

音声対話管理システム CAMMIA の開発

○畑岡信夫、大淵康成(日立中研) 赤堀一郎、△立石雅彦(デンソー基礎研)
△Scott Judy、△Jeongwoo Ko、△Teruko Mitamura、△Eric Nyberg (CMU/LTI)

1 はじめに

音声技術の車載応用は、音声のキーアプリケーションであるとも言われ、耐雑音対策、システム構成等の研究開発が積極的に推進されている。車載情報サービスの世界では、安全性を考慮した観点から、音声によるヒューマン・マシン・インタフェース(HMIs: Human Machine Interfaces)が、「チャレンジング」な試みとして期待されている。特に、音声対話は、効率的なHMIsを構築する上で重要性が高まっている。

我々は、VoiceXMLを拡張した対話管理方式CAMMIA (Conversational Agent for Multimedia and Mobile Information Access)を提案し[1][2]、車載情報サービスを応用とした評価プロトを構築した。本稿では、具体的な対話シナリオの生成と評価プロトに関して報告する。

2 音声対話システム CAMMIA の構成

図1に、我々が提案している対話管理方式CAMMIAの車載情報システムでの応用構成を示す。CAMMIAは、対話シーケンスの表現能力と共に、タスク遷移(例えば、ルート計画や各種案内のタスク等)に追従する柔軟な対話管理となっている[1][2][3]。

端末側の中心となるのはVoiceXMLインタプリタ(VXI)で、音声認識(ASR)と音声合成(TTS)を通してユーザとのやり取りを行う。VXIは、端末内での他のアプリケーションとの非同期通信を行うための入出力モジュールを持っている。この結果、GPSをベースとしたルートガイダンス(図1ではNavigation Simulator)からの割り込みにも対応することが可能となっている。

センタ側では、対話タスクの遷移に追従した対話シーケンスを記述したVoiceXMLの自動生成を可能とする対話管理(DM)方式を考案した。

Development on Speech Dialog Management System CAMMIA

N.Hataoka, Y.Obuchi (Cent. Res. Lab., Hitachi Ltd.)
I.Akahori, M.Tateishi (DENSO CORPORATION)
S.Judy, J.Ko, T.Mitamura, E.Nyberg(Carnegie Mellon U.)

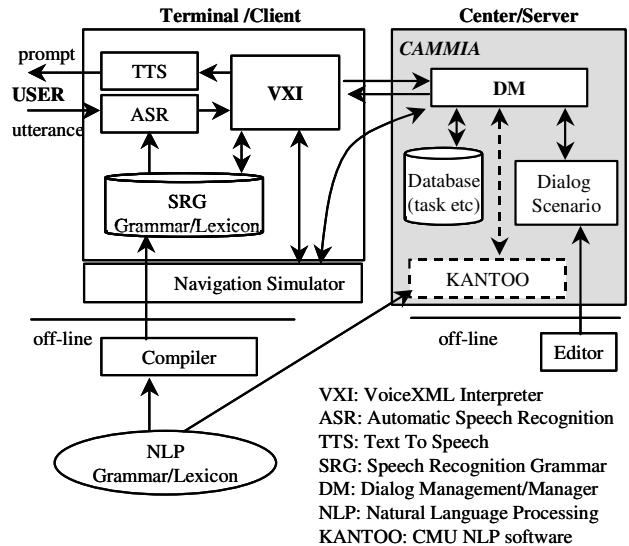


図1 CAMMIA 全体システム構成

この結果、対話タスクが遷移した場合も、現状のタスクを一時保持しながら、次のタスクへと遷移し、その後、元のタスクへ戻る事が可能となっている。さらに、対話コンテンツの設定は、シナリオ設定エディタを用いる事により、音声認識、VoiceXML等の詳細知識を持たなくても容易に構築する事が可能となっている。

3 対話遷移に追従する対話管理方式

対話管理方式は、VoiceXMLの自動生成を可能とするために、VoiceXMLを拡張して、ScenarioXML、DialogXMLを上位概念として持つ3階層構成となっている[1][2]。ScenarioXMLは、対話タスクを表現し、DialogXMLは具体的な対話内容を表現する。システムはDialogXMLを経由して生成されたVoiceXMLにて記述された対話シーケンスを実行する[3]。

4 文法と語彙の設定と対話シナリオの生成

4.1 文法と語彙の設定

具体的には、車載情報サービスにおけるルートガイダンス(道案内)、観光案内、施設案内等のタスクを想定して、音声コーパスを解析して文法と語彙を設定した。現在は、人手での作成となっているが、将来はKANTO[4]での単一化文法をベースに、音声認識で使用可能な簡略化

された正規化文法（ネットワーク文法）への自動変換する方式を開発している[3]。

4.2 対話シナリオの生成環境

コンテンツ提供者が、対話内容を設定し、容易に対話サービスシステムを構築できる事は、音声対話システム普及の観点から非常に重要な機能である。本稿では、図2に示すように、テンプレート・ベースの入力から ScenarioXML を生成し、最終的に VoiceXML を自動生成する環境を構築している。処理の手順は、

- ① create a dialog template;
- ② create a ScenarioXML draft for the template;
- ③ edit/refine the ScenarioXML;
- ④ compile ScenarioXML to VoiceXML (via DialogXML).

図3には、天気案内タスクでの対話テンプレート (dialog template) の例を示している。現在、作成効率の評価と GUI ベースの環境（ツール）への移行を行っている。

5 車載応用プロトと予備評価結果

5.1 プロトタイプ・イメージ

図4に車載情報サービスを応用とした評価プロトの車載機画面イメージを示した。JR 名古屋駅を出発地として、目的地を名古屋城にし、途中で駐車場情報や各種レストラン、施設の情報を入手する想定で、操作が進められる。

5.2 処理時間・メモリ規模の評価

天気案内をタスクとして、タスク読み込み時間と生成される VoiceXML サイズを測定した結果、状態遷移をメモリに格納しない初期プロトではタスク読み込み時間 4271.25ms、VoiceXML サイズ 2873.83byte、メモリ格納を効率的に行う現プロトでは各々 1198.5ms (削減率 71.94%)、578byte (削減率 79.89%) であった。この結果、

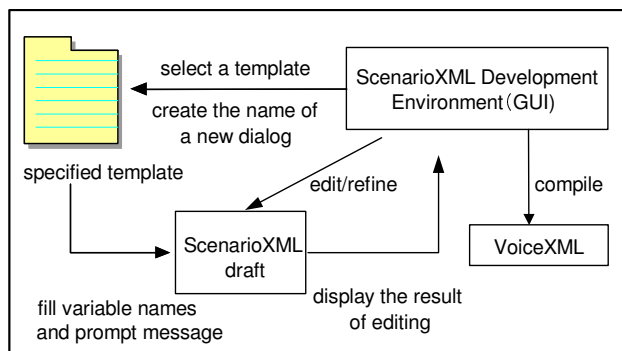


図2 ScenarioXML 構築の環境構成

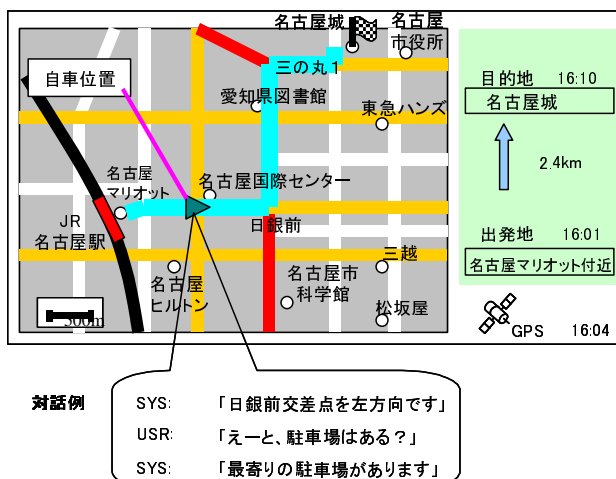


図4 車載機の動作イメージの例

対話管理処理での効率的なメモリ利用により、大幅な処理時間と端末へ配信する VoiceXML 容量の削減が可能となることが分かった。

6 まとめ

対話タスクの遷移に追従できる音声対話方式を提案し、評価プロトを完成した。さらに、対話シナリオを容易に生成できる環境を構築した。

参考文献

- [1] E.Nyberg, T.Mitamura, P.Placeway, M.Duggam and N.Hataoka, DialogXML: Extending Voice-XML for Dynamic Dialog Management, *Proc. of HLT-2002* (2002.3)
- [2] 畑岡信夫、他、VoiceXML をベースにした音声対話管理方式の開発、音講論 2-9-12 (2002.9)
- [3] Y.Obuchi, et al., Robust Dialog Management Architecture using VoiceXML for Car Telematics Systems, *Proc. of WS on DSP in Vehicular and Mobile Systems* (2003.4)
- [4] E.Nyberg, T.Mitamura, *Proc. of AMTA* (2000)

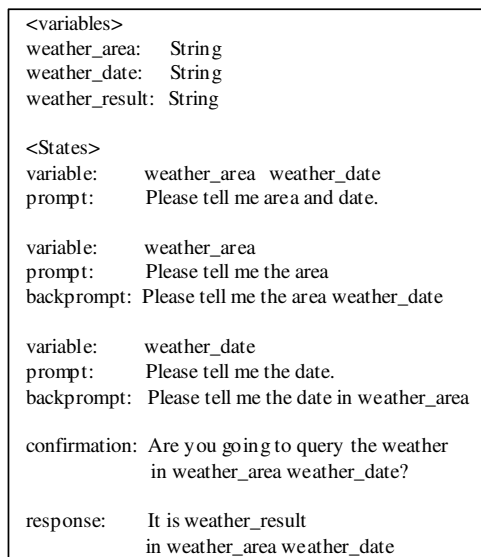


図3 Weather Dialog Template の例