

Title	音声インターフェイスを用いたJAIST内PHSナビゲーションシステムの構築
Author(s)	渡辺, 功
Citation	
Issue Date	2000-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/661
Rights	
Description	Supervisor: 藤波 努, 知識科学研究科, 修士

修士論文

音声インターフェイスを用いた JAIST 内 PHS ナビゲーションシステムの構築

指導教官 藤波 努 助教授

北陸先端科学技術大学院大学
知識科学研究科知識社会システム学専攻

渡辺 功

2000 年 3 月

要旨

本論文では，Japan Advanced Institute of Science and Technology (JAIST) 内において，来訪者，および構内の情報を欲する人にナビゲートをするシステム「CommNavi」について述べる．

JAIST 内では，研究者の情報，構内の情報をリアルタイムに引き出すという行為が難しいという問題がある．この問題は，情報を欲する来訪者に対しても当てはまり，早急な解決が望まれている．

よって，本研究では問題解決を主眼においたシステム「CommNavi」を構築した．この CommNavi は，構内に設置されている PHS (Personal Handyphone System: PS 端末) を介した音声で対話ができ，その結果，ユーザは欲する情報を手に入れることができる．

実験は，道案内時における情報提供，対話の構築のあり方に重点をおいた．また，評価方法は，定性的評価（プロトコル分析，アンケート）と定量的評価（道案内時におけるユーザの所要時間）を用いた．

その結果，PS 端末を用いた道案内，およびユーザの情報取得を早める対話の構築の有効性が示せた．

目次

1	序論	1
1.1	研究の背景	1
1.1.1	研究者における背景	1
1.1.2	来訪者における背景	2
1.2	目的	2
1.3	関連研究	3
1.3.1	個人化戦略	3
1.3.2	位置情報を利用したツール	4
1.3.3	音声情報に関する媒体	5
1.3.4	音声メディア	5
1.3.5	対話	6
1.4	本研究の位置付け	7
1.5	本論文の構成	8
2	システム構成	9
2.1	CommNaviの特徴	9
2.1.1	音声インターフェイスの導入	9
2.1.2	ナビゲーション	9
2.2	ハードウェアの構成	11
2.2.1	PS	12
2.2.2	CS	13
2.2.3	PBX(Positioning Server)	13
2.3	ソフトウェアの構成	13

目次

2.3.1	音声認識と音声合成	14
2.3.2	ナビゲーションエンジン (Navigation Engine)	15
2.3.3	位置管理モジュール (Position Manager)	15
2.3.4	スケジュール管理モジュール (Schedule Manager)	16
2.4	データベースの構成	16
2.4.1	個人情報 (Personal DB)	16
2.4.2	位置情報 (Personal Position DB)	16
2.4.3	CS を中心とした構内マップ情報 (CS Map DB)	17
3	実行方法	18
3.1	対話における情報取得方法	18
3.2	会話の構築	20
4	実験	28
4.1	実験の評価方法	28
4.2	実験 1: 時間毎の情報提供	29
4.2.1	概要と目的	29
4.2.2	結果	30
4.2.3	考察	31
4.3	実験 2: ユーザに応答する情報提供	32
4.3.1	概要と目的	32
4.3.2	結果	32
4.3.3	考察	35
4.4	実験 3: ハンドオーバー時に情報を提供	35
4.4.1	概要と目的	35
4.4.2	結果	36
4.4.3	考察	38
4.5	実験 4: 最終実験	38
4.5.1	定量的評価	39
4.5.2	定性的評価	40

目次

4.5.3	考察	44
5	結論	45
5.1	本研究の成果	45
5.2	今後の課題	47
5.3	応用	48
	謝辞	49
	参考文献	50
	発表論文	53

目 次

2.1	ハードウェアの構成	12
2.2	ソフトウェアの構成	14
3.1	対話の分岐	19
3.2	対話の流れ	27
4.1	音声合成	30
4.2	Positioning Server の誤差情報	31
4.3	ルートの指示	33
4.4	位置補正	35
4.5	ユーザの問いかけ	36

表 目 次

1.1	対話の成立要因	6
1.2	対話形態の違い	7
3.1	図 3.2 の見方	20
4.1	本実験の評価方法	29
4.2	プロトコル分析の特徴	29
4.3	対話を始めて目標までたどり着いた時間	39
4.4	音声合成のストレスを感じたか?	40
4.5	音声認識のストレスを感じたか?	41

第 1 章

序論

1.1 研究の背景

1.1.1 研究者における背景

研究者は斬新な発想を生む能力とその発想を実現する能力が求められている。そして、この発想が促進されやすい条件の一つとして高密度な場の形成 [野中 96] を作ることがあげられる。高密度な場とは、専門的な知識を所有する者が集まり、お互いの知識を共有することによって、欲するアイデアが生まれやすくなる環境を指す。その視点から、JAIST¹を分析すると、研究者一人一人が多種多様で洗練されたバックグラウンドを取得しており、かつ多人数²であるので高密度な場の形成ができる可能性を十分に秘めている。

しかし、実際には発想を妨げる問題がいくつか存在している。その一つに、研究者間の環境のギャップがあげられた。研究者は一人一人異なる研究活動、時間、場所などの環境情報を持っているが、この環境情報はリアルタイムで他人に伝えにくく、コミュニケーションの場が作りにくくなっている。ゆえに、最適な発想の場でありながら、個人個人の持つ環境の違いにより、発想を妨げてしまっているものであった。

¹Japan Advanced Institute of Science and Technology, HOKURIKU

²教官: 約 220 名, 生徒: 約 910 名 (1999 年現在)

1.1.2 来訪者における背景

近年における JAIST の状況として、研究者に対する内外からの来訪者は増加傾向にある。しかしながら、このような状況にも関わらず、決定的な対策はまだ何も施されていない。つまり、面会時の時間配分、研究者自身の研究活動、または会議などの打ち合わせのほとんどは、研究者自身の管理で行っており、研究者自身の負担がかなり大きくなっている。また、研究者は各々独自の異なる研究活動、時間、場所などの環境情報を持ち、それに沿って行動している。よって、その人の環境情報は他人に知られにくく、結果として、それらに関する情報伝達に悪影響を及ぼしている。さらに、研究者の予定が突如変更する場合を想定すると、その研究者のスケジュールに関わっている他人は情報を速やかに知る手段が少なく、貴重な時間を無駄にする可能性が考えられる。特に、JAIST の来訪者はこの問題にあたる可能性が極めて高く、これら問題を解決する必要性が早急に求められている。

1.2 目的

本論文では、JAIST 内における研究者のリアルタイムな環境情報、及び位置情報の得難い問題を解決することを目的としている。そしてその目的を遂行し、情報を速やかに引き出せ、その結果、時間を有効活用できるシステム「CommNavi」を構築する。特に、CommNavi は以下の3点に重点を置いて開発する。

1. 研究者間の環境によるギャップを補完するための的確な情報提供

JAIST の内線 PHS 端末 (以下 PS³) を介して情報提供を行う。また、情報はナレッジマネジメントの個人化戦略を基にして提供する。その際、相手に関する何の情報を与えればよいのかを追究する。

2. 音声インターフェイスのみによる道案内の追求

道案内では、カーナビゲーションに代表されるように、仮想空間を作り上げて表示するディスプレイを用いたものが主流である。しかしながら、

³Personal Station

本研究ではPSによる音声インターフェイスのみでナビゲートを行う。特に、音声合成、認識が発展している今現在において、もう一度考察する必要があると考えている。もし、ディスプレイなしでナビゲートができれば、仮想空間を実世界に置き換えたナビゲートができる証明になる。

3. 機械対人間による対話のあり方の見直し

人对機械との会話については、本実験を通して情報取得スピードに関しての見直しを行う。その結果、ユーザに応じて早く情報を見つけられるような会話空間の構築をする。そして、構築する点で何が重要であるかを明確にする。

これら上記の目的を評価するうえで、定性的（プロトコル分析、アンケート）、定量的（道案内における所要時間）な視点から考察し、最終的にシステムの有効性を確かめる。

1.3 関連研究

CommNaviを構築する上で、最近の関連した研究や動向を簡単にまとめ、その上で本研究の位置付けを行う。

1.3.1 個人化戦略

ナレッジマネジメントの調査により、ナレッジマネジメント戦略はコード化戦略と個人化戦略の2種類に大別できることが報告されている [Tierney 99]。まず、コード化戦略とは、知識をコード化（形式化）、データベースに格納、再利用、のステップから成り立つものであり、リエンジニアリングなど、既成のサブセットにあたるプロジェクトの遂行においては非常に有効な戦略である。よって、コード化戦略を採用するには、人から抽出した形式知をどのように格納し、どのように再利用できるかがポイントとなる。

これに対し、個人化戦略は人と人の場の共有、すなわち「対話」を重要視する。個人の持つすべての暗黙知を相手に伝えるには、お互いが話し合い、理解を深めあうことにより伝えられるという考えを基底にしている。コード化できない暗黙知や、専門的な知

識が必要にある新規的なプロジェクトの遂行は個人化戦略の方が適する。ゆえに、個人化戦略を採用するには、コミュニケーション環境の整備が鍵となる。しかし、ここで注意しなければならない点があり、それは、コード化戦略、個人化戦略の特性を踏まえた上で用いなければならないことである。もし、用い方を誤ると本来の力を出せないばかりでなく、悪影響を及ぼす可能性も十分にあり得る。

1.3.2 位置情報を利用したツール

位置情報を利用したサービスは、複数の低軌道衛星を使用して実世界における位置を割り出す GPS⁴を使用したカーナビゲーションが普及している [後藤 99]。特徴としては、精度面で大変優れており一般的な誤差が約 10m ~ 100m である。ナビゲーションにおいてはディスプレイを用いた仮想空間のグラフィックと音声を併用している。この GPS はさらなる応用として、天災、緊急時などに役立てようとする方向にある。また、情報サービスである VICS⁵と組み合わせて、文字、簡単な図形、などの情報提供をしている。

この GPS と比較対象される位置情報検出として、PHS⁶を使ったものがある [後藤 99, 村田 99]。これは、PHS が使用している PHS 基地局を利用した位置検出である。数年前は、車などによる高速走行時において、位置登録中にハンドオーバーしてしまうことがあり、そのため位置情報を拾えない状況があったが、最近の技術進歩により回避されつつある。しかし、位置情報検出の点から考えると、PHS 基地局のカバーは約半径 100m であり、GPS に比べると誤差が大きいく欠点となっている。また、同じ構成で規模を縮小したものに構内 PHS システムがあり、高さがある建物内などでは GPS に比べて有効なものである。この基地局を用いた位置情報検出サービスとしては、NTT DoCoMo 社の「いまどこサービス」[DoCoMo] が有名である。その他の利用としては、痴呆徘徊者の位置情報検出、ペットの位置検出に应用されているが、同時にプライバシーの問題を含んでおり、その解決手段が望まれている [桜井 99]。

別のアプローチとしては、実世界指向インターフェイスである「WorkNavi」[Nagao 96] があげられる。これは、GPS や WWW⁷を組み込んだものであり、ディスプレイを使用

⁴Global Positioning System

⁵Vehicle Information and Communication System

⁶Personal Handy-phone System

⁷World Wide Web

したナビゲートと情報案内を行っている。なお、本研究に関連する特徴としては、歩く人を対象にしたナビゲート、及び、自然言語対話を組み込みがあげられる。

1.3.3 音声情報に関する媒体

本研究に関する音声メディアの媒体、PHS と CTI⁸ について述べる。近年、携帯電話、CTI、など音声を扱う媒体の普及は目覚ましいものとなっており、次世代 IT 技術の基盤となる要素を十分に秘めている。まず、携帯電話、PHS に関しては、通信手段として世の中に深く浸透し、生活の一部となっている。さらに、音声情報を扱うだけでなく、E-Mail の送受信、銀行の振込みなど「話す端末」だけでなく「使う端末」として日々進化している [村田 99]。

一方、CTI は、コンピュータと電話を統合する技術の総称であり、主にコールセンターの支援環境として普及している。そして、将来の予測として、E-Mail、グループウェア、VoIP⁹、などインターネット技術との融合 [小山 97] により、オフィス環境の支援が拡張されると期待される。

1.3.4 音声メディア

音声に関する媒体が増えるにつれて、音声認識、合成、理解に関する技術が重要視され、同じく急速な進化を遂げた。音声メディアは、人がコミュニケーションをとる上で昔から扱ってきているものであり、人が訓練なしでかつ理解しやすい形で情報を渡せる、可視化できない空間にも伝えることができる、などのメリットがある。

しかしながら、音声インターフェイスはそのほとんどがマンマシンインターフェイスの一部でしか使われてない [田中 99]。原因となる要素は、音声認識の不完全さと自然言語解析がロバストでなくまだ研究段階であることがあげられる。また、人間が行動する時は、視覚から取り入れる情報を中心にしたものになるので、聴覚を中心とした音声メディアはその扱いを間違えると人は認識不足、勘違いをする可能性が高い。ゆえに、実世界もその問題を考慮して視覚から情報を読み取ることを主体とした標識などが普及し

⁸Computer-Telephony Integration

⁹Voice over Internet Protocol

ており，聴覚から得られる情報メディアは少ないのが現状である．

将来として，音声技術は，音声情報を伝える媒体の普及にあわせて進化しており，媒体の性質によってその音声インターフェイスも変化して発展すると考えられる．

1.3.5 対話

表 1.1: 対話の成立要因

協調性の原則	現在進行形の話題に対して，自らその話題の進展，発展に貢献する．
量の公理	必要な情報を与え，不必要な情報は持ち込まない．
質の公理	嘘，不確定な情報は持ち込まない．
関連の公理	話題に関連した情報だけ提供する．
様態の公理	発話する情報は常に明瞭であり，無駄に長い情報提供はしない．

対話の分類においては，対話となる主題がはっきりしている目的指向型，主題がはっきりしない自由展開型，そして両者の中間である混合型の3つに分類されている[山梨 86]．

また，Grice[Grice 75]は，対話における情報交換効率を最大化を目指すための提案（対話に対する原則と4つの公理）を行っている．これを表 1.1 に簡潔にまとめた．

しかしながら，この表 1.1 は，人対人の対話の原則であり，人対機械の対話ではその様相が少し変わってくることが知られており，電話対話においての人対人，人対機械の比較実験がされていた[有田 87, Cohen 84]．これを以下の表 1.2 に示す．

現在においては，製品としての人対機械の対話は，自然言語を取り入れたものは少なく，予め決められた単語を入力することによっての対話が主流である．例として，IDO社の「乗り換えナビ」[IDO]があり，携帯電話からの音声入力も受け付けている．自然言語処理に関しては，日本語連続音声理解システム SPOJUS-SYHO・Y [堂下 98]があり自然言語対話における有効性が示されている．

表 1.2: 対話形態の違い

分析視点	人対人	人対機械
タイプ	同時性	逐次性
言い直し	多い	少ない
復唱		
間投詞		
同定行為の比率	高い	低い
対話の目的とする内容率	低い	高い

1.4 本研究の位置付け

オフィス環境における支援ツールの動向として、グループウェアとCTIを組み合わせる動きが見られる[小山 97]。特にCTIに関しては、企業内の仕事の効率化、個人管理に大変優れているが、ナレッジマネジメントの観点では定義分けができておらず、かつ、その統合としてはコード化戦略を中心としたものであった。対して、本研究は、CTIと個人化戦略を統合している。もともと、個人化戦略としてのツールは少なく新しい手法である。また、CTIの開発に関しては、企業内の仕事の効率化を重視[BizIT]しているため、来訪者に対する配慮がなされておらず、閉じたものとなっている点が問題であり、その解決手段を本稿で述べる。

近年の音声認識の急激な発達に伴い、機械の人間との対話のあり方が問われる時代に来ていると感じた。現存するシステムの機械対人間の対話においてそのほとんどは、人間側の情報収集に用いられているが、会話の構築方法は上記の表 1.1, 1.2 を原則としている。しかしながら、機械は最終的な目標として情報提供を行うように構築されており、そのための情報収集スピードという視点に関する会話形態の追求は少ない。

道案内に関しては、ディスプレイを使ったものが多いが、本システムは音声インターフェイスのみを使用する。また、位置情報は Positioning Server から取得する。このような組み合わせの道案内システムはまだない。

1.5 本論文の構成

本論文は、本章も含め5章から構成される。

2章 初めに、システム CommNavi の特徴を述べる。その後、システムの構成を、ハードウェア、ソフトウェア、データベースの概要を示す。

3章 対話するにあたっての実行方法を述べた後、CommNavi に実装した会話についての詳細を説明する。

4章 まず、実験の評価方法を示す。そして、CommNavi を使用した実験を述べる。なお、予備実験は3つ、最終実験は1つである。

5章 前章の実験の成果を目的に基づいて考察し、システムの有効性を示す。その後に課題、応用例を提案して本論文をまとめる。

第 2 章

システム構成

2.1 CommNavi の特徴

2.1.1 音声インターフェイスの導入

本システムは、JAIST 内においてのコミュニケーション・ネットワークを構築するために、PBX¹に登録された PS 使用する。なお、PS 端末は、JAIST 内の各研究者に配られているものである。

CommNavi の構築にあたり、メインとして音声インターフェイスのみの実装を試みた。これによりユーザは、PS を介して CommNavi と対話し、JAIST 内の環境情報を抽出する。よって、CommNavi からの情報はすべて音声によって伝えられ、情報検索においてもユーザの音声入力によって行われる。

しかしながら、音声インターフェイスだけの提供は、視覚情報に比べて誤認識や間違いなど弊害の生まれる可能性が高い。よって、これらを踏まえた上で解決策を提案し、音声インターフェイスを導入する。

2.1.2 ナビゲーション

ここで述べるナビゲーションは、カーナビゲーションに代表されるような道案内の他に、面会予定者 (Target Partner) の環境情報を与え、ユーザが時間を有効に使えるよう

¹Private Branch eXchange

に助言も行うことである。

道案内

JAIST 内の特徴として、敷地が広い、同じような空間が多い、高さがある、行き先方が複数存在することがあるなど、来訪者にとって迷いやすい特殊な状態にある。ゆえに、来訪者を迷うことなく目的地、または目的とする人（以下、ターゲット）まで案内することは、来訪者の負担が軽減され、その結果、有効に使える時間を増やすことができる、案内役を付ける必要もなく成るなどのメリットがある。

ターゲットした所までの行き先方が分からないなど、JAIST 内に慣れていない来訪者は、CommNavi と対話することによって道案内を選択することができる。ここで、来訪者に関しては JAIST 内に入り次第貸し出すものと前提としている。

一般の PHS が基地局を利用していると同様に、JAIST 内においても範囲が狭いアンテナ（以下 CS）が各所に配備されている。よって、JAIST 内にある PBX はその付加機能として、各々の PS に対して CS²を使用した位置取得機能を持たせることができる。つまり、一般の PHS システムに対して規模は小さいものの、CS にアクセスしている PS があれば、その持ち主の場所も特定できるというものである。よって PS を来訪者に貸し出すことにより、JAIST 内における来訪者の現在地を特定できることになる。

また、カーナビゲーションと大きく異なる所は、特定の場所への道案内の他に、面会予定者に直接向かう道案内もできる。それは、面会予定者が位置情報を特定できる PS を携帯しているので、来訪者との距離の差分を計算して道案内ができることであり、CommNavi の大きな特徴でもある。研究者は研究室を始めとして、さまざまな位置で研究活動を行っているし、突然面会の場を変更したい場合も想定できる。このような事態に対しても本システムは大変有効なものになる。

なお、カーナビゲーションに代表される道案内は、ディスプレイを用いたものが中心であるが、本システムは音声インターフェイスだけを用いて道案内を開始する。そのため、ユーザはディスプレイを使用した環境に比べて、自分がどこにいるのか分かりにくいという問題が生じる。さらに、JAIST 内は場所によってはなにも目印がない、似たような空間、PS がハンドオーバーできない場所も存在する。よって、それら問題の対策を

²Cell Station

作成し，実験によって音声インターフェイスのみの道案内の有効性を検証していく．

環境情報の提供

従来から相手への連絡手段は，内線電話，E-Mail や研究室のドアの前の張り紙がほとんどであった．これらの非同期的な連絡手段では，伝言が相手に伝わる時のタイムラグの発生が否めない．このため，相手への連絡の伝わりが遅くなり，結果としてすれ違いや，時間を無駄にってしまう可能性があった．また，お互いPS を使って連絡を取り合うこともできるが，この同期的な方法では，相手が忙しい状態であるなどの場合，連絡が取れなくなってしまうことがある．つまり，人に対する連絡手段はその時の状況に応じて同期的，非同期的な方法を使い分けなければならない．

この問題に関しては，CommNavi の構築にあたって，コミュニケーション・ネットワークを形成し，解決を目指す．それは，相手と意思疎通ができない場合，CommNavi は仲介役となり，ユーザが時間を有効に使えるよう指示するものである．まず，非同期的な形として個人情報の一部に付け加える形で，ボイスメッセージ (Voice Message) を保存できるようにした．それは，急な用事が入るなどしたら，相手への伝言として利用できるものである．これは，ユーザからもボイスメッセージを追加できるものであり，一般の留守番電話機能のオープン化を目指したものである．

一方，会議の延長，急な用事が入るなど予測できない状況においては，影響を受ける他人に速やかに伝えたほうがよい．しかし，予定しなかったことが起る場面では，なかなかその情報を相手に伝えにくい．この時，PS を持っているので同期的な会話が望ましい．しかしながら，CS の範囲外での研究活動，会議などつながらない事があるので，これらの事態を解消するアルゴリズムを実装する．例えば，ユーザが来訪者であれば図書館などに案内させ，内部の人間であれば終わるまでに自分の研究をするよう助言する．

2.2 ハードウェアの構成

ハードウェアの構成を図 2.1 に示す．

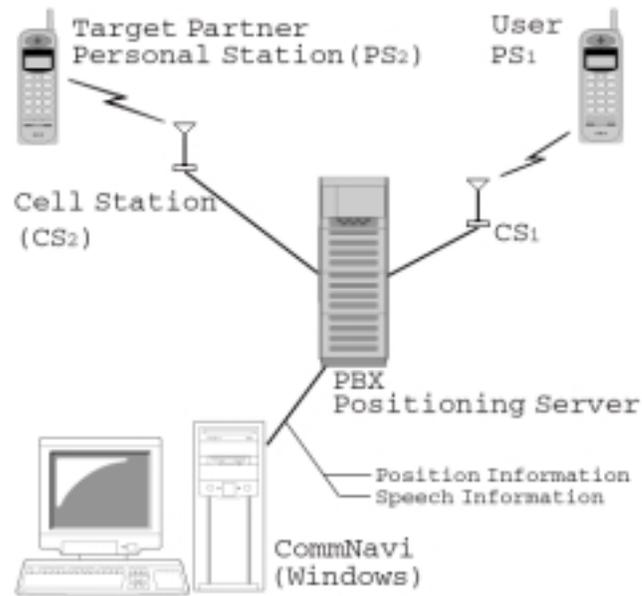


図 2.1: ハードウェアの構成

2.2.1 PS

PBXに登録されたPS端末は、ユーザ、および求める相手の人が必ず持っていることを前提としており、CommNaviを間に挟むことによって、コミュニケーションネットワークを形成する。ユーザは直接相手にPSを介しても連絡が取れるが、相手側の都合により出れない場合がある。さらに、ユーザが来訪者であった場合、相手までの道のりが分からない場合も想定できる。この対策に、PSを使ってCommNaviと対話し、相手側の環境情報を取得、ユーザは自分自身の持つ時間を有効に使えるものである。また、道案内時にも、CommNaviから伝えられる位置情報などの音声情報を頼りにしてユーザは行動する。

PSの特徴として、CSの範囲にいる限りそのCSを絶えず利用しているため、完全にCSの圏内を出ないと次のCSに位置登録ができない。また、PSはCSの圏内を出る電源切れの行為によって使用ができなくなる。

2.2.2 CS

CSは、JAIST内に設置されているものであり、PSを使った音声情報の送受信の時に使用するものである。JAIST内には多数のCSが配置されているが、場所によって配置されてない所もある。また、1フロアに固まって配置されている所もあれば、広い場所に1つのCSしか置いていない所もある。

そしてCS1つの性能は、同時に3つの位置登録ができ、約20m~30mの範囲でPSを補足できる。しかしながら、そのCSの位置の建物の構造上によって範囲が可変するものであり、道案内などの位置情報を扱う時には注意を払わなくてはならない。さらに、PSを持っている人の行動が複数噛み合わさると、同じ場所でも違うCSを使うことになり、場所によって確実な予測は不可能な所もある。

また、CSのもう一つの特徴としては一般のPHSシステムと同様の位置情報を扱うことができる。つまり、CSの位置が分かれば、それを使っているPSもCSのエリア範囲にいることが分かる。よって、CSが設置されている場所やCSを使っている人の検索をして、ユーザに案内させることができる。

2.2.3 PBX(Positioning Server)

位置情報サービスは、PBXの付加機能であるPositioning Serverが管理している。このPositioning Serverは、PSが位置登録、発信、着信、ハンドオーバー時に使用するCSのIDとPSのIDをPCに出力するものである。Positioning ServerとCommNaviが実装されているPCの間は、現在のところ情報センターからの光ファイバケーブルとRS232Cで結ばれており、ここで位置情報を流している。

2.3 ソフトウェアの構成

図2.2にCommNaviを構成するソフトウェアを示す。

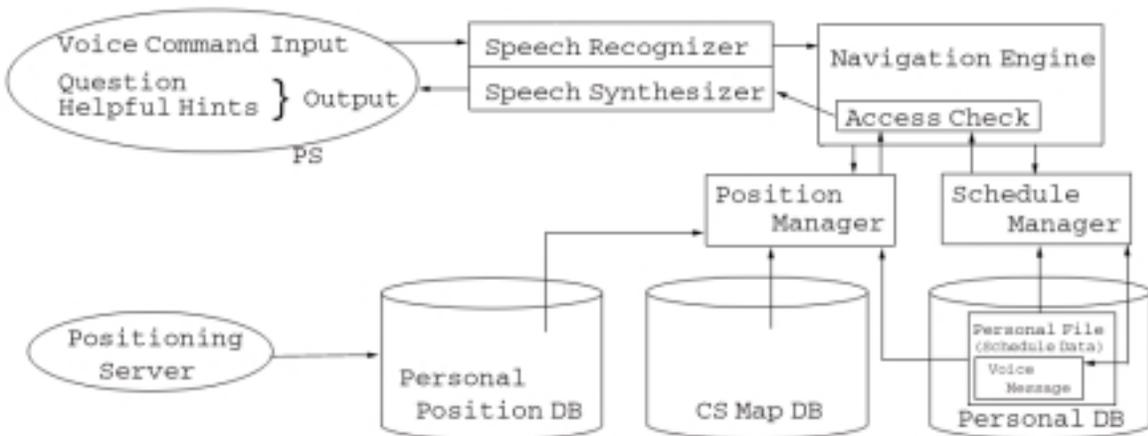


図 2.2: ソフトウェアの構成

2.3.1 音声認識と音声合成

ユーザは対話を行うことによって CommNavi から情報を引き出す。そのために、音声認識、合成 (Speech Recognizer, Speech Synthesizer) を用いた。なお、エンジンは IBM 社製の ViaVoiceTM SDK for Windows[®] Version 1.5 である。これは、特定話者の認識エンジンである。本来ならば、不特定多数の人を想定してるので、エンジンは不特定話者認識エンジンを使うことが望ましい。よって、人により音声入力を受け付けにくいと考えられた。そのため、実験的に幅広く受け入れられるワードを探して、それらを使用する方法を用いた。

CommNavi への入力、および出力は PS を使った音声インターフェイスを用いた対話で行われる。図 2.2 のボイスコマンド (Voice Command) はユーザからの入力である。この孤立単語認識される語を入力することによって CommNavi は応答をする。それに答えて CommNavi は 次の質問 (Question) か、ユーザの欲する情報 (Helpful Hints) を与える。また、ボイスコマンドはユーザがすべて覚えなくてもよく、CommNavi から次の対話にふさわしいボイスコマンドを次の質問といっしょに提供する。つまり、ユーザは与えられた単語をしゃべるだけで、知りたい情報を得ることができる。

2.3.2 ナビゲーションエンジン (Navigation Engine)

音声認識モジュールによってユーザの音声入力を解読した後、ナビゲーションエンジン (Navigation Engine) は意味解析をし、下位のモジュールに対応する指示を与える。位置情報、及びJAIST内の研究者は、すべての人がオープンに情報を与えることはないものであり、その対策としてアクセスチェック (Access Check) を設けた。その判別作業をこのモジュールで行う。

また、下位から情報をもらうと、その情報をユーザの聞き取りやすい文にして音声合成に送るものである。

2.3.3 位置管理モジュール (Position Manager)

ナビゲーションエンジンから情報をもらうと、3つのデータベースを利用して、対応する答えを返す。その情報を以下に示す。

- ユーザ自身の位置情報を与える。なお、対話で一番初めに与える位置情報は、現在使用しているCSの範囲で答える。この情報は主にJAIST内に慣れていない来訪者を想定して渡すものである。また、ユーザ側の誤差補正があった場合、それに応じて位置情報を与える。
- ユーザが知りたい相手の位置情報を渡す。この情報はすべての対象者を想定しているものであり、頻繁に利用される情報である。なお、相手側の位置情報は、相手の持つPSが使用しているCSを使って与える。
- 道案内に関しては、ユーザの位置情報、相手側の位置情報、及びCSを基にしたデータベースを使って、ユーザに進むべき方向を音声で伝える。これは、ユーザがJAISTに慣れていない来訪者などの利用を想定している。同時に、最短距離の導出、特殊な環境情報 (エレベーター内、及びCSが少ない場所ではハンドオーバに失敗する可能性があること) も付加させている。これによって、ユーザは迷わず相手の位置にたどり着けるようになる。

2.3.4 スケジュール管理モジュール (Schedule Manager)

急な予定や、会議の延長など必ずしもスケジュール通りに行かない問題がしばしば発生している。このような時、影響が及ぶ他の人にいち早くその情報を与えなければならない。しかし、研究者のスケジュールはなかなか分かりにくいものであり、研究室に出向いて知るか、E-Mail 等でのコミュニケーション手段を取るしかなかった。そこで、CommNavi にスケジュールデータを扱わせ、音声を通して研究者のスケジュール情報の入出力をすることにした。これより、ユーザがどの位置にいても自分の情報を即座に更新でき、また別の人とは CommNavi と対話することによって、どの位置からでも対応ができる。また付加価値として、家庭電話機に見られる留守番電話機能とほぼ同一であるボイスメッセージ (Voice Message) の入出力を行った。留守番電話機能と異なる点としては、録音されたボイスメッセージはテキストデータで保存をすることで、扱うファイルの容量が少なくなることであること、誰でも聞けることにある。

2.4 データベースの構成

2.4.1 個人情報 (Personal DB)

個人情報データベースには、個人情報 (mailID, アクセス権, 内線番号, 初期状態における個人の位置情報など) とスケジュール情報が収められ、位置情報管理モジュール、スケジュール管理モジュールに関連しているものである。さらに、ボイスメッセージを格納するスペースも用意されている。位置情報管理モジュールに関しては、PS コード、アクセス権を渡す。対して、スケジュール管理モジュールに関しては、スケジュール情報とアクセス権を渡す。

学内においては、スケジュール情報がばらばらに保存されていることが多いので、それらを統合することにより、さらなる時間の効率化が期待できる。

2.4.2 位置情報 (Personal Position DB)

位置情報データベースは PBX の位置情報検出機能から送られて来る情報を基に作成されている。また、このデータベースは午前 0 時をもって DB はクリアされるものであ

る．構成は個人のPSコードと，感知したCSコードからなり，それらを位置情報管理モジュールに渡す．

2.4.3 CSを中心とした構内マップ情報（CS Map DB）

位置情報データベースにはCSコード，場所，次にハンドオーバーする可能性のあるCSコード，及び相対距離などのJAIST構内における情報が収められている．このデータベースは位置情報モジュールに呼び出され，それに応じた値を渡すものである．各CSの場所は，特徴を明確にしておかないと指示があいまいなものなる．しかし，細かく設定しすぎると，誤差が発生した時にユーザに不信感を与えてしまうものであり，その調整は実験によって算出される．よって，このデータベースは道案内に関する情報に特に大きく影響を与えてしまうものである．

第 3 章

実行方法

3.1 対話における情報取得方法

ユーザは、CommNavi に直接電話をかけることにより対話を行える。この時、CommNavi は 2, 3 択からなる質問をするものであり、その対話の分岐例を図 3.2 に示す。例 1 では CommNavi の問いに対してユーザが「スケジュール」(注 1) と答えている。すると、CommNavi はスケジュール情報をユーザに伝える。しかし、例 2 では「位置情報」(注 2) と答えている。すると、CommNavi は位置情報を渡している。このようにして、対話は CommNavi の質問、ユーザ側は選択肢に対する回答をして進めていき、最終的にユーザ側の欲する答えが導き出される。

しかし、これは原則的な会話である。熟練者の使用になるにつれ、このステップが煩わしいものになってくると想定した。この時、CommNavi から与えられる選択肢を無視したショーカットもできる。例 1 の(注 3)においては、3 択から選ばせている。しかし、この選択を無視して「國藤進」とすれば、ターゲットする相手が変わり、それに応じた答えを CommNavi が返すものである。

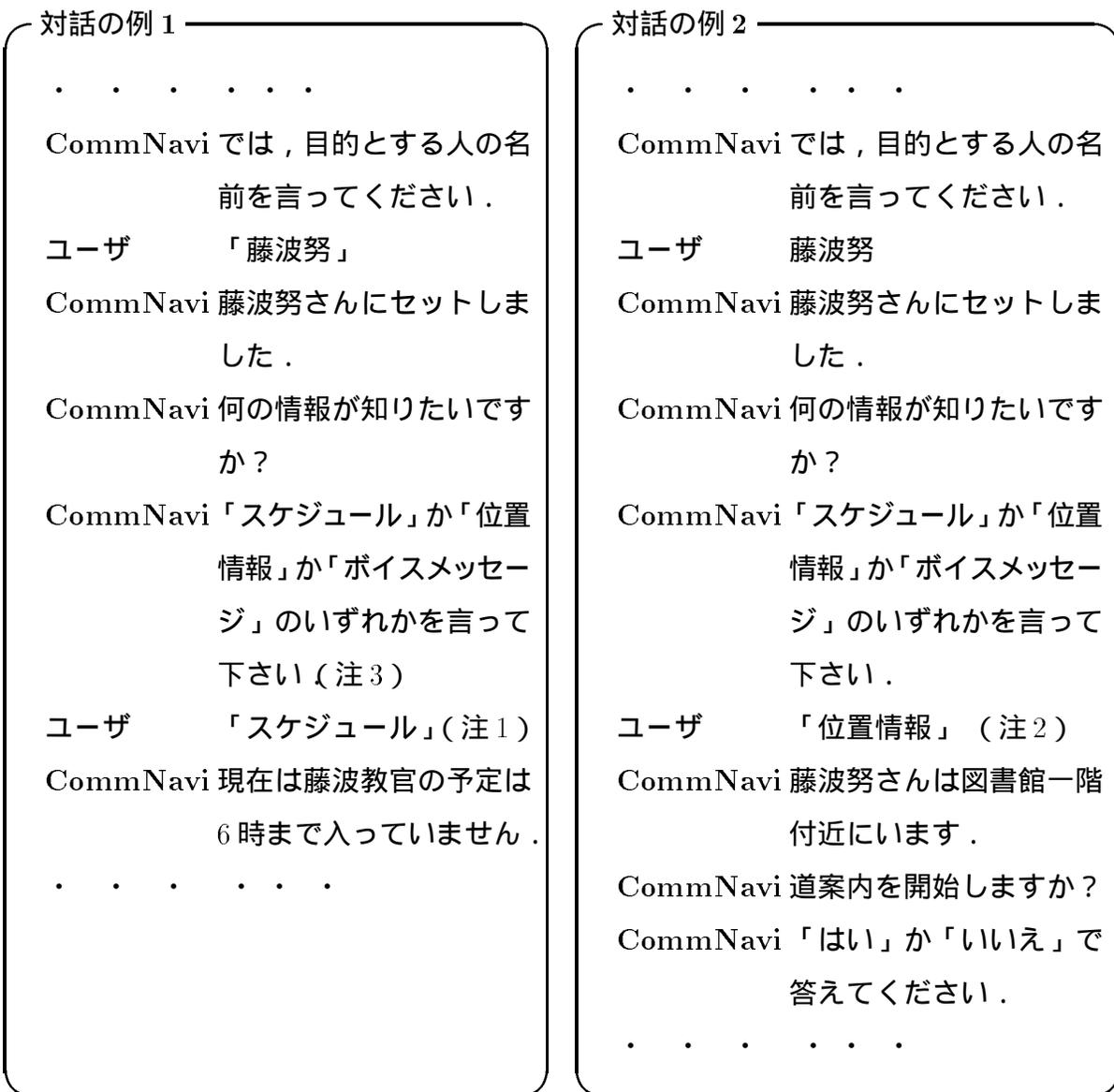


図 3.1: 対話の分岐

3.2 会話の構築

本システムを使う上で、理想的な会話の流れ（図 3.2，表 3.1）を導出した。

対話の流れは、大きく3つのブロックに分けることができる。それは、JAIST 内におけるユーザおよび研究者の状況に関するブロック、ユーザ自身の環境情報の更新によるブロック、ボイスメッセージ扱いによるブロックである。なお、最初の2つのブロックは互いに独立しているが、ボイスメッセージの扱いのブロックは、2つのどちらかを進めるようになっている。よって、会話の流れを整えるため、フラグを付けて制御している。また、最初の2つのブロックは対話のルートとなる地点を持っており、そこから実質的な対話が始まる。

ユーザの知りたい情報一つに対して、3～5ステップからなる対話を行うことによって導き出される。また、ユーザが行う選択肢は2及び3から成り立つものである。理由として、使い慣れていないユーザは4つ以上の選択に支障をきたす恐れがあった。また、位置情報を聞いた後で、スケジュール情報が聞けるというショートカットなどもできるようにした。操作の熟練者になると、この操作でいち早く情報を引き出すことを可能にしている。

なお、会話において終わりとなる選択肢はない。ゆえに、ユーザが十分情報を引き出せたと感じたときにPSの切断により終了するものである。

以下、会話の構築詳細を述べる。

表 3.1: 図 3.2 の見方

図における記号	説明
→	対話の流れを表す
[]	CommNavi 中心の対話
< >	ユーザ中心の対話
[< >]	ユーザと CommNavi が協力した対話
0 or 1	0 は空であることを示し、1 は中身がある状態を示す
*1 - *23	一連の対話と動作におけるステップ ID ナンバー

以下にイベントの詳細を示す。

対話のスタート

1. [Start]

ここから、対話がスタートする。ユーザは CommNavi サーバーに直接電話をかけて、CommNavi は受信をした旨をユーザに言うステップである。

2. <Navigation><My Self>

選択肢は、「学内のナビゲート」か「自分の環境情報の更新」の2択である。学内のナビゲートは、研究者の環境情報、JAIST 内における道案内などを行う。それに対して、自分の環境情報の更新は、もともと実装しているスケジュールに違いが生じた場合、またはボイスメッセージを入れる場合に音声を使って更新するものである。

ターゲットの情報取得に関するイベント一覧

3. <All People><Target Partner><Room>

対話2を選ぶと、この対話になる。この対話に入ると、これ以降は学内ナビゲートの対話のルートとなる。つまり、ユーザが対話の流れが分からなくなった時、ユーザの判断でリフレッシュすると、この対話になる。また、ユーザが知りたい情報を渡した後はここに戻ってくる。選択肢は、「全員の研究者の検索」、「特定の研究者の情報検索」か、「部屋の位置情報検索」である。全員の研究者の検索は、対話する特定の人がない場合において、時間的に余裕のある人をピックアップするものである。それに対して特定の人を探したい場合、その人の名前をフルネームで述べる。また、会議室などの場所に直接行きたい場合は部屋情報を選択する。

4. [<Check_Time> Check_Schedule]

対話3の研究者全員のチェックを選択すると、この対話になる。CommNavi とユーザが協力し時間帯を中心として、空いている人を抽出する。

5. [Pattern Match Output]

対話4より、該当する研究者を選び出した後、候補となる研究者と、その研究者は何時まで時間が空いているかを CommNavi がしゃべる。この後、対話3に戻る。

6. <Schedule><Position><Voice Message>

ユーザは求めようとする特定の人を指名した場合、対話3からこの対話に流れる。ここでは「スケジュール情報」「位置情報」「ボイスメッセージ」の3択で答える。スケジュール情報は、現在における環境状態、位置情報は、学内においてどこにいるか知りたい場合、「ボイスメッセージ」は、その人に対してのメッセージを聞く、入れるなどの動作をする場合に選ぶ。

7. [Helpful Hints_1] , [Helpful Hints_0]

スケジュールデータを読み込んだ後、相手側に用事があった場合、及びない場合に分けてそれぞれ音声で出力する。それに応じてユーザは最適な行動を取ることができる。一番のメリットは、情報の柔軟さである。PSをお互い持っているので、同期的な対話で情報の交換もできるが、相手側が出れない場合、または来訪者がJAISTに到着しておらず、PSを持っていない場合に備えて、非同期的な情報を提供する。情報提供が終わると、対話3へ戻る。

8. [Position Infomation
Output]

対話6で相手の位置情報を選択した場合、ここで、すぐに位置情報を示す。理由は、ユーザがJAIST内を知っている人間であれば、何をしているか、推測がつくからである。しかしながら、アクセス保護がなされていた場合に限り、情報は渡さず、その旨を伝える。それらの情報を与えた後、対話9に行く。

9. [Accept_Navi
<No><Yes>]

対話8から続けてこの対話に入る「道案内をしますか」というCommNavi質問にユーザが答えを出す。もしJAIST内に詳しい人間であれば「いいえ」と答えることを想定している。対して、来訪者など、JAIST内部の構造を良く知らないユーザであれば、「はい」と答え、道案内の対話に移る。

10. [Navigation_P]

ここでは、音声インターフェイスのみを用いて、相手側のいる場所へ道案内を開始する。ユーザに的確に、かつストレスを感じさせないようにナビゲートを行うことを心がけている。なお対話に関しては実験により明らかにする。

11. [<Room_ID>]

対話3で、特定の部屋に行く選択をした場合、この対話になる。部屋には一意のIDが割り付けられている。よって、このID番号をユーザが述べることにより、行き先が分かる。また、CommNaviからは、部屋情報を述べることができる。

12. [Navigation_R]

対話11でCommNaviとユーザで行き先を決めた後、この対話になり、道案内が始まる。ナビゲーションのアルゴリズムは、対話10と同じであり、実験によって示す。

ユーザの情報更新に関するイベント一覧

13. [Identification] [<Error><OK>]

対話2でユーザ自身の環境情報の更新を選んだ場合、この対話になり、本人かどうかの整合チェックが行われる。CommNaviは、本人と判断したなら次の対話に進むことができ、違っていると判断したなら、繰り返し照合させる。

14. <Voice Message><Schedule>

対話 13 のユーザの照合が終わった後、この対話になる。ユーザはボイスメッセージの操作か、スケジュールの更新のいずれかを選択することができる。また、会話の流れが分からなくなって、リフレッシュしたい時には、この対話に戻ることができる。ユーザ自身の更新の対話ブロックにおけるルートとなるステップである。

15. <Erase_Message><Message Input><Hearing>

対話 14 において、メッセージの操作を選ぶとこの対話になる。ここでは、「メッセージの消去」「メッセージの入力」「メッセージの出力」を選ぶことができる。CommNaviはそれに応じて、メッセージ操作を行う。

16. [Erase_Message] , [Message Input] , [Hearing] , [Erase_Schedule]

ユーザからの入力によって、それぞれ、メッセージの消去、メッセージの入力、メッセージの出力、スケジュール消去のいずれかの処理を行う。また、この音声認識エンジンは不完全なものであるので、そのつど確認を行う。

17. <Do Nothing><New Schedule>[<Schedule_ID>]

対話 14 において、スケジュールの更新を選び、今現在の予定がない場合に発生する対話である。ここの対話は3つの選択肢に分かれていて、それぞれ「何もしない」「新しいスケジュールを入れる」「既存のスケジュールを移動する」をユーザが選択する。なお、スケジュールの移動に関しては、スケジュールに割り当てられた番号を述べて、現在のスケジュールに割り当てる。

18. [Check_Time
Schedule Input]

対話 17 において、新しいスケジュールの作成を選ぶことによって、この対話になる。ここでは、ユーザがスケジュールの入れたい時間帯を入力

すると、システムはそこに新しいスケジュールのフラグを立てる。そして、その時間になると、ユーザは予定が入っているということになる。

19. [Move_Schedule]

スケジュールの予定にはそれぞれID がつけられている。このID を指定することにより、システムはその予定の移動をすることができる。

20. <Do Nothing><Erase_Schedule>

対話14でスケジュールを選択し、現在の時間においてユーザのスケジュールで予定があると、この対話になる。ここでは、「現在の予定をそのままにする」か、「予定を消す」かのいずれかをとることができる。

ボイスメッセージの操作イベント一覧

21. <Do Nothing><Message Input>

この対話は、上位の対話選択において、ボイスメッセージを選択し、対するボイスメッセージが入っていない時に発生する。ここでは、「何もしない」、「ボイスメッセージの入力」のいずれかを選ぶ。何もしないを選ぶとそのまま対話のルートに戻る。ボイスメッセージの入力を選ぶと、入力メッセージ対話に入る。

22. <Message Input><Do Nothing><Hearing>

上位の対話選択において、ボイスメッセージの操作を選び、かつ相手のボイスメッセージが入っている時に起こる。「メッセージの入力」、「何もしない」の選択は対話21と同じである。相手へのボイスメッセージを聞きたい場合、メッセージの出力を行えるものである。この場合、何度でも聞きなおしを行うことができるようにループが組まれている。

会話の整流によるフラグ

23. [Set_Flag] , [[Check_Flag] [0][1]]

対話の流れは，JAIST 内のナビゲートとユーザ自身の環境更新に分けることができる．しかし，ボイスメッセージの操作で，ボイスメッセージが入っていないと，その後の対話は共通したものを使っている．そのため，ボイスメッセージの操作が終わると，どちらから対話が流れてきたかをチェックする必要がある．このため，ユーザ自身のナビゲートからの対話の場合，そのためのフラグを立てる操作 (Set_Flag) が入る．これは，フラグチェック (Check_Flag) で認識され，それに沿って対話の流れを調節する．

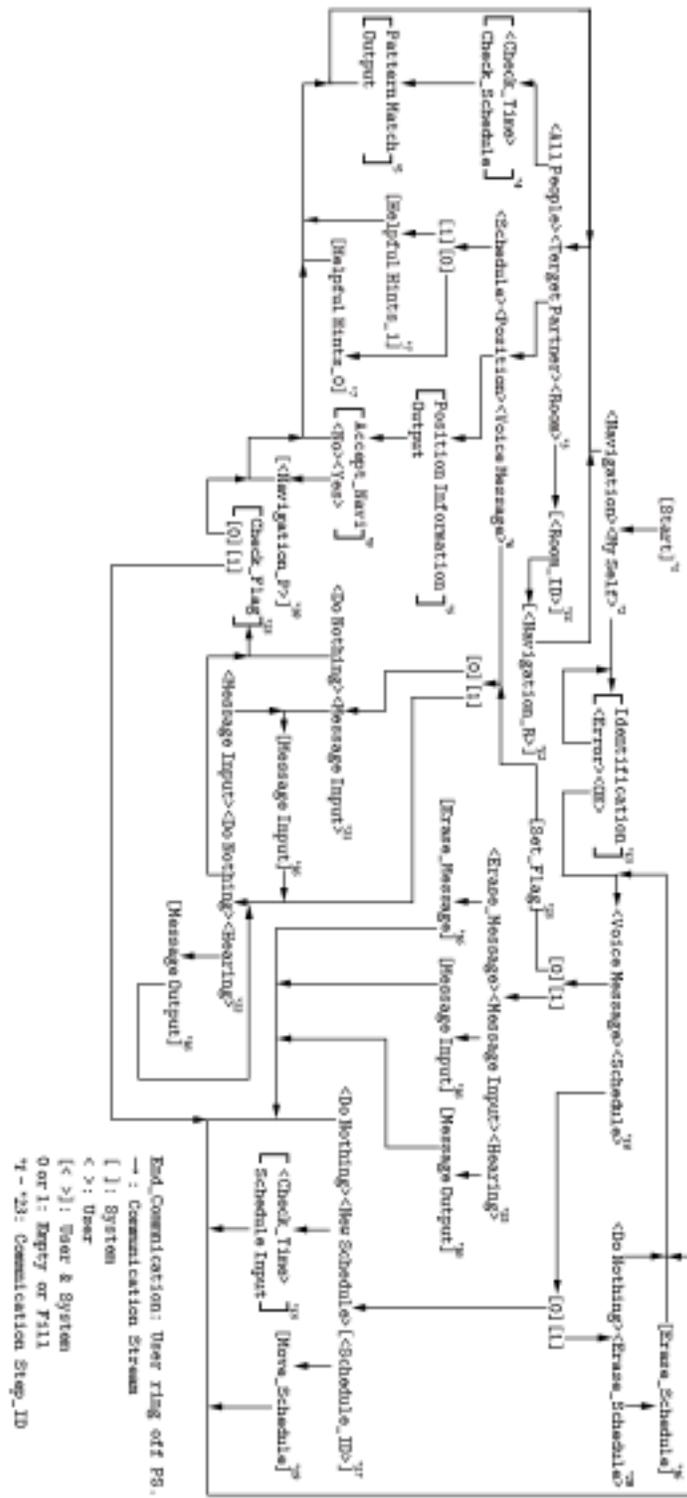


図 3.2: 会話

第 4 章

実験

4.1 実験の評価方法

本実験における評価は、研究者間の環境によるギャップを補完するための的確な情報提供 2 音声インターフェイスのみによる道案内の追求 機械対人間による対話のあり方の見直しを主眼にした。

まず、道案内システムでの主な実験は、情報を与える視点から大きく 4 つに分けられる。

実験 1 一定の時間になったら情報を提供する。

実験 2 ユーザが情報を欲する時のみ提供する。

実験 3 ハンドオーバー時に情報を提供する。

実験 4 実験 1~3 までを踏まえた上で、最終的な実験をする。

そして、被験者は実験 1~3 までは 4 人~7 人程度、実験 4 においては 8 人（経験者 4 人未経験者 4 人）である。

また、これら実験の評価方法として、以下の方法を選択した。

プロトコル分析においては、発問法 [海保 93] を用いた。その特徴を表 4.2 に示す。用いた理由としては、本実験において新しい概念をいくつも取り入れておりユーザの学習過程を理解しなければならないこと、及び、音声認識がされにくい時の対応をすぐに行えるからである。

表 4.1: 本実験の評価方法

実験の種類	定性的評価	定量的評価
実験 1	プロトコル分析 (発問法)	なし
実験 2	プロトコル分析 (発問法)	なし
実験 3	プロトコル分析 (発問法)	なし
実験 4	アンケート [辻 97]	道案内時の時間計測

表 4.2: プロトコル分析の特徴

メリット	デメリット
初心者による学習理解, 過程において強力な手法	学習理解に限られる
実験において慣れてない人が自然に発話できる環境	解答の設定基準が難しい

4.2 実験 1: 時間毎の情報提供

4.2.1 概要と目的

まず, Positioning Server のみから送られてくる情報を使っての実験を行った. 目的は以下の通りである.

- Positioning Server からの位置情報を使った動作確認
- Positioning Server の位置情報に関するポテンシャルの検証
- 音声合成の動作確認

まず, ユーザに学内を歩いてもらい自分自身の位置を確認してもらった. この時, ユーザに情報を伝えるタイミングとしては, 一定の時間毎にユーザ自身の位置情報を提供するものである. また, この時の状況として, 音声認識のパラメータの設定ができないなど難しい状態であったので, 一定の時間になると情報を提供することにした. よって, ユーザは何もしゃべらず音声合成の案内を聞いているだけでよい.

4.2.2 結果

実験を数人の被験者にしてもらい、プロトコルデータを基にした結果を出した。

Positioning Server の動作確認

Positioning Server からの位置情報をユーザに届けることができ、ユーザはそれを認識することに成功した(図 4.1)。これにより、システムの信頼性は確かめられた。



図 4.1: 音声合成

Positioning Server のポテンシャル

ユーザすべての意見としては、位置情報が曖昧過ぎるというものだった。先ほど通った場所に関する位置情報を与えてしまったり、まったく別の場所の位置情報を提供してしまったりした(図 4.2 では、ユーザが5階にいるのに対して4階と案内している)。以上より、正確な情報は少なく、誤差が大きいことが大きな問題であることが分かった。理由としては、CS の範囲が大きいものであることが原因であった。対策としては、CS の範囲を小さくすることがあげられる。しかしながら、理論的に誤差が小さくなるが、PS 本来の対話エリアが小さくなるなどの本来の性能に響いてしまうものである。これより、別のアプローチによる対策が必要であることが分かった。



図 4.2: Positioning Server の誤差情報

音声合成の動作確認

聞き取りにくい意見が多く出たが、案内を何度も聞くことによりユーザは理解度を増していた。また、繰り返し情報提供するので、聞き間違いによる大きな誤解は引き起こさなかった。また、会話のスピードを遅くすることによってユーザの認識率が変わるものであった。

4.2.3 考察

更新時間を少なくすると、情報提供が少なくなり、ユーザが知りたい時は待たなければならないという現象が発生した。逆に、更新時間を多くすると、繰り返し同じ情報を CommNavi から出されるので冗長性があることが判明した。特にこの冗長性は表 1.1 の会話の様態の公理に反するものであり、注意が必要であった。

位置情報の補正が必要不可欠であることが分かった。音声合成は聞きにくい所があるものの、ユーザにある程度の情報を伝えることができた。また、エレベーター内や電波が通じにくい場所など特殊な環境における指示も必要であることが分かった。

4.3 実験2: ユーザに応答する情報提供

4.3.1 概要と目的

小実験を重ねて、PS から孤立単語認識ができるようになった。また、認識されやすい単語の抽出もできるようになった。このため、図 3.2 対話空間の道案内における過程の実行をした。また、実験1ではCommNaviからしゃべることしかできなかったが、今回は対話を用いて、ターゲットの場所を理解し、さらにそこに進めるようにした。なお、道案内に関しては、ユーザが情報を欲する時のみ提供することにした。以下に今度の実験の目的を示す。

- 音声認識を使った対話の実行の確認
- CS を使った相手との距離を算出し、ユーザに伝えることの検証
- 特殊な環境の対応と重み付けの検証
- ユーザ側からの位置情報の補正による検証

4.3.2 結果

PS を使用した音声認識の確認

本システムの音声認識エンジンは特定話者に対応、および特定のマイクを利用したものである。よって、本システムと組み合わせると、弊害が生まれる可能性が十分にあった。理由としては、CommNavi ユーザは不特定多数である。よって、どのような人でも認識させなければならない。さらに、肉声とかなり変わる PS を介して認識させる。そのため、デフォルトのままでは認識させることができなかった。しかしながら、ある特定の単語を学習させ、認識をさせる形態を用いることにより認識ができるようになった。よって、その単語を中心として対話の流れを構築した。

最初の音声認識では、かなり評判がよくなか、ユーザはストレスを感じていた。しかし、何回も認識させる行動を繰り返すことによって、音声認識エンジンを働かせることに成功した。

距離の算出とルート設定

道案内においての付加機能として、ターゲットまでの距離を算出させるアルゴリズムを作った。これは、ユーザのPSが使ってるCSの場所と目的とする人が使っているCSの場所の距離を算出し、それをユーザに伝えるものである。これにより、相手側に近づいているか遠のいているかが分かるものである。これによって、方向の指示が可能になった(図4.3)。また、ルート設定においては、特殊な環境(CSが少ない場所の回避)の重みを付加した最短距離を導出し、相手にそのルートで向かわせた。

その結果、ユーザは誤差補正と組み合わせてターゲットまでたどり着くことができた。また、距離を導出し、それを相手に教えることに関してはかなりの好評だった。



図 4.3: ルートの指示

特殊な環境への対応

JAIST 内には、エレベーターが複数存在し、そのエレベーターにおいては、ハンドオーバができなかったり、動作が不安定になる場所があった。またCSのエリア外など、特殊な場所が多数存在する。そこで、まずその環境に応じて注意を促すようにした。さらに、ルート設定時においては、そのような場所を重みを多くして回避するようにさせた。

特殊な環境情報の提供については好評だった。しかしながら、特殊な場所を迂回させ

ることは賛成派と反対派に分かれた。賛成派の意見としては、そのままナビゲートしてもらえるので確実に目的まで行けるので安心である、というものである。反対派の意見としては、最短距離で行く方法があったら教えてもらったほうがよい、または、急いでいる人には、よくないのではないか、という意見もでた。

ユーザ側からの位置補正

実験1において、Positioning Serverの位置情報の曖昧さが目立つ結果となってしまった。その対策として今回は、ユーザ側の位置補正をすることにした。それは、ユーザ自身が現在見えているものをCommNaviに教え、CommNaviはそれを判断して誤差補正する。しかしながら、見えているものは、さまざまなものがあり、さらに同じような場所や何もないような場所も存在する。また、見えているものはユーザの知識表現によってさまざまな言い方に変化する。例えば、JAIST 正面玄関奥に黒い像がある。これを、CommNaviに伝える場合「黒い像」「人の像」「玄関前にある像」など、人によって多種多様な表現を使うことになる。つまり人と人の対話であったら理解できるものの、機械においては判断と予測がつきにくいものである。さらに、CommNaviでは、ある決められた単語しか認識できず、その単語を機械に伝えることは難しい。むしろ、そのようなことを伝えていては、膨大な数を機械に教えなくてはならない可能性もある。また、人間の思いつく単語をすべて登録すると際限がなく、誤認識結果が増える可能性も出てくる。つまり、ユーザの知識表現に左右されにくいものを選ばなくてはならない。

今回、ユーザ側からの位置補正として用いたものは部屋に割り振られたIDである(図4.4)。各部屋ごとのIDは一意であり、「C21」「図書館」などである。これは、「C21」のユーザのしゃべりは「シーにじゅういち」「シーにいいち」など限られたものになる。なお、この誤差補正の仕方は、口頭でユーザに伝えた。

結果は大変によく、誤差補正の手法に関する大きな有効性を示せた。しかしながら、音声認識がされにくいという現象も発生していた。それは、その空間に応じた声の反響などが大きな要因であると推測ができた。このため、音声認識エンジンのパラメータ調整がさらに必要であることが分かった。



図 4.4: 位置補正

4.3.3 考察

この実験における情報の与え方は、ユーザが情報を欲するときのみ与えた(図4.5)。実験1では、ユーザが知りたくなくても情報を与えてしまう欠点を重視して解決を試みた。

その結果、ユーザが知りたい時に CommNavi の情報を引き出せるので解決ができた。しかしながら、こちらから言わないとなにもしないので冷たい感じがする、または、いつハンドオーバーしたのか分からないといった意見も出された。

さらに対話において、CommNavi から伝えられる情報が長すぎるといった意見も多く出された。特殊な場所における対応の仕方、ユーザ、および相手がどこにいるかの情報、どちらにいけばよいかの案内などを合わせると約30秒かかっていた案内もあり、様態の公理に反するので、この問題を回避しなくてはならないと感じた。

4.4 実験3: ハンドオーバー時に情報を提供

4.4.1 概要と目的

いつもフレッシュな情報を与えるため、ハンドオーバーした時点でユーザにその旨を伝えるアルゴリズムを構築した。なお、この時、ユーザからの誤差補正はないものとする。



図 4.5: ユーザの問いかけ

- 道案内時における目印の指示
- 対話の冗長性を削減
- 対話において重要性のある情報を優先させる
- 音声認識, 合成のパラメータ調整
- ユーザのストレス度を量る

4.4.2 結果

目的物の指示

道案内においては, 目立つ目印, 案内板などの情報を提供した。つまり, ルート設定する時に, もし目印になるものがそばにあり次第, それをユーザに伝えるものである。しかし, 注意すべき点は, 目印が小さい, 見る方向では死角になる, という状態ではユーザはそれを見つけることができず, 逆にパニックに陥る可能性もある。よって, あちこちに点在している JAIST のマップや売店など分かりやすいものだけを選んだ。ゆえに, ディスプレイを使ったナビゲートではランドマークを使えるという強力な手法であるが,

音声インタ - フェイスでは諸刃の剣であることが分かり多用は控えた。

ユーザはうまくそれを見つけた場合、ユーザ自身の位置情報をうまく把握できていることが多かった。しかも、指示が遅く通り過ぎてからの案内もあった。さらに実験により調べていく必要があると感じた。

対話の冗長性を見直し

実験2において、対話に冗長性があるとの指摘が多かった。また、少しでも自然な対話を目指そうとして、丁寧な言い回しをしていた。しかしながら、被験者の関心が薄く、伝えなければならない情報だけを言ってもらったほうがよいという分析結果が得られていた。このため、相手に伝える文章の見直しをして、情報量については、実験2と同じにした。

結果は、多少緩和されているものの、あまり変わりのないという意見が占めた。しかしながら、情報量についてはさらに求める傾向が見られた。

情報の重み付け

一連の情報提供において、一番必要な情報を先に持っていくことにした。なぜなら、長い対話において、重要な部分を聞き逃してしまう可能性が高かったからである。例えば、実験2における道案内において、一番必要な情報は何かをユーザに質問した所、「どこに行けばよいのか分かる方向」が多勢を占めるものであった。よって、これらを反映した情報の重み付けを考えた。

結果、先に来る情報がユーザの行動を決定づける重要な要素であることが判明した。

音声合成、認識のパラメータ調整

PS を使うにあたって、情報の入出力である音声は、システムの信頼性を高める上で重要な位置を占めることが、実験1, 2で示されていた。よって、ユーザに聞き取りやすくするため、かつ認識されやすい単語の洗練をした。

その結果、ユーザがなれてきたこともあったが、入力されやすくなってきたという評価がでてきた。また、聞きなおしの回数も減った。

しかし、まだまだ入力されにくい事実があり、原因としては、その場の環境による入力音声の変化があげられる。この使用しているエンジンは、使っている場のノイズを取るといった作業が行われている。つまり、ユーザはどこでも使うという設定ではないので、ノイズ計算の狂いが生じる可能性が高い。調べてみた結果、廊下などは反響音が多く、認識できにくい現象が多数を占めた。

4.4.3 考察

実験2では、「言われなければ答えてくれない」「聞いても同じことを答えてくる」と言った弊害が生まれていた。また、図書館など、音声入力できない場所も存在する。

よって、今回はハンドオーバーした時点で情報を与えた。すると、予想通り絶えず新しい情報をユーザに分け与えることができた。

しかし、CS の位置やその場における環境により情報提供のタイミングが変わってしまい、情報提供に不確実さが増した。特に位置登録が少ないとユーザに与える情報提供も比例して少なくなる傾向があった。

実験1の頃と比べると実験サンプルが採れ、また小実験を繰り返すことにより音声認識しやすい言葉などが解明されてきた。ゆえに、システムにおける信頼性が増してきた。特に、何度も使用してもらっている被験者によっては、スムーズに対話ができるようになっていた。

4.5 実験4: 最終実験

実験1~3を通して、プロトコルデータを集めた。これらを分析して実験4に反映させた。

まず、道案内における情報を与えるタイミングであるが、実験2と実験3の利点を組み合わせ、弱点を補った。つまり、ハンドオーバーした時点で、CommNaviからの情報提供をさせた。また、ユーザからの入力、及び誤差補正も組み入れた。今回は相手の環境情報として、スケジュールのスタブを組んで相手に教えた。

今回の被験者は8人であり、構成は本システムを今までに使ったことがある人が4人、初めて使う人が4人である。また、この被験者らは学内の状態を知っているものである。

なお、今回の実験はアンケート調査によって評価する。また、道案内システムを使ってもらい、ターゲットへたどり着けたかのどうか、またはたどり着けた時の時間的评价をした。条件としては、まず求める人の居場所に十分近くなった時点でたどり着いたとみなした。そして、ルート設定においてCSの配置に対する重み付けは、対して変わらないという意見を採用して行わなかった。さらに、目的とする人は動かないことと、その人が持っているPSによる位置情報はある程度正確なものとした。

4.5.1 定量的評価

結果は、表4.3のとおりである。

表4.3: 対話を始めて目標までたどり着いた時間

ユーザ名	たどり着いたか?	たどり着いた場合の所要時間(分)	CommNavi使用経験	たどり着いた人の平均時間(分)
ユーザA		10	あり	9.25
ユーザB		8		
ユーザC		6		
ユーザD		13		
ユーザE		8	なし	11.75
ユーザF		12		
ユーザG		12		
ユーザE		15		
参考時間	-	2	-	-

なお、参考時間は、JAIST内の構造を知っている人が標準に目的地まで歩いた時間である。まず、JAIST内を知っている被験者であるためか、全員がたどり着いていた。また、CommNaviを使用した経験のあるユーザの方が早くたどりついていた。理由としては、CommNaviを始めて使う人が音声認識をすることにてこずっている印象があった。その点、なれている人は歩きながら情報を引き出す操作を多くしていた。また、経験者

のほうがより多くのショートカットを使っていた．その差が時間となって出ているのではないかと分析した．

4.5.2 定性的評価

定性的評価に関しては，アンケートを使って評価してもらった．その項目と，結果を以下に示す．

音声合成

音声合成に関して，ストレスを感じていたかを質問した．その結果，表 4.4 ようになった．

表 4.4: 音声合成のストレスを感じたか？

	はい	多少	いいえ	合計
CommNavi 経験者	1	2	1	4
CommNavi 未経験者	1	2	1	4
合計	2	4	2	8

これより，経験者，未経験者に限らず多少ストレスを感じている人が多かった．また，経験者は慣れているせいか，聞き直しをする回数が未経験者に比べて少なかった．さらに，突然話しかけられて困るという意見もあった．回避する方法として，余剰語の付加，または人間の肉声をデータベースに蓄えて使う手法が挙げられる．

音声認識

音声認識に関して，ストレスを感じていたかを質問した．その結果，表 4.5 ようになった．

認識に関しては，予想通りストレスを感じている人が多かった．さらに時の心理的状态によって，後の CommNavi との対話の姿勢にかなり影響を及ぼしていたものだった．今回，特定話者認識エンジンで PS を介した音声認識だったので多少無理があったよう

表 4.5: 音声認識のストレスを感じたか？

	はい	多少	いいえ	合計
CommNavi 経験者	2	1	1	4
CommNavi 未経験者	3	1	0	4
合計	5	2	1	8

に感じた．そして，システムにおける入出力の重要さが浮き彫りになった．しかしながら，携帯からの音声認識もできるようになってきたので，そちらのエンジンを使えば認識率の向上，及びシステムにおける信頼性の向上が期待される．

音声による情報提供

情報の提供に関しては，量と質の視点から意見を出してもらい，主要を以下にまとめた．

情報の量

- 適量である (4/8)
- CommNavi からの話が少し長い状態になる時があった (2/8)
- まだ話全体が多すぎる (1/8)
- 少ない (1/8)

この実験では，道案内時に与える情報は，ユーザまたは相手の位置情報，方向，距離と正しく進んでいるかの確認，特殊な環境情報だった．さらに，聞き逃した時は，少し短めにした対話を提供した．

なるべくシンプルな情報提供を心がけたが，まだ量が多すぎるという意見もあった．対策は，情報提供の形を文章とは違い重要な単語だけを言う方法が挙げられる．しかし，いきなり重要な単語を言われてもユーザの心構えができておらず聞き逃しが多く発生することが予想される．対策として，対話の中でユーザにカスタマイズされた情報提供ができるような条件分岐を作ってもよい気がした．

気になった意見としては、対話の時間と、沈黙の時間がはっきりしすぎているので、もう少し甘くしてもよいのではないかという提案が出されていた。これは、相手が聞き逃した時に、もう一度聞き直す時間を与えたつもりであったし、継続的な情報提供であると、必要な情報を抽出しにくいのではないかと思う所からきている。どちらがいいのか対話において選択させることが必要であるように感じた。

情報の質

- 的確である、またはある程度的確だった (4/8)
- 位置情報で若干のずれが生じた (3/8)
- 誤差情報が認識されなかったので不正確だった (1/8)

道案内における他のアプローチによる情報提供の要望はなかった。結果としては的確、ある程度は的確であるという人が50%占めた。道案内では、誤差補正をメインにして使用する人が多かった。つまり、通る道にある部屋のIDを見つけたい積極的に音声入力して、補正していく傾向が見られた。よって、ハンドオーバにおける位置情報は補助的なものとなっていた。

今回は、スケジュールのスタブを作り、相手に提供した。スケジュールの内容は、目的とした人の現在の予定、及びそれは何時まで続くかという内容であった。その結果、未経験者でも、スケジュール情報を聞き出してその場所に歩き出せており、便利だったという評価があった。しかし、人によってはスムーズに情報を引き出せる人と出せない人がいた。その要因としては音声認識の精度によるものであった。

対話の流れ

対話の流れについて、意見を述べてもらった。

- 対話の流れはスムーズであった (6/8)
- 対話した気がしない (2/8)

対話の流れとしては、スムーズであったという意見が多かった。

残りの人の意見としては、自分の音声を拾ってくれなかったことが大きな要因であるとの意見が出されていた。

今回の実験では対話のショートカットをつけていた。例えば、図 3.2 では、道案内時に相手のスケジュールを聞くことはできない。しかし、ショートカットを作ることにより、短時間でより多くの情報を引き出すことが必要であると感じ、この機能を付加した。よって、道案内をしながら、相手のスケジュールを聞けるという作業、または、いきなり人の名前を言って情報を引き出すという動作が可能になる。

未経験者の場合は、そのショートカットをしなかった。原因として、CommNaviに入力できる言葉を知らなかったことにある。それに対して経験者は、ある程度使った人ならショートカットを簡単に用いる傾向が見られた。これは、使用経験を重ねることによって対話の内容を覚えていることから来ていた。しかし、この動作は表 1.1 の会話における協調性の原則から逸脱したものである。これに対して機械はユーザの答えに対して原則及び公理を乱さずに答えている。よって、情報引き出しのスピードにおける視点からは、機械との対話において人間の学習に合わせた会話を構築する必要があった。

CommNavi の来訪者に対する有効評価

本システムは、初めて来られる来訪者に対しても環境情報の提供、道案内などを行うことにより、さらに力を発揮するものである。

しかしながら、今回は、JAIST の構造に関する知識を持っている被験者しか集めることができなかった。そこで、アンケート項目に、来訪者を含めた場合、本来の有効性が保てるかに関して質問した。

- 有効性は保てるものである (5/8)
- 音声認識の精度が上がれば期待できる (2/8)
- 視覚情報がほしい (1/8)

以上のような結果になった。これより、有効性があると感じている人、音声認識の向上によって期待できる人を合わせると 87.5% になった。よって、来訪者に対しても高い有効性がある可能性が分かった。

4.5.3 考察

この実験においては、音声インターフェイスのみを使っての道案内は有効であることが分かった。さらに、情報の引き出しにおいては、人間の学習形態に合わせた会話の構築があった。そのために、構築された会話を認知している人間側は協調性ある会話をする必要がなくてもよいことが分かった。しかし、機械は、会話の原則と公理を守っていないとユーザに不信感を与える結果となっていた。

第 5 章

結論

5.1 本研究の成果

本論文では，Positioning Server と計算機，PS を組み合わせ，音声インターフェイスで情報を提供するシステム「CommNavi」を構築した．

これにより，研究者間にある環境の溝を埋め，相手の状況を知ることによりユーザの持つ時間を有効に使えるものである．また，JAIST 環境を知らない来訪者にも環境情報の提供，および道案内の支援ができるものを目指した．

本研究で提案し，実現したことを以下に示し，今後の課題と応用を述べる．

- 音声インターフェイスでナビゲート

ユーザは JAIST 内の情報を音声のみで引き出し，目的地までたどり着くことができた．これは，現在主流であるディスプレイを用いなくて，かつ音声のみでのナビゲートができることを示す結果である．

- 道案内においての情報を与えるタイミング

Positioning Server のみでの情報提供は状況に応じて不確定な要素が多く信頼性に欠けるものだった．また，情報提供のタイミングは，このシステムから出される状況変化の情報が流れた時，及びユーザが情報をほしがっている時に与えることがベストであった．

- 誤差補正

Positioning Server のみでの位置情報は曖昧なものであることが分かった。よって、ユーザが曖昧さを感じたときには、ユーザからの誤差補正をさせた。つまり、ユーザがどこにいるか、何が見えるかを言わせて判断するものである。この時の注意点として、ユーザの知識表現に左右されることの少ないものを選ぶ必要があった。この誤差補正を取り入れるとシステムにおける信頼性が飛躍的にあがり、情報の質が向上した。

- 距離の算出

ユーザに道案内させる時に最短距離ルート設定をし、ユーザにその情報を伝えた。正しい方向に歩いているかを確認するため距離の算出を行い、その距離に応じてユーザに指示を与えた。ユーザのストレス軽減につながった手法であった。

- 特殊な環境への対応

エレベーター内、CS の範囲外の環境においては、事前にユーザに伝えた。これにより、ユーザに心構えを持たせることができた。また、この特殊な環境におけるルート設定時の重み付けは、賛成意見と反対意見に分かれた。

- 目標物の指示

ユーザの通るルート上、目立つものがあると、それに関する情報を流した。しかし、誤解や勘違いを招く恐れが多いため、本当に目立つものを中心とした目印となるものだけを取り入れた。この点はディスプレイを使ったナビゲートと大きく異なる所であった。

- 会話の構築

対話の流れにおいては、対話の協調性の原則と公理を基底にして構築した。また、コンセプトとしては、情報の引出しのスピードに焦点を絞った。よって、音声入力の特長を生かすため、ショートカットなど、対話の流れにかなわず沿わなくても情報を引き出せるようにした。結果、初

めて使用するユーザから経験者までを視野に入れた対話の構築ができた。
これによって、速やかな情報取得ができるようになった。

5.2 今後の課題

今後の課題として、以下のことがあげられる。

- 相手側の位置情報に関する誤差補正

ユーザ側に関する誤差補正は強力なものであった。しかし、相手側の誤差補正も視野に入れなければならない。その対策として、位置情報ブローカーアーキテクチャ[丹 97]を取り入れる方法があげられる。

- PS を持っていない人への対応

本研究では、目的とする相手側は必ず PS を持っているという前提であった。しかし、実際には、持っていない人や、PS を持ち歩かない人も多数存在する。よって、それらを考慮したスケジュール作りの試みが必要である。これに関しては事情通口ボットに関する研究[本村 99]を取り入れる方法が考えられる。今回はスケジュールのスタブを作成しユーザに提供したが、実際に組み入れて問題を抽出した方がよいと感じた。

- プライバシーとセキュリティー

位置情報は、大変デリケートなものであり、その扱い方次第ではプライバシーを侵害してしまう恐れが十分にある[桜井 99]。今回は、人によってアクセス権をつける対処であったが、さらに暗号化などの処理を用いる必要がある。

また、音声による人の認識に関してもなりすましなどのセキュリティーの点で危ぶまれる。音声によるパスワードと平行して声紋分析のチェック等を視野に入れる必要がある。

5.3 応用

本研究は、JAIST 内に特化したシステムであった。しかし、大きな建物の中で PS と位置情報の使い道である分野はさらに残っている。例えば、大企業では、その巨大さに特定の人に会うことや、他人の環境情報が分かりにくい状態にある。また、巨大なデパートでは特定の場所へ行く方法は看板や標識を見て判断するしかない。つまりそこまで行かなくてはならず、探すのにも労力が必要である。

これらの問題に対しては、本システムをカスタマイズすることによって、便利なツールになると考えられる。

今後の発展として、本研究での成果を生かすとともに、音声認識の精度を上げるなどさらに充実した環境を整えることによって、開かれたオフィス環境の支援、および個人化戦略に特化した知識共有の場の形成に多大な貢献ができることを期待する。

謝辞

本研究を進めるにあたり，お世話になった方々に謝辞を述べたいと思います．

終始，御指導を賜り，快適な研究環境を配慮していただいた指導教官の藤波努助教授に深謝致します．日頃から有益な御助言をいただき，多方面に渡って励ましていただいた國藤進教授に感謝致します．

また，研究設備におきましては，多大な御支援をいただきました，丹康雄助教授に感謝致します．

そして，本論文をまとめるに当たってご協力を頂いた國藤・藤波研究室の方々に心から厚く御礼申し上げます．

最後に私事で恐縮ですが，これまでの学生生活を支えてくれた両親に深く感謝致します．

平成 12 年 2 月 12 日

渡辺 功

参考文献

- [野中 96] 野中 郁次郎, 竹内 弘高, 梅本 勝博/訳: “知識創造企業”, 東洋経済新報社, 1996;
- [Tierney 99] T. Tierney, N. Nohria, M. T.Hansen, 黒田由貴子/訳: “コンサルティング・ファームに学ぶ「知」の活用戦略”, Diamond Harvard Business, Vol.9, pp60-74, 1999;
- [山梨 86] 山梨 正明: “対話理解の基本的側面”, 対話行動の認知科学的研究研究会編, 1996;
- [Grice 75] H. P. Grice: “Logic and Conversation”, In P. Cole and J. Morgan(eds.), Syntax and Semantics, Vol.3, 1975;
- [有田 87] 有田 英一, 小暮 潔, 野垣内 出, 飯田 仁: “メディアに依存する会話の様式”, 情報処理学会研究報告 (NL61-5) , 1987;
- [Cohen 84] P. Cohen: “The Pragmatics of Referring and the Modality of Communication, Computational Linguistics, Vol.10, No. 2, 1984;
- [堂下 98] 堂下 修司, 新美 康永, 白井 克彦, 田中 穂積, 溝口 理一郎: “音声による人間と機械の対話”, 株式会社 オーム社, 1998;
- [田中 99] 田中 和世: “音声処理応用システム・音声インターフェイス”, 財団法人 日本情報処理開発協会 先端情報技術研究所 人間主体の知的情報技術に関する調査研究 , pp.44-51, 1999;

参考文献

- [本村 99] 本村 陽一, 松井 俊浩, 麻生 英樹, 浅野 太, 原 功, John Fry: “事情通口ロボットによるオフィス環境における知的作業支援”, 人工知能学会 AI シンポジウム Technical Report, 1998;
- [Nagao 96] K. Nagao, J. Rekimoto: “Agent Augmented Reality: A Software Agent Meets the Real World.”, In Proceedings of the Second International Conference on Multi-Agent Systems (ICMAS-96), pp.228-235, 1996;
- [丹 97] 丹 康雄, 日比野 靖: “位置情報ブローカの提案とそのアーキテクチャ”, モバイルコンピューティング研究会 (MBL) 第 2 回 研究報告会, 1997;
- [小山 97] 小山 俊明, 三宅 亜男, 星 徹: “コンピュータ・テレフォニー新時代をリードする CTNET-Server システム”, 日立 IT 技報, Vol.1, 1997;
- [海保 93] 海保 博之, 原田 悦子: “プロトコル分析入門 発話データから何を読むか”, 株式会社 新曜社, 1993;
- [辻 97] 辻 新六, 有馬 昌宏: “アンケート調査の方法 実践ノウハウとパソコン支援”, 株式会社 朝倉書店, 1987;
- [桜井 99] 桜井 哲真, 萩野 輝雄, 鈴木 義武, 西野豊: “超小型端末の光と影 位置情報サービス ”, 電子情報通信学会誌, Vol.82, No.4, pp.397-403, 1999;
- [後藤 99] 後藤 敏, 阪田 史郎: “ポイント図解式 モバイル・コンピューティング教科書”, 株式会社 アスキー, 1999;
- [村田 99] 村田 嘉利, 中川 純一: “移動体データ通信がわかる本”, 株式会社 毎日コミュニケーションズ, 1999;
- [NEC] 日本電気 (株): “でんわなび”,
<http://www.psinfo.nec.co.jp/dennavi/index.htm>;
- [BizIT] 日経 BP 社: “CTI サイト”,
<http://bizit.nikkeibp.co.jp/it/cti/cti.shtml>;

参考文献

- [IDO] 日本移動通信株式会社: “乗換ナビ”,
<http://www.ido.co.jp/square/voice/navi.html>;
- [DoCoMo] NTT 移動通信網株式会社: “いまどこサービス”,
<http://www.nttdocomo.co.jp/mc-user/phs/ichi.html>;

発表論文

- [渡辺 99] 渡辺 功, 藤波 努, 丹 康雄: “携帯電話で使える面会支援システム”, 人工知能学会, AI シンポジウム (SIG-J-9901), pp.85-90, 1999;