

ゆかりプロジェクトにおける分散環境行動 DB と ロボット対話インタフェースの連携メカニズム

佐藤 淳[†] 上田 博唯[‡] 木戸出 正継[†]

[†]奈良先端科学技術大学院大学 情報科学研究科

[‡]独立行政法人 通信総合研究所 分散協調メディアグループ

E-mail: {jun-sat, kidode}@is.aist-nara.ac.jp
hiro-u@crl.go.jp

1. はじめに

ゆかり (Universal Knowledgeable Architecture for Real-Life appliance) プロジェクト[1]では、家庭内における各種センサによって得られるデータと、ネットワーク化された家電機器から得られる内部状態情報などに基づいた分散環境行動 DB というフレームワークを提案している[2]。このフレームワークで得られるコンテキストを利用した高機能サービスを構築するためには、ロボットとユーザの対話によるユーザインタフェースが1つの重要なキー技術となる。本稿では、この分散環境行動 DB とロボット対話インタフェースが連携して、ユーザからの曖昧な要求をより正確に受け取ってシステムを制御し、同時にユーザに家庭内にある各種機器の状況をよりの確に理解してもらうことができるメカニズムについて提案する。

2. 音声対話システム

筆者らは、擬人化音声対話エージェントのツールキットである galatea[3]を使用した音声対話システムを計画している。これを図 1 に示す。点線で囲んだ部分が galatea である。このシステムでは、ユーザ発話をワードスポットティングによって一度単語列に分解する。そして、得られた単語列に属性を付加する。例えば、「テレビ」であれば<放送型メディア再生>、「阪神」であれば、<野球チーム名>などである。このような単

語属性の組み合わせからロボットの行動を決定する。ここで、ユーザ発話が行われてからロボットが行動するまで順を追って説明する。まず、ロボットに対してユーザ発話が行われると、ロボットはユーザ発話を受け取り、入力音声として SR-M に引き渡す。この時 SR-M は GSR-DB を参照する。GSR-DB には音声認識されるべき全ての文が記述されている。入力音声は GSR-DB に記述されている中で最も適合する文に変換される。SR-M によって得られた入力文は、WD-SPT-M に引き渡され、抽出単語列として出力される。WD-SPT-M は WD-SPT-DB を参照する。ここには、抽出単語列に与えるべき属性が記述されている。もし、ユーザが「阪神」と発話すると「阪神」<野球チーム名>が返ってくる。この抽出単語列(属性付)を DPS に渡す。DPS は RAR-DB を参照する。RAR-DB には、抽出単語列(属性付)が得られた場合に、ロボットがとるべき行動を記述しておく。ユーザとロボットの対話例と、RAR-DB に記述されている規則例を示す。

例1 U: ユーザ R: ロボット

U: 「阪神が見たい」

R: 「19:00 から 4ch で放送するよ」

RAR に記述されている規則例

```
if ユーザ発話に<放送型メディア再生>かつ
<野球チーム名>が含まれる
then TV 番組表(EPG)を検索
    if EPG に<野球チーム名>が存在
    then 「<放送時間>から<ch>で放送するよ」
        と発話
```

VC-DB はユーザ語彙に関する多義性の辞書となる。例えばユーザが「阪神」と発話したとき、「阪神タイガース」のことが「阪神百貨店」のことに

The cooperation mechanism for a robot dialogue interface and a distributed database of environment and human actions in UKARI project

[†]Jun SATO, Masatsugu KIDODE

GraduateSchool of Science, Nara Institute of Science and Technology

[‡]Hirotsada UEDA

Distributed and Cooperative Media Group, Communication Reserch Laboratory

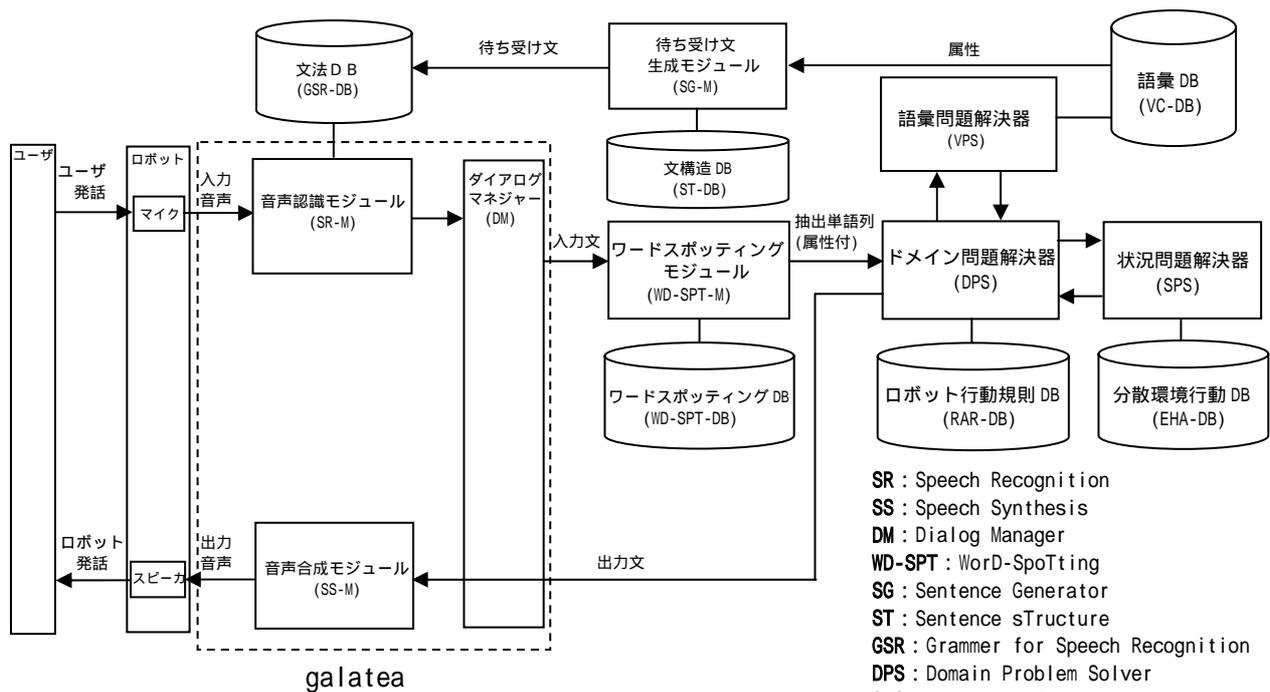


図1 音声対話システム図

よってロボットの行動戦略は変わってくる。語彙情報は DPS が VPS に問い合わせることによって得られる。VPS は VC-DB を参照して DPS に答えを返す。GSR-DB に記述する待ち受け文を自動生成するために、ST-DB には待ち受けるべき文構造を記述する。例えば、「阪神が見たい」の場合であれば「<野球チーム名>が<放送型メディア再生>」という文が記述されている。ここで<野球チーム名>、<放送型メディア再生>の部分に入れる単語を記述しなければならない。そのため、VC-DB には単語の属性も記述する。

EHA-DB に問い合わせるのは SPS である。DPS は必要に応じて SPS をも利用する。今回のシステムでは、DPS が問題解決者の核となり、他の問題解決者に命令を下している。EHA-DB には、ユーザ情報、環境情報などが記述されている。ユーザ発話によっては、これらの情報を使用することもある。対話例を下に示す。

例2 U : ユーザ R : ロボット

状況 : 例1の直後。テレビの電源をONにした

R : (EHA-DB の<お父さんの好み>参照)

R : 「お父さんも阪神が好きだよ」

U : 「お父さんもきっと見たがるよね?」

R : 「お父さんは書斎にいるよ」

この例はユーザの問いに対し、ロボットが EHA-DB に記述されている<お父さんの居場所>を

参照することでユーザの質問に答えている。以上のような動作を組み合わせることによって、例えば母親の例1の「阪神見たい」という発話から TV 番組を検索。TV の電源を ON。例2へ続き、父親も阪神ファンであることを知り、EHA-DB から父親の居場所を検索して、書斎にいる父親に TV で放送されていることを知らせる、というような一連の動作をロボットに行わせることができる。

このように EHA-DB の情報を使用することで今までに無いサービスをユーザに提供する事が可能となる。

3. おわりに

本稿では、UKARI プロジェクトにおける音声対話システムを提案した。今後は、この音声対話システムを用いて、ゆかりプロジェクトで提案されている各種サービスと結びつけ、検証していく。

5. 参考文献

- 1) 美濃導彦、「ゆかりプロジェクトの目的と概要-UKARI プロジェクト報告 No.1 --」、情報処理学会第 66 回全国大会(2004)
- 2) 桶上義彦他、「ゆかりプロジェクトにおける分散環境行動データベースの構成法に関する検討」、情報処理学会第 66 回全国大会(2004)
- 3) <http://hil.t.u-tokyo.ac.jp/~galatea/galatea-jp.html>