

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年5月10日現在

機関番号：14303

研究種目：基盤研究(C)

研究期間：2010～2012

課題番号：22500153

研究課題名（和文） オブジェクト指向的な音声対話システム開発手法の研究

研究課題名（英文） Studies on object-oriented development method of spoken dialogue system

研究代表者

荒木 雅弘 (ARAKI MASAHIRO)

京都工芸繊維大学・工芸科学研究科・准教授

研究者番号：50252490

研究成果の概要（和文）：本研究ではオブジェクト指向開発方法論に基づいた音声対話システム開発向けの対話記述言語の設計と、その処理系としてのフレームワーク開発を目的とした。研究成果として、MVC(Model-View-Controller)モデルに基づいた音声対話システムのコンポーネントを、セマンティック web の標準オントロジーを継承したデータモデル定義から自動生成する手法の開発に成功した。

研究成果の概要（英文）：The aim of this research is to design a new object-oriented dialogue description language for spoken dialogue systems and to implement a framework as an executable platform of the proposed language. As a result of the research, we developed a generation method of executable components of spoken dialogue system based on MVC (Model-View-Controller) model from the data model definition inheriting semantic web ontology.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	1,700,000	510,000	2,210,000
2011年度	1,200,000	360,000	1,560,000
2012年度	500,000	150,000	650,000
年度			
年度			
総計	3,400,000	1,020,000	4,420,000

研究分野：音声対話処理

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：(1) 音声対話システム (2) オブジェクト指向

1. 研究開始当初の背景

音声対話システムの開発・移植の効率向上に関する研究は、対話コーパスからの(あるいは対話シミュレータを用いた)対話状態遷移規則の機械学習に関する研究と、音声認識に用いる統計的言語モデルを少量のタスク内コーパスを用いて適応させる研究が主たるアプローチで、システム全般を開発するプロセスに関しては、CMUのLet's Goプロジェクトなど一部を除いて、取り組みが遅れて

いるのが現状である。

前述のLet's Goプロジェクトや、CSLU Dialogue Toolkitが提供しているものは、近年のソフトウェア工学的な見方で位置づけると、最も表層レベルのツール群であり、その土台となるべき手法や開発プロセスの検討は不十分である。一方、近年のWebアプリケーション開発においてはソフトウェア工学分野で確立されているオブジェクト指向開発が主流となりつつあるが、音声対話シ

システムにおいては、文法・言語モデル・意味解析規則・ユーザモデリング・対話制御など、その開発において特殊な知識や技術が必要であり、いまだに開発方法論のソフトウェア工学的な検討には踏み込めていない。

2. 研究の目的

本研究では、音声対話システムの開発に適したオブジェクト指向の開発プロセスを確立することを目指す。この開発プロセスの特徴は、(1)音声認識文法・意味解析規則とデータ処理を直結し、カプセル化することによる再利用性・保守性の向上、(2)データモデルからの対話パターンの自動生成による開発効率の向上である。具体的には、継承を有効に利用した対話記述言語およびその処理系の開発と、データモデルからインタラクション記述までをカバーする統合開発環境の実装を目標とする。

3. 研究の方法

(1) アーキテクチャの選択

音声対話システムの開発方法論を考案するにあたり、今後の携帯情報端末への展開を考慮すると、音声入出力のみをモダリティとして考えるのではなく、マルチモーダルインタフェース (MMI) への対応を前提としたアーキテクチャが必要になる。既存の MMI システム開発手法の問題点から、我々は MMI システム開発方法論の要件として以下の 3 件を重視する。

- MVC の明確な分離
- GUI ベースの Web アプリケーション開発方法論との共通性の確保
- モダリティの追加・変更等に対する保守性・機能拡張性の維持

W3C の提案する MMI アーキテクチャでは対話管理部とデータモデルの結合が未定義であり、MIT の Galaxy-II アーキテクチャでは MVC 記述がハブスクリプトに混在可能であり、また MVC いずれの要素の変更に対してもハブスクリプトを書き換える必要が生じるため、保守性・機能拡張性が低くなってしまふ。これらと比較して、ITSCJ MMI システムアーキテクチャは明確に分離された 6 階層で構成され、上記 3 点の要件を満たすことから、本研究ではこのアーキテクチャに沿った開発を行う。図 1 に ITSCJ アーキテクチャの構成を示す。

(2) オブジェクト指向の導入

MMI システムは人間を含んだ系であるため、開発過程において機能の追加や修正は不可避である。従来の CSLU Toolkit や XISL、商用の音声対話システム記述言語 VoiceXML は、状態遷移に基づく設計がなされている。状態遷移モデルは機能の実現に状態遷移を用いているため、新たな検索手段の追加など、

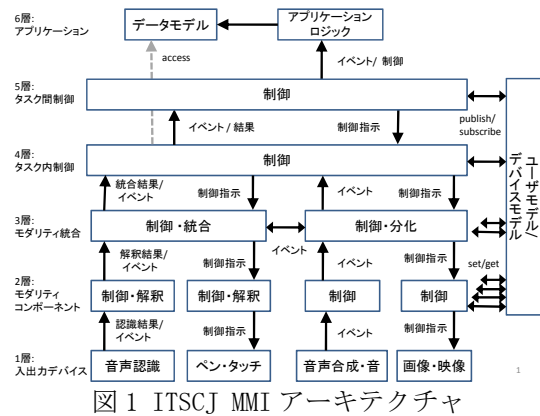


図 1 ITSCJ MMI アーキテクチャ

```

[SystemInitiative
@DBSearch
class MyBook extends Book {
  int price
  static constraints = {
    name( onsearch:"ilike" )
    author.name( onsearch:"ilike" )
    thumbnailUrl( )
    price( )
    isbn( )
  }
}

```

図 2 MrailsScript の例

機能を追加・修正するためにはシステムの基本構成である状態遷移そのものを変更する必要がある。一方で、オブジェクト指向設計はデータモデル定義を軸に機能を付加していくため、機能の追加・修正に対して柔軟に対応できる。オブジェクトベースで対話システムを構築するフレームワークには Queen's Communicator があり、オブジェクトの差し替えによる機能追加・修正が可能だが、対話制御を特権的なオブジェクトである DiscourseManager に集中させるため、対話の流れを制御するコードが複雑になるという欠点がある。

本研究ではオブジェクト指向設計に基づいた Grails フレームワークを基礎とし、データモデル定義記述を MMI システム開発の中心的記述と位置づける。

4. 研究成果

(1) 対話記述言語 MrailsScript

Grails ドメインクラス定義を基礎として、対話のためのタスクと主導権の設定をアノテーションで付加し、またオントロジーとして採用した Schema.org クラスの継承を行うことでセマンティクスを取り入れた、対話記述言語 MrailsScript を設計した。MrailsScript はタスク分類・主導権を決定するアノテーションと、データモデル定義からなる。

図 2 は設計した MrailsScript によって記

述した本の検索システムの例である。このシステムはタイトルあるいは著者名による検索をシステム主導型対話で行う。

MMI においては、例えば音声によってタスクを誘導するケースなどにおいて、タスクの種類や主導権に応じて適切に対話をコントロールする必要がある。そのため、タスク・主導権の分類に基づき対応するテンプレートを用意し、アノテーションによって指定する方法を取った。

対話タスクは、システムとユーザの間を流れる情報の方向という観点から以下の3つに分類できる。

- @Explanation 説明タスク
道案内やチュートリアルなど、システムからユーザへ情報を提供するタスク。
- @Slotfilling スロットフィリングタスク
アカウント取得のための個人情報の入力など、ユーザがシステムへ情報を入力、提供するタスク。
- @DBSearch 検索タスク
時刻表の検索やトラブルシューティングなど、ユーザが検索条件を入力し、システムが結果を返すタスク。

MrailsScript では、以上のタスクをアノテーションで指定し、それぞれのプロパティに必要な条件を、データモデル定義の制約規則中に記述する。検索タスクでは図2の6行目と7行目のように、検索に使用する要素と制約規則を onsearch 属性として追加し”ilike” (部分一致), ”eq” (全一致) などの検索基準を指定する。スロットフィリングタスクでは入力が空であることを許す nullable 属性に真偽値を指定する。現段階では典型的な3つのタスクを実装した。対話の主導権は、達成するべき目標や想定されるユーザによって、以下の3つに分類できる。

- @SystemInitiative システム主導
システムが決められた手順に沿って、ユーザに順次情報の入力を求めることでタスクを達成する形式。
- @UserInitiative ユーザ主導
ユーザが自由な順序・まとまりで情報を入力し、タスクを達成する形式。
- @MixedInitiative 混合主導
ユーザとシステムの間で主導権が入れ替わりつつタスクを達成する形式。

入力要素の一部が空であった場合の動作、全ての入力要素が埋まった場合の挙動、などを3つそれぞれに個別に設定した。タスクの場合と同様に、プロパティ中に、そのプロパティがどのような手続きで動作するか条件を追加することも可能である。

(2) MMI 開発フレームワーク

MrailsScript で指定された要素によって拡張された Grails のドメインクラス・ビュー・コントローラを生成する、Mrails フレームワークを実装した(図3)。

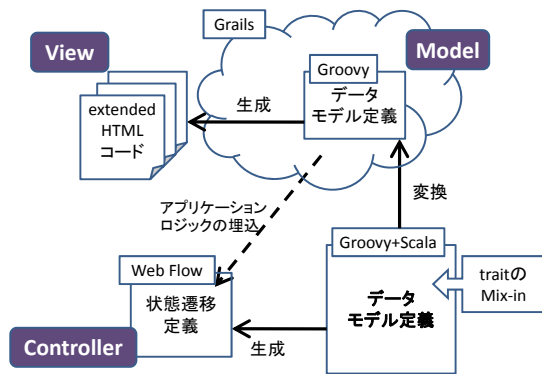


図3 MVC 実行コードの生成

データモデルは、MrailsScript から Grails 標準のドメインクラスを生成した。対話記述であるタスクと主導権のアノテーション及び継承の宣言を取り除き、継承した Schema.org クラスのオントロジー名に基づいたフィールド宣言と Grails 制約規則を自動生成する。なおプロパティ型の制約は、一般的なデータ型を網羅している HTML5 の input 要素の type 属性による分類を使用し、後述のビューコードと対応させた。

コントローラコードは、タスクアノテーションに基づいて、ビューを適切に制御するために必要なアクションを定義したコードとして生成する。予め定めたフローに従って、データベース検索タスクでは、クエリの入力、検索結果の一覧、詳細表示の順に画面が遷移する。クエリの入力画面では、MrailsScript の制約規則中の onsearch 属性に基づいて、ビューから送られてきたリクエストを元にデータベースへの検索クエリを生成する機能を生成する。

ビューコードの生成では、主導権アノテーションに基づいて、各々のプロパティについて値の取得順序や方法を定義したコードを生成する。本実装では、Grails ビューコードを拡張し音声インタラクションを可能にしたものを MMI に対応したビューコードと位置づける。データベースに対しての基本操作のみを対象とした標準の自動生成機能に対して、主導権設定と音声モダリティ依存の情報付加したビューコードを自動生成した。

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 9 件)

① 竹腰 大輔, 荒木 雅弘: データモデル定義に基づくマルチモーダル対話記述言語の設計と実装, 人工知能学会論文誌, 査読有, Vol. 28, No. 3, pp. 261-266, 2013.

DOI:

<http://dx.doi.org/10.1527/tjsai.28.261>

② Masahiro Araki: Multimodal Motion Learning System for Traditional Arts, In Proc. AHFE2012, 査読有, pp. 5274-5281, 2012.

③ Masahiro Araki and Daisuke Takegoshi: A Rapid Development Framework for Multilingual Spoken Dialogue Systems, In Proc. COMPSAC 2012, 査読有, pp. 195-202, 2012.

④ Masahiro Araki: Rapid Development Process of Spoken Dialogue Systems using Collaboratively Constructed Semantic Resources, In Proc. SIGDial 2012, 査読有, pp. 70-73, 2012.

⑤ Masahiro Araki and Daisuke Takegoshi: Framework for the Development of Spoken Dialogue System based on Collaboratively Constructed Semantic Resources, In Proc. NAACL-HLT Workshop, 査読有, pp. 25-28, 2012.

⑥ Masahiro Araki and Yuko Mizukami: Development of a Data-driven Framework for Multimodal Interactive Systems, In Proceedings of IWSDS 2011, 査読有, pp. 91-101, 2011.

⑦ Masahiro Araki, Kana Shibahara and Yuko Mizukami: Spoken dialogue system for learning Braille, In Proceedings of IEEE COMPSAC 2011. 査読有, pp. 152 - 156, 2011.

⑧ Masahiro Araki and Yu Funakura: Impact of Semantic Web on the Development of Spoken Dialogue Systems, In Proceedings of the Second international conference on Spoken dialogue systems for ambient environments (IWSDS 2010), 査読有, pp. 144-149, 2010.

⑨ Masahiro Araki and Takashi Hattori: Proposal of a Practical Spoken Dialogue

System Development Method: Data-management Centered Approach, In W. Minker et al. (eds) 査読有, Spoken Dialogue Systems Technology and Design, Springer, pp.187-211, 2010.

[学会発表] (計 5 件)

① 西牟田勇哉, 荒木雅弘: セマンティック Web のための音声質問文の検索クエリへの変換, 第 75 回情報処理学会全国大会, 4N-6, 2013 年 3 月 7 日, 仙台市.

② 小澤克貴, 荒木雅弘: マルチドメイン構造を用いたセマンティック Web 検索対話システムの提案, 情報処理学会研究報告, 2013-SLP-95(2), 2013 年 2 月 1 日, 熱海市.

③ 竹腰大輔, 荒木雅弘: データモデル定義に基づくマルチモーダル対話記述言語の設計と実装, 2012 年度人工知能学会全国大会論文集, 102-OS-18-6, 2012 年 6 月 12 日, 山口市.

④ 荒木雅弘: オブジェクト指向的な対話システム開発言語の提案, SIG-SLUD-B102-04, 2011 年 10 月 13 日, 東京都.

⑤ 荒木雅弘: Rails フレームワークを用いたマルチモーダル対話システム, SIG-SLUD-B002-05, 2010 年 10 月 28 日, 東京都.

[図書] (計 1 件)

① 荒木雅弘, 森北出版, フリーソフトで学ぶセマンティック Web とインタラクション, 2010, 198.

6. 研究組織

(1) 研究代表者

荒木 雅弘 (ARAKI MASAHIRO)

京都工芸繊維大学・工学科学研究科・准教授

研究者番号: 50252490

(2) 研究分担者

()

研究者番号:

(3) 連携研究者

()

研究者番号: