返すことにより、利用者が自ら表現することが困難であったサービスやコンテンツに関するニーズを引き出す。引き出されたニーズに対して、多種多様かつ大量のサービスやコンテンツからマッチするものを探し出し、マッチしたサービスやコンテンツを提供する。

「ラダリング型検索サービスシステム」の概観を図1に示す。ユーザとの対話を司る対話エンジンは、主に次の4つのモジュール（オントロジー、意図解析、ラダリング対話制御、マッチング）から構成される。以下概説する。

① オントロジー [6]

ユーザの発話の意図を解析したり、その解析結果によってシステムが次に何を発話するかを決定するには、意図を解析するための言語知識（形態素辞書や構文辞書）だけでなく、その話題となるドメインに関する詳細な言語知識が必須である。これらの言語知識をオントロジーと呼ぶ。オントロジーは、意図解析やラダリング対話制御に利用されるだけでなく、検索の対象となるデータを拡張する（同義語、類義語、異表記語にも検索対象範囲を拡張する）ためにも利用される。

② 意図解析

ユーザがテキスト入力した文の意図を解析する¹。その解析された意図に基づいて、次のシステム発話文を決定し、ユーザとシステムとの対話が進行する。意図解析では、上記のオントロジーを利用して、形態素解析、構文解析[7]の結果をもとにユーザの意図を得る。本システムにおける意図とは、システムの質問に対する「YES(はい)」、「NO(いいえ)」、「MAYBE(曖昧)」等の一般的な返事の他、オントロジーに存在する値そのものを指す。

③ ラダリング対話制御

¹ 音声入力のIFも持っているが、実証実験ではユーザは文をタイプして入力した。

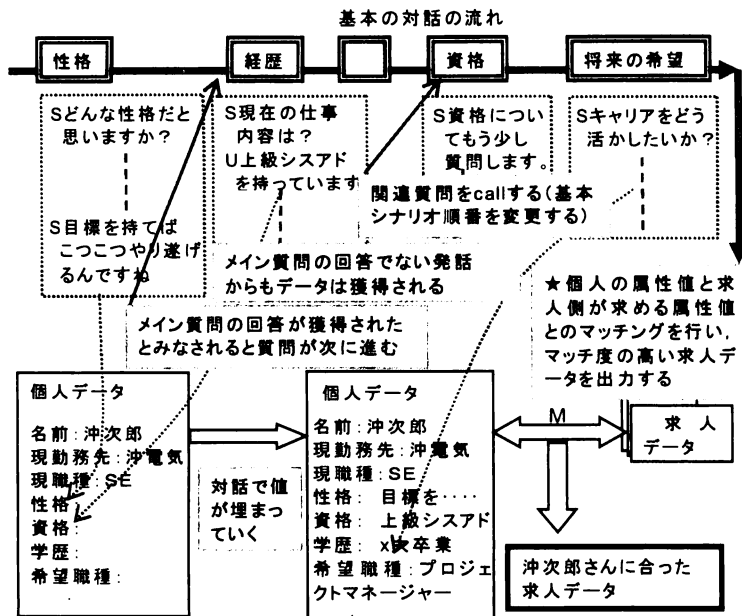


図2 対話進行の模式図

意図解析によって得られた意図、対話シナリオ及びマッチング状況に基づいて、次のシステム発話文を決定する。本モジュールについては次章で詳説する。

④ マッチング

検索したいサービスやコンテンツに関するデータと意図解析技術によって獲得した個人データとのマッチングを行い、ユーザのニーズに合ったデータを出力する。

3. ラダリング対話制御

3.1. 設計方針

ラダリング対話制御の目的は、意図解析によって得られたユーザの意図に対して対話をどのように進行させるかを決定することである。さらに、ラダリング対話制御では、ユーザの深層のニーズや価値観を引き出すことが求められるため、システムがユーザに対して一方的に質問し、その回答を得るだけでなく、システムがユーザの意見を受け止めたり、ユーザの回答に対してさらに掘り下げた質問をするという対話の技法が求められる。これらの技法を実現するために、我々はラダリング対話について調査、分析を行い、その結果を受けて、対話シナリオ²及び対話制御方式の設計を行った。本システムで実現が望まれる対話進行を模式的に表した図を図2に示す。

対話制御モジュールの基本動作は、図2のMマ

²本稿では「対話シナリオ」とは、システムがもつ対話ルールの集合を指す。一方、実際のカウンセラーが作成した想定対話集を「対話サンプル」と呼ぶ。

クのマッチングにおいて必要な個人データを順に聞いていくことである。ベースとなる質問の内容や順番は、カウンセラーの知見に基づいてあらかじめ決定することができる。しかし、ユーザからより自然に個々の値を引き出すためには、単純なアンケートのように順番に質問するのではなく、カウンセラーが実際行う対話のように、その時点での対話内容に沿って話題を転換していくことが望まれる。ユーザの回答によって質問の内容が変わったり、質問の仕方(質問の表現方法)が変わったりする必要がある。さらに、実際の対話では、ユーザは、システムがする質問以外の個人データを提供することも多く、質問に直接関係のない回答も値として獲得できる必要がある。

上記に述べた基本動作を実現するための対話シナリオの仕様及び対話

方式を以下にまとめる。なお、以降、ユーザ発話から獲得すべき個人データの項目(図2の個人データ内の「資格」、「希望職種」)を属性と呼び、実際獲得される値そのもの(図2の個人データ内の「上級シスアド」、「プロジェクトマネージャー」)を属性値と呼ぶ。

- (1) 対話シナリオは、属性毎に作成する。対話シナリオは、ある属性の値を得るための対話ルールの集合である。
- (2) 対話全体は複数のテーマ(図2の二重枠(「性格」、「経歴」等))から構成される。
- (3) 各テーマは、そのテーマを代表する属性(代表属性)を持つ。代表属性に対して質問の優先順位を設定することで、基本的な対話の流れを決めることができる。
- (4) 全ての属性は、テーマ別に分類される。なお、属性には代表属性の他に、深掘り質問のための属性(サブ属性)がある。
- (5) 属性値は、インスタンス型、ユーザ発話を標準化したストリング型、数値型、時間型の4種類がある。インスタンス型の属性値には、オントロジーを割り当てることで、そのオントロジーの値がセットされる。
- (6) 対話シナリオには、意図解析結果によって、さらにその属性について質問したり、関連する属性のシナリオを呼び出したり、シナリオの優先順位を変更するルールが記述できる。

テーマ	シナリオ名 (属性)	関連属性	応答開始条件	応答開始結果					
				開始応答文	応答ルール				
					応答条件			応答結果	
					意図	属性	属性値	応答文	応答アクション
転職理由	転職理由	転職理由	-	今回、なぜ転職したいと思われましたか？	ATTRIBUTE	転職理由	休日	お休みが少ないのでしょよね。	call(現在の勤務体系)
		会社の不満			ATTRIBUTE	転職理由	年収	年収に関して問題があるのですね。	call(現在の年収)

図3 一般シナリオの例

3.2. 対話シナリオ

本システムにおける対話シナリオとはユーザとシステムとの対話のルールを記述したものである。

対話シナリオには、属性を1つの単位とするシナリオの他に、ユーザ発話の意図を単位とするシナリオが存在する。前者を一般シナリオ、後者を特別シナリオと呼ぶ。

a. 一般シナリオ

一般シナリオの代表的な記述項目をピックアップした例を図3に示す。一般シナリオは図3の2列目のシナリオ名(属性)を単位として記述する。3列目の関連属性とは、そのシナリオ使用時ににおいて意図解析により優先的に値が抽出される属性である。

ルール本体は応答開始条件と応答開始結果からなる。図3では応答開始条件は省略したが、①前シナリオの属性、②現時点での個人データの属性/属性値、③一つ前のユーザ発話で獲得された属性/属性値、の条件を記述することができる。

応答開始結果は、応答文(システム発話文)と複数の応答ルールからなる。図3の5列目の開始応答文は、本シナリオにおける最初の質問文である。なお開始応答文や9列目の応答文には、ユーザ発話から意図解析により得られた属性値やオントロジーの値を変数として組み込んで生成することができる。

応答ルールは、応答条件と応答結果からなる。応答条件には、開始応答文の回答として得られた YES,NO 等の意図や属性/属性値を記述し、応答結果には、その回答が得られた場合のシステムの応答文と次のアクション(応答アクション)を記述する。応答アクションには、①さらにこの属性に関する質問を続ける、②別のシナリオを呼び出す、③シナリオ優先順位を変更する、④終了する、等の処理を記述することができる。

b. 特別シナリオ

一般シナリオは、主にマッチング時に利用されるユーザ情報をトリガーとするのに対して、特別シナリオは、ユーザ発話の意図解析結果の属性以外の意図をトリガーとする。例えば

Sys: 希望される職種を教えてください。

User: テクニカルマーケティングはどんな仕事ですか？

というようなユーザからの想定外の質問は一般シナリオでは対処できない。これに対処するのが特別シナリオである。意図解析結果から「EXPLAIN: テクニカルマーケティング」という意図を受け取り、特別シナリオがこの意図を受けて、システムは

Sys: 「テクニカルマーケティング」とは、製品を使ったデモンストレーションやプロトタイプを作ることで、製品のプロモーションを技術の観点からサポートする仕事です。

という応答をする。特別シナリオには、EXPLAIN(説明要求)の他、意図解析が出力する PASS(無回答)、QUIT(終了希望)、DAMN(罵り)、MAYBE(曖昧)に関するシナリオと、START(対話開始)、END(対話終了)シナリオの計7種類がある。

3.3. 対話制御方式

本システムにおけるユーザとの対話は、特別シナリオ内の START シナリオで開始され、予め決められたテーマの流れに沿ったシナリオ選択を基本として、シナリオ内のルールに基づいて進行する。ユーザ発話は、意図解析され、対話制御モジュールに渡される。意図解析結果に基づいて、本モジュールは次の5つのいずれかの処理を行う。

- (1) 一般シナリオのルールを継続する(現属性に関する質問を続けることで、対話の深掘りをする)。
- (2) 異なる属性の一般シナリオに遷移する(関連する属性の対話シナリオの質問を続けることで、対話の深掘りをする)。
- (3) 一般シナリオは継続せず、特別シナリオを呼び出す(ユーザ回答が、シナリオが想定する回答から逸脱した場合、特別シナリオを用いて対話する)。
- (4) ある属性に関する一連の質問は終了し、シナリオ優先順位に従って、別の属性のシナリオが続く。
- (5) ある属性に関する一連の質問は終了し、マッチングから新しく情報要求(属性)を受け取る。

対話制御モジュールは、1つのシナリオが終了する度に、対話ログにシステムとユーザの対話結果を書き込む。

本システムでは、(5)のように、次に質問する属性（どの属性に関するシナリオを用いるか）を決定する際、シナリオ内のルールだけでなく、マッチング結果も利用する。マッチングモジュールは検索を最適にする属性を対話制御に与え、対話制御はその属性のシナリオによって対話を継続することができる。また、対話の終了を判断するのもマッチングモジュールである。対話終了と判断された場合、特別シナリオ内の END シナリオによって対話を終了する。

4. 実証システム

4.1. 対話シナリオの構築

対話シナリオを作成するために、キャリアカウンセラー自身のノウハウに基づいて、「転職者向け紹介ドメイン」における全 12 の対話のテーマを設定した。この 12 テーマに対して、代表属性とその属性に関する質問である代表質問文を作成した。表 1 に選定した対話テーマの例を示す。

テーマ (代表属性)	代表質問
転職理由	どうして転職しようと思ったのですか？
仕事の経験	直近の仕事内容を教えてください。
仕事の実績	その仕事で評価されたことは何ですか？

表 1 対話テーマの例

さらに、カウンセラーは上記 12 テーマの代表質問に対して 10~20 個のユーザの想定回答を作成し、想定回答毎にその後続くことが想定される対話サンプルを作成した。

次に、これらの全対話サンプルのユーザ発話から「属性/属性値」を抽出した。全対話サンプルから抽出された属性/属性値を整理し、最終的に 99 種類の属性を決定した。また、これらの属性から、対話シナリオ構築に必要なオントロジー（例：転職理由オントロジー、会社の特徴オントロジー）を決定し、ここで抽出された属性値は、各オントロジーの値として登録した[6]。

99 種類の属性は、予め設定した 12 のテーマに分類され、属性毎に対話シナリオを作成した。なお、対話シナリオは、その作成及び変更を容易にするため、図 3 をより詳細化した Excel のファイルに記述し、VBA プログラムで対話エンジンが読み込む XML ファイルに変換する。

対話シナリオの基本ルールは、システムからユーザへの質問文と、ユーザの想定回答とその回答が得られた場合のシステムの次の動作（さらなる質問等）であ

る。ユーザの想定回答とは属性値（その属性に設定されたオントロジーの値）と意図解析結果から出力される意図であるので、各属性値、意図を自動的付与した雛形を作成して、最初に全属性に対してベースとなるシナリオを作成した。

次に、個々のシナリオに対して、ラダリング対話を実現するための動作を人手で追加し、対話の流れが自然になるように、個別に対話シナリオをチューニングした。例えば、一つ前のシナリオによって、質問の仕方（質問文の表現）を変えたり、ユーザの回答によって、次に質問する属性を変えたりした。

さらに、チューニング後はシステム試験の結果に基づいて、シナリオの改良を行い、最終的な実証実験用のシナリオを構築した。

現時点における一般シナリオの登録数を表 2 にまとめる。

シナリオ数	99
応答開始条件数	170
応答ルール数	5,613
属性に関する質問をさらに続ける	46
シナリオ優先順位を変更する	516
別のシナリオを呼び出す	3,892
その他（シナリオ終了等）	1,159

表 2 一般シナリオの登録数

今回、カウンセラーは対話サンプル（対話の元原稿）を作成し、システム開発者が対話シナリオ（Excel ファイル）を作成した。カウンセラーは、その他にも、質問文やシナリオ優先順位等の確認、修正を行ったり、テストサイトを使って全体の対話の流れが適切かを確認し、システム開発者へフィードバックすることで、シナリオ開発をサポートした。

4.2. 実証実験における対話例

以下に、実証実験[6]で得られた対話ログで、ユーザが独自では表現できなかった情報を引き出している例を表 3 に示す。

表 3 の対話例では、対話開始時にはユーザが自分では質問に対して明確に答えることができなくても、システムがユーザの回答に応じて、ユーザにとって答えやすいようにその質問を具体化したり、問題を複数の質問に分けて答えやすくすることで、ユーザから深い情報を引き出すことに成功している。

興味深い点は、対話ログを実際のキャリアカウンセラーに見せたところ、「この人には、小さな店舗の副店長の仕事を紹介する」と判断できた点である。これは、開発システムとの対話によって、ユーザから転職先紹介に関する必要な情報を引き出した良い例であろう。

シナリオ名	システム発話文	ユーザ回答	取得属性
仕事の経験	今どういったお仕事をされていますか？	書籍販売, 図書館司書をしていました.	現在の職種: 図書館司書 所有する技能: 書籍
仕事の実績	わかりました. 書籍, 音楽, 販売の経験をお持ちなのですね. 現在あるいは過去で, 仕事で評価されたことや褒められたことを教えてください.	まじめな勤務, 接客態度.	仕事上でのあなたの強み: まじめさ 所有する技能: 接客
:	:	:	:
希望する仕事	どのような仕事を希望しますか？	まだよくわからない	希望する仕事: 不明
仕事上で強みが発揮される時	よくわからない, ということですね. では, 仕事をする上での強みや何か得意分野について, すでにお話しいただいています. ではそうしたあなたの強みが発揮された結果, 何がどう変わりましたか？	お客が親しみを持って来店してくださるようになった.	仕事上で強みが発揮される時: お客が……
:	:	:	:
仕事上での課題	あなたが仕事をする上で苦手と思っていたり, 自分自身の課題だと考えていることがありますか？	売上など, 実際のデータを分析し仕事に活かすこと.	仕事上での課題: 売上など, 実際の……
課題を指摘する人	そうですか. あなた自身の苦手な分野や課題を, あなたの周りで指摘してくれる人はいますか？	いない.	課題を指摘する人: なし
課題が生じる理由	ではその苦手分野や課題は, どういった理由で生じていると思いますか？	やり方がわからないから.	課題が生じる理由: やり方が……
課題を解決したいか	苦手分野や課題を, あなた自身は解決したと思いますか？	思う.	課題を解決したいか: 思う
課題の改善方法	わかりました. それでは苦手分野や課題に対して, どうしたら克服, あるいは改善していけると思いますか？	情報収集する	課題の改善方法: 勉強する
	勉強するということですね.		

表 3 実証実験における対話例

5. まとめと今後の課題

本稿は, ラダリング型検索サービスシステムについて対話エンジンを中心に説明した. さらに, 実証実験の対話ログで, 開発システムが, ユーザが単独では表現できなかった求めるサービスやコンテンツに関する情報を引き出した例を紹介した.

一方で, 実証実験の結果, 実際のサービスを行う際の課題が明らかになった. 第一はシナリオ開発の工数の問題である. 今回対話シナリオの作成はシステム開発者が行ったが, 実用システムではシステム開発者ではなく, カウンセラーやドメインのエキスパートがシナリオを開発できなければならない. 今後, 開発ツールを整備したり, ベースとなる対話シナリオを用意して人間が記述すべき部分は最小限にしたりする工夫が必要となる. また, ステレオタイプのシステムへの応答は機械学習等を用いてユーザの意図を判断し, 質問を自動生成する[4,8]試みも有用である.

第二の課題は対話品質である. 対話品質の向上のためには, 現在の対話シナリオの記述能力, 制御機能を強化するだけでなく, 意図解析の解析能力の向上, オントロジーとして有する言語知識の質及び量の向上, さらなるマッチングとの連携が必要となる. 一方, 実証実験で得られた, ユーザの対話ログも対話品質向上のための有益なデータになると考える.

本研究は, 経済産業省平成19年度情報大航海プロジェクト「ラダリング型検索サービスの研究開発」の一環として, 株式会社リクルート殿と共同で行われた.

参考文献

- [1] 清田陽司, 黒橋嶺夫, 木戸冬子, “大規模テキスト知識に基づく自動質問応答-ダイアログナビ-”, 自然言語処理, Vol.10(4), 2003
- [2] K. Acomb, et al., “Technical Support Dialog System: Issues, Problems, and Solutions”, Bridging the Gap: Academic and Industrial Research in Dialog Technologies Workshop Proceedings, pp.25-31, 2007
- [3] 大塚裕子, 山本瑞樹, 乾孝司, 丸本聡子, 奥村学, “市民参画型道路計画における対話支援-対話型アンケートシステムプロトタイプ版の実装-”, 言語処理学会第13回年次大会発表論文集, pp.302-305, 2007
- [4] 丸元聡子, 鈴木泰山, 大塚裕子, 伊藤裕美, 乾孝司, 奥村学, “空港計画における対話型意見収集システムの実装と課題”, 言語処理学会第14回年次大会発表論文集, pp.5-8, 2008
- [5] M. Spitters, et al., “Learning to Compose Effective Strategies from a Library of Dialogue Components”, ACL 2007, pp.792-799, 2007
- [6] 下畑さより, 北村美穂子, 介弘達哉, 池野篤司, 坂本仁, 折原幾夫, 村田稔樹, “ラダリング型検索サービスのためのドメイン知識構築, 及び, 実証実験”, 情報処理学会研究報告 2008-NL-185, 2008
- [7] M. Kitamura and T. Murata, “Practical Machine Translation System allowing Complex Patterns”, Proceedings of MT Summit IX, pp.232-239, 2004
- [8] 鳥澤健太郎, “一般ユーザーにインタビューする対話エージェント”, 情報処理学会研究報告 2007-NL-180, pp.25-30, 2007