

資 料

人間主体の知的情報技術  
に関する調査研究 V

平成 14 年 3 月

財団法人 日本情報処理開発協会  
先端情報技術研究所

**KEIRIN**



この事業は、競輪の補助金を受けて実施したものです。

## 人間主体の知的情報技術調査ワーキンググループ

|    |        |                                              |
|----|--------|----------------------------------------------|
| 主査 | 奥乃 博   | 京都大学 大学院情報学研究科 知能情報学専攻 教授                    |
| 委員 | 大須賀 昭彦 | (株)東芝 研究開発センター コンピュータ・ネットワーク<br>ラボラトリー 主任研究員 |
| 委員 | 加藤 俊一  | 中央大学 理工学部 経営システム工学科 教授                       |
| 委員 | 國藤 進   | 北陸先端科学技術大学院大学 知識科学研究科 教授                     |
| 委員 | 児島 宏明  | 産業技術総合研究所 電子技術総合研究所 知能情報部<br>主任研究官           |
| 委員 | 杉本 重男  | 図書館情報大学 図書館情報学部 教授                           |
| 委員 | 長尾 確   | 名古屋大学 工学研究科 情報工学専攻 助教授                       |
| 委員 | 新田 克己  | 東京工業大学 大学院総合理工学研究科<br>知能システム科学専攻 教授          |
| 委員 | 平田 圭二  | NTT コミュニケーション科学基礎研究所<br>メディア情報研究部 特別研究員      |
| 委員 | 宮田 一乗  | 東京工芸大学 芸術学部 メディアアート表現学科 助教授                  |
| 委員 | 山名 早人  | 早稲田大学 理工学部 情報学科 助教授                          |
| 幹事 | 北島 重信  | (財)日本情報処理開発協会<br>先端情報技術研究所 技術調査部 主任研究員       |
| 幹事 | 伊藤 文英  | (財)日本情報処理開発協会<br>先端情報技術研究所 技術調査部 主任研究員       |

# 人間主体の知的情報技術に関する調査研究 V

## 目次

|                                                                  |    |
|------------------------------------------------------------------|----|
| 1. 総論.....                                                       | 1  |
| 1.1 調査ワーキンググループの活動方針 .....                                       | 1  |
| 1.2 調査の概要 .....                                                  | 2  |
| 1.2.1 調査方針 .....                                                 | 2  |
| 1.2.2 調査対象 .....                                                 | 2  |
| 1.3 調査活動の方針 .....                                                | 5  |
| 1.4 調査結果の要約 .....                                                | 6  |
| 1.5 調査結果の概要 .....                                                | 7  |
| 1.5.1 ロボットの耳は2つで十分か .....                                        | 7  |
| 1.5.2 エージェントの研究動向 .....                                          | 7  |
| 1.5.3 人間中心の情報環境 —感性的相互作用の視点から— .....                             | 8  |
| 1.5.4 新世代グループウェアのためのアウェアネス研究の最新動向.....                           | 8  |
| 1.5.5 音声情報処理の本格的アプリケーション構築<br>に向けての展望と課題 —VoiceXML を中心として— ..... | 9  |
| 1.5.6 Digital Libraries, Metadata, Interoperability .....        | 9  |
| 1.5.7 デジタルコンテンツのトランスコーディングとインタフェースロボット.....                      | 10 |
| 1.5.8 法的推論システムの研究動向と新しい展開 .....                                  | 10 |
| 1.5.9 音楽情報処理研究の新しい潮流.....                                        | 10 |
| 1.5.10 インタラクティブグラフィックスの動向.....                                   | 11 |
| 1.5.11 検索履歴を用いた情報検索サポート .....                                    | 11 |
| 1.5.12 Webマイニングとその将来像 .....                                      | 11 |
| 1.5.13 3次元視覚システムの技術動向と展望 .....                                   | 12 |
| 2. 米国政府が支援する情報技術研究開発の動向.....                                     | 13 |
| 2.1 はじめに .....                                                   | 13 |
| 2.1.1 米国連邦政府研究開発予算の動向 .....                                      | 13 |
| 2.1.2 研究開発投資効率の改善 .....                                          | 16 |
| 2.2 ネットワーキングおよび情報技術研究開発計画.....                                   | 18 |
| 2.2.1 2002年度の概要 —BlueBook 2002.....                              | 18 |
| 2.2.2 プログラムコンポーネント領域 (Program Component Area: PCA) .....         | 19 |
| 2.2.3 NIITRD 計画の計画・執行組織 .....                                    | 22 |
| 2.2.4 2003年度の重点課題.....                                           | 24 |

|       |                                                                |     |
|-------|----------------------------------------------------------------|-----|
| 2.3   | ヒューマンコンピュータインタフェースおよび情報管理領域.....                               | 25  |
| 2.3.1 | IITR (Information Technology Research) プログラム.....              | 27  |
| 2.3.2 | デジタルライブラリ・イニシアチブ<br>(電子図書館構想、Digital Library Initiative) ..... | 38  |
| 2.4   | おわりに.....                                                      | 40  |
| 3.    | 研究開発の新しい展開と内外の動向.....                                          | 43  |
| 3.1   | ロボットの耳は2つで十分か.....                                             | 43  |
| 3.1.1 | はじめに.....                                                      | 43  |
| 3.1.2 | 我々が日常聞く音は混合音.....                                              | 44  |
| 3.1.3 | 複数話者実時間追跡システム.....                                             | 48  |
| 3.1.4 | おわりに.....                                                      | 53  |
| 3.2   | エージェントの研究動向.....                                               | 55  |
| 3.2.1 | エージェントとは.....                                                  | 55  |
| 3.2.2 | エージェントの移動性.....                                                | 56  |
| 3.2.3 | エージェントの協調性.....                                                | 59  |
| 3.2.4 | エージェントの自律性.....                                                | 61  |
| 3.2.5 | その他の特性.....                                                    | 63  |
| 3.3   | 人間中心の情報環境 —感性的相互作用の視点から—.....                                  | 66  |
| 3.3.1 | はじめに —情報システムにおける「人間」の位置づけ—.....                                | 66  |
| 3.3.2 | インタラクションの形態.....                                               | 67  |
| 3.3.3 | インタラクションのモデル化.....                                             | 73  |
| 3.3.4 | 事例紹介 インテリアコーディネート支援システム.....                                   | 81  |
| 3.3.5 | 人間中心型の情報技術の研究開発動向.....                                         | 83  |
| 3.4   | 新世代グループウェアのためのアウェアネス研究の最新動向.....                               | 87  |
| 3.4.1 | はじめに.....                                                      | 87  |
| 3.4.2 | アウェアネス研究の分類と動向.....                                            | 88  |
| 3.4.3 | 浅い感性を支援する個別アウェアネス研究.....                                       | 90  |
| 3.4.4 | 浅い感性を支援する統合アウェアネス研究.....                                       | 93  |
| 3.4.5 | 深い感性を支援するアウェアネス研究.....                                         | 97  |
| 3.4.6 | おわりに.....                                                      | 100 |
| 3.5   | 音声情報処理の本格的アプリケーション構築に向けての展望と課題<br>—VoiceXML を中心として.....        | 104 |
| 3.5.1 | はじめに.....                                                      | 104 |
| 3.5.2 | 音声情報処理技術とアプリケーションの動向.....                                      | 104 |
| 3.5.3 | 音声ポータル現状.....                                                  | 105 |
| 3.5.4 | VoiceXML.....                                                  | 107 |

|                                                         |     |
|---------------------------------------------------------|-----|
| 3.5.5 音声アプリケーションの課題 .....                               | 111 |
| 3.6 Digital Libraries, Metadata, Interoperability ..... | 114 |
| 3.6.1 はじめに .....                                        | 114 |
| 3.6.2 Digital Library .....                             | 115 |
| 3.6.3 メタデータについて .....                                   | 121 |
| 3.6.4 Interoperability .....                            | 129 |
| 3.6.5 おわりに .....                                        | 131 |
| 3.7 デジタルコンテンツのトランスコーディングとインタフェースロボット .....              | 135 |
| 3.7.1 デジタルコンテンツのトランスコーディング .....                        | 135 |
| 3.7.2 セマンティックトランスコーディング .....                           | 136 |
| 3.7.3 アノテーションとトランスコーディング .....                          | 138 |
| 3.7.4 インタフェースロボット .....                                 | 150 |
| 3.7.5 会話ロボット Pong と QB .....                            | 151 |
| 3.7.6 おわりに .....                                        | 156 |
| 3.8 法的推論システムの研究動向と新しい展開 .....                           | 159 |
| 3.8.1 はじめに .....                                        | 159 |
| 3.8.2 法的推論の概要 .....                                     | 159 |
| 3.8.3 法的推論研究の動向 .....                                   | 160 |
| 3.8.4 新しい展開 .....                                       | 166 |
| 3.8.5 まとめ .....                                         | 168 |
| 3.9 音楽情報処理研究の新しい潮流 .....                                | 170 |
| 3.9.1 はじめに .....                                        | 170 |
| 3.9.2 計算機への実装に有望な音楽理論 .....                             | 170 |
| 3.9.3 音楽知プログラミング .....                                  | 174 |
| 3.9.4 関連する国際会議 .....                                    | 178 |
| 3.10 インタラクティブグラフィックスの動向 .....                           | 181 |
| 3.10.1 はじめに .....                                       | 181 |
| 3.10.2 インターネット上のグラフィックス .....                           | 181 |
| 3.10.3 リアルタイム CG のハードウェア技術 .....                        | 182 |
| 3.10.4 家庭用ビデオゲームの現状 .....                               | 185 |
| 3.10.5 インタラクティブグラフィックスの応用事例 .....                       | 187 |
| 3.10.6 おわりに .....                                       | 188 |
| 3.11 検索履歴を用いた情報検索サポート .....                             | 190 |
| 3.11.1 はじめに .....                                       | 190 |
| 3.11.2 情報フィルタリング .....                                  | 191 |
| 3.11.3 ユーザごとの情報検索サポート .....                             | 192 |

|        |                                      |     |
|--------|--------------------------------------|-----|
| 3.11.4 | 連続するクエリの出現パターンの分類とユーザの検索意図の推定.....   | 196 |
| 3.11.5 | おわりに .....                           | 201 |
| 3.12   | Web マイニングとその将来像 .....                | 203 |
| 3.12.1 | はじめに .....                           | 203 |
| 3.12.2 | Web マイニングとその概要.....                  | 205 |
| 3.12.3 | おわりに～次世代 Web に向けて.....               | 211 |
| 3.13   | 3次元視覚システムの技術動向と展望 .....              | 218 |
| 3.13.1 | コンピュータビジョンの歴史 .....                  | 218 |
| 3.13.2 | 汎用システムの必要性.....                      | 219 |
| 3.13.3 | 3次元視覚システム .....                      | 220 |
| 3.13.4 | モデルの階層 .....                         | 224 |
| 3.13.5 | 応用システム .....                         | 224 |
| 3.13.6 | 将来への展望 .....                         | 227 |
| 付属資料 1 | 海外調査報告.....                          | 229 |
| 付属資料 2 | 自然言語処理の最近の動向 — 「言語理解と行動制御」を中心に—..... | 241 |

# 1. 総論

## 1.1 調査ワーキンググループの活動方針

この数年の間にパソコンや携帯電話(携帯端末)およびインターネットが急速に普及し、社会は情報技術 (IT) による大きな変革の波にさらされている。現在、わが国においても IT 革命が叫ばれているが、これを真に成功させ、国民の幸福に結びつける為には、インターネット上の価値あるコンテンツを効果的に開発・維持・検索・活用するための技術と仕組み、および、人間とコンピュータとの間のインタフェースの能力を現在のレベルよりも格段に優れたものへと飛躍的に向上させることが必要である。

従来の情報技術は、コンピュータの高速化、高性能化、高信頼化、大容量化などを廉価に実現し、人間の作業を機械で置き換えることによって、いかにしてコンピュータを既存の人間社会に適応させて行くかが主要課題であったと言えよう。これらは依然として重要であることに変わりはない。しかしながら、インターネットの出現により生活様式の変化や、新たな文化の創出が生まれつつある現実を見据え、情報技術によるより良い社会を実現するためには、新しい「人間・コンピュータ共存系」におけるコンピュータの役割、すなわちコンピュータが人間に与えるもの(メリット)に、より一層着目する必要がある。これは、インターネットアプリケーションも含めて、広義の「コンテンツ」ともとらえることができよう。

また従来、コンピュータ化しやすい部分は非常に大きな自動化の効果をあげてきたが、知的精神活動に直接かかわるような、コンピュータ化しにくいと言われてきた部分においては、いまだに多くの研究余地が残されており、感性の扱いなどを含め、さまざまな要素を考慮したユーザインタフェースのさらなる向上が期待される。さらに、B-to-C の電子商取引や G-to-C の電子政府、そして教育などにおいては、身障者、老人、子供を含めたすべての人が快適に情報技術の恩恵を受けられるような知的なユーザインタフェースの実現が必須であろう。

以上の視点から、調査ワーキンググループでは情報技術の重要領域として、「知的ユーザインタフェース」と「広義のインターネットコンテンツ」に特に重点を置いた「人間主体の知的情報技術」の調査・検討を行うことにした。今年度は各分野の専門家からなる委員や講師の最新知識を集積して、これらの領域における研究開発のリーディングエッジを探り、これを元に今後注力すべき技術分野の検討や、世界におけるわが国の技術ポテンシャルの水準を評価するための元データを得ることを目指した。



## 1.2 調査の概要

調査ワーキンググループで取り上げるべき人間主体の知的情報技術は、整理すると次の4つになろう。

- ・ 計算機と通信ネットワークを通じて、人々が各種情報を容易に利用できるようにする技術。
- ・ 人間の知的活動を増幅する技術。
- ・ 使いやすさ、使い心地のよさを重視する技術。
- ・ 社会生活で幅広く利用できる情報環境を提供する技術。

### 1.2.1 調査方針

本年度は、次の調査方針に基づき、昨年度に引き続き、調査を進める。

- ・ 日本で国の研究プロジェクトを走らせる場合を想定し、研究テーマの新しい切り口を示して情報技術の研究シナリオを描く。
- ・ わが国が研究レベルにおいて米国などと互角のものを選んで調査分析し、わが国の有する人間主体の知的情報技術のどれを伸ばしていけばよいかを明らかにする。
- ・ 潜在需要や需要シフトをとらえ、研究を実施する動機を明確化する。
- ・ 技術がどのように役立つかを、アプリケーションなどを示して明らかにし、社会への波及効果を分析する。

特に、昨年度調査が不足していたような事項を掘り起こすことにより、IT関連技術のより広範囲でより深い報告を行うように努める。

### 1.2.2 調査対象

技術分野を次の3つの技術を軸として整理する。

- ・ 情報端末としてのコンピュータシステム（パソコン、モバイル端末など）。
- ・ インフラストラクチャとしてのネットワーク、データベース、情報リポジトリ、知識ベース。
- ・ マルチメディアを対象とするヒューマンインタフェースおよび人工知能技術。

調査ワーキンググループでは主にこれらのミドルウェア層、および、ユーザインタフェース層を構成すると思われるソフトウェア技術、および関連する基礎技術に重点をおいて調査する。

調査対象と考えられる研究分野について以下にリストアップする。これらの分野について、研究テーマの実現上の問題点、利点、社会的インパクト、研究開発投資額、期間などを分析する。

また、5年から10年先における基礎技術の開発に注目し、現在、商品が出ている領域は除外し、将来における土台となる技術をリストアップしていく。

(1) ネットワーク上の処理を含むデータベース技術

- ・ 探索境界がオープンな分散データベース管理
- ・ 情報リポジトリ
- ・ インターネット上の情報検索
- ・ 情報フィルタリング
- ・ コンテンツ処理、意味理解
- ・ データマイニング

(2) ネットワークを含めたコンピュータの新しい利用形態

- ・ エージェント指向コンピューティング
- ・ 発想支援、知的創造支援環境
- ・ 共同作業支援

(3) マルチモーダルインタフェース技術や関連する人工知能技術

- ・ 音声、図形、画像などのパターン認識処理や知識処理技術の利用
- ・ 操作の簡単化のための人工知能技術の利用
- ・ 状況依存処理
- ・ 仮想現実感
- ・ 拡張現実感
- ・ 感性情報処理

(4) モバイルコンピューティングなどの新しいコンピュータ技術と利用

- ・ 端末となるコンピュータとその機能（ユーザインタフェース）
- ・ ウェアラブルコンピュータ
- ・ ワイヤレス
- ・ ネットワーキング（媒体とプロトコル）

(5) 社会サービスおよびそれを構成するに必要な情報処理技術

- ・ 電子図書館
- ・ 電子美術館
- ・ 遠隔教育システム
- ・ 遠隔医療診断システム
- ・ 高齢者介護
- ・ その他の公共サービス

(6) その他

調査ワーキンググループでは、主にこれらのミドルウェア層、および、ユーザインタフェース層を構成すると思われるソフトウェア技術および関連する基礎技術に重点をおいて調査を行うこととした。また、5～10年先における基礎技術の開発に注目し、現在、商品が出始めようとしている領域は除外し、将来の基盤技術あるいは新しい技術の萌芽となるような技術あるいは研究テーマをリストアップするように心がけた。

上述したように、人工知能およびネットワークに関連する研究分野は拡大し、発展している。このような状況においては、わが国の研究開発力がすべての分野において卓越することは不可能と言っても過言ではないであろう。したがって、分野を選別し、選別した分野に人材費用という資源を集中化するという戦略が重要となる。調査ワーキンググループでは、そのような戦略をたてるための調査を主眼とする。

### 1.3 調査活動の方針

具体的な議論の第 1 段階として、次のことを検討する。すなわち、中長期的な研究は、将来の産業に技術シーズとなるようなものが望ましく、米国のネットワーキング・情報技術研究開発計画などを参考にして、わが国として重要と思われる分野やテーマを選択し、それらの研究や技術内容の特徴や水準を分析し、わが国の技術的な位置づけの評価と今後の取り組み方について議論する。つまり、主査および委員と幹事は、以下のような方針で調査に望むこととする。

主査および委員は、専門として取り組んでいる分野、興味のある分野、さらには将来的に市場へのインパクトがありそうな分野について調査を行い、それらの分野におけるわが国の技術的な位置づけの評価を行う。そして、当該分野に対して、わが国は国策としてどのような取り組みをしたらよいかについて議論を行う。

幹事は、米国のネットワーキング・情報技術研究開発計画の中身や進捗情報に関して、Blue Book 2002 などの資料を元に調査を行い、また、関連する技術資料などの情報を調査ワーキンググループに提供する。

調査ワーキンググループの具体的な活動は、次の通りである。

#### (1) 海外調査

大須賀委員による International Conference on Autonomous Agents 参加、Carnegie Mellon 大学および SRI International 訪問。

#### (2) 第 1 回会議

活動方針の討議。大須賀委員による「Agents2001 会議報告と最近のエージェント技術動向について」の報告。

#### (3) 第 2 回会議

田中講師による「自然言語処理の最近の動向 —"言語理解と行動制御" を中心に—」の講演。新田委員による「法的推論の新しい展開」の報告。

#### (4) 第 3 回会議

宮田委員による「インタラクティブ CG の動向」の報告。平田委員による「音楽情報処理の新しい展開と SoundComplete プロジェクト構想」の報告。

#### (5) 第 4 回会議

富田講師による「3 次元視覚システムの技術動向と展望」の講演。津田講師による「ウェブマイニング技術の動向」の講演。

#### (6) 第 5 回会議

各委員による報告と討議。

## 1.4 調査結果の要約

本ワーキンググループの調査活動を通じて、以下のような特徴が明らかになってきた。

(1) コンテンツの高度利用を促進するために、情報提供側が提供すべきコンテンツに対してより正確な情報を付加するための枠組みが重要であるという方向性がより明確になってきた。

(2) 情報収集や情報検索の高速化や高度化のための技術開発で、競争が激化している。

(3) 組織が保有するナレッジ（知）を創造し、共有し、再利用するプロセスを情報処理技術を活用して高速化し、効率化する「ナレッジマネジメント」が新たなビジネス展開法として不可欠である。特に、ナレッジには形式知のほかに暗黙知があり、両者がなくてはビジネスへの展開はできない。この暗黙知については、日本が一早く注目し、その重要性を訴えてきたものであり、それが海外で注目を浴びて、情報処理技術と統合されてナレッジマネジメントとして提案されている。

(4) 情報提示にあたっては、いかに自然に利用者に情報を提示するかが利用者拡大へのキーワードであり、そのために従来にない技術の開発が不可欠である。また、情報入力についても、自然なインターフェースが重要である。

## 1.5 調査結果の概要

調査報告では、下記の点に留意した。

- ・ 特徴、利点、実現にあたっての問題点、実現時期、必要なインフラストラクチャ。
- ・ 現在の研究状況、研究体制（産学協同、国際協同研究など）。
- ・ 普及の条件、情報産業や社会への波及効果。
- ・ 日米比較（技術力、市場の受け入れの容易さ、インフラストラクチャなど）。

報告の概要を以下に示す。詳細については3章に記載する。

### 1.5.1 ロボットの耳は2つで十分か

ロボットが人間社会の中に入り込み、共生していくためには、混合音が扱えること、アクティブオーディション、動きながら聞く機構、未知環境での音の知覚、画像処理などのほかの処理との統合、実時間処理が大きな課題である。混合音の処理では、音源定位が重要であり、頭部伝達関数（HRTF）が使用されることが多い。しかし、HRTFは環境での音響的な伝達特性と組み合わせ使用しなければならないので、音響的な特性が動的に変化する実環境では使いにくく、HRTFを使わない手法が必要となる。そこで、これらの課題に対して、マイクロフォン2本で十分かについて考察を行い、2本のマイクロフォンで実現可能な機能について報告する。体を動かして聞くというアクティブオーディション、あるいは、画像処理とモータ処理を統合して、体全体で聞くという情報統合が重要である。そのために、方向通過型フィルタや聴覚エピソード幾何学、実時間処理方法を開発し、複数の実験で有効性を確認した。

### 1.5.2 エージェントの研究動向

最近「エージェント」という言葉はさまざまな定義で用いられている。本報告では、プログラミング言語としてのエージェントをオブジェクト指向言語の次の技術という位置づけでとらえる。プログラミングパラダイムの変遷は、計算の世界と現実の世界を近づけるための抽象化の歴史とみることができ、オブジェクトの次に位置づけられるエージェントもこの延長上にあると考えられる。オブジェクト指向にどのような特性が加えられ、それによって何が抽象化されたかという観点でエージェントを分類し、分類ごとの研究動向や事例を紹介する。

### 1.5.3 人間中心の情報環境 —感性的相互作用の視点から—

本報告では、「人間の感性」という切り口から、「人間中心」の「人にやさしい」情報システムの実現に重要な役割を果たす、人間と情報機器・情報システムとのインタラクションに焦点を絞って、技術課題の分析と最近の研究開発の話題を紹介する。

特に、情報提供サービスにおける利用者とデータベースとのインタラクションを、インタラクションの形態と、インタラクションを構成する要素、システムが提供すべき機能について、4つの面（感性、コンテンツ、状況、タスク）からインタラクションの特徴をモデル化する枠組みを提案する。

感性モデルは、インタラクションの過程で、だれが、あるいは、だれのために、どのような判断基準で情報の取捨選択を行うかを定めるとともに、その人の感性的な状態（内部状態）を記述するために必要となるモデルである。

コンテンツモデルは、インタラクションの過程で、何を情報処理や検索の対象にしているのかを記述するために必要となるモデルである。利用者の感性の違いに関係なく、コンテンツ自身の性質から情報処理の仕方や取捨選択の基準が定まる場合もある。

状況モデルは、ある空間的な状況、時間的な状況で、情報処理の対象や、利用者とシステムとのインタラクションがどのような状況で進行しているかを記述する。同じ利用者によっても、利用者の置かれた状況が違えば、同一の操作などに対して異なったインタラクションの形態を提供する必要がある。物理的空間・情動的空間・ユビキタスの空間における状況を考えることができる。

タスクモデルは、利用者と情報機器あるいはインターネット上で運用される情報提供サービスとが、どのような文脈で、どのようなサービスを提供するかを定義するモデルである。

### 1.5.4 新世代グループウェアのためのアウェアネス研究の最新動向

グループウェア研究が一段落した1990年代になって、同期・対面環境ではあたりまえの存在感、実在感や臨場感が双方向通信（分散）環境では欠落していることに気づき、存在感、実在感や臨場感などのアウェアネスを補完する研究が勃興してきた。そこで報告者らは存在感、実在感、あるいは臨場感のアウェアネスを伝達する新世代グループウェア環境の研究開発を行った。本報告では、まず、同期・対面環境で自明の存在感、実在感、臨場感を、双方向通信（分散）環境でも伝達できる遠隔アウェアネス通信の研究の体系的サーベイを行う。次に、浅い感性を伝達する個別アウェアネス研究と浅い感性を統合するアウェアネス研究を紹介する。また、深い感性を伝達するアウェアネス研究について報告する。その結果、適切な「高品位」A-V環境を構築し、「高品位」なテキスト、手書き文字、

静止画像や動画画像を提供できれば、対面環境と同等あるいはそれ以上の存在感、実在感、および臨場感を伝達できる A-V 環境を構築できることを実証する。

### 1.5.5 音声情報処理の本格的アプリケーション構築に向けての展望と課題 — VoiceXML を中心として —

音声情報処理の技術に関しては、この数年間画期的な革新はなく、現在確立されている技術を用いて適応範囲を拡大していくフェーズであった。そのフェーズが終盤にさしかかった今、これまでの試みを集約した大規模で実用的なアプリケーションを構築することが、当面の目標となってきている。そのようなアプリケーションを構築するための共通の基盤として注目されているのが VoiceXML である。本報告では、VoiceXML を中心として、音声対話を利用した大規模アプリケーションの可能性に関する展望と課題について述べる。

### 1.5.6 Digital Libraries, Metadata, Interoperability

1990 年代のインターネット、WWW の発展は情報の発信と流通に大きな変化をもたらした。Digital Library (電子図書館、デジタル図書館) はこうした背景の下に注目され、研究開発が盛んに進められたトピックの 1 つである。Digital Library のための新しい技術に関する研究開発は、アメリカでの Digital Library Initiative をはじめとして、計算機科学、情報学分野を中心とする学際研究として進められてきた。一方、その間、図書館を中心とする現場においても、資料の電子化、サブジェクトゲートウェイ、増大する電子出版物への対応などさまざまな取り組みが進められてきている。また、ネットワーク上での情報発信と蓄積、流通を支えるには、流通する情報に関する情報をいろいろな視点から記述することが必要である。そのため、情報に関する情報、あるいは情報を表現した実体 (情報資源) に関する記述であるメタデータが注目され、さまざまな分野でメタデータの開発が進められてきた。

本報告では、はじめに、Digital Library に関する 1990 年代から進められたいろいろな研究開発活動を振り返って概観を述べる。ここでは計算機科学や情報学を中心とする研究開発指向の活動と、図書館を中心とする実際的な Digital Library 開発、それら以外のネットワーク上での情報資源の提供と共有の取り組みについて述べる。次に、メタデータに関する概要とネットワーク上での情報資源のためのメタデータとして広く認知されている Dublin Core Metadata Element Set に関して解説したあと、Digital Library における Interoperability について、メタデータの視点を中心として述べる。



### 1.5.7 デジタルコンテンツのトランスコーディングとインタフェースロボット

Web ページや MPEG ビデオなどのデジタルコンテンツをより高度に活用するための技術として、トランスコーディングがある。今後は、コンテンツを作成する技術と、それを多目的に再利用するためにコンテンツにメタ情報を付加する技術、そして、トランスコーディングのようなコンテンツ加工・変換技術に分かれて研究・開発が進むと考えられる。もちろん、それら技術の間の相互作用も考慮すべきであるが。また、ユーザの使用するクライアントデバイスもさまざまなものが開発されていくだろう。最も一般的なものは現在の携帯電話の延長にあるモバイルデバイスであると考えられるが、それ以外にも、カーナビやテレビ、PC、PDA の発展型も依然として存在し続けると思われる。それら、さまざまなクライアントデバイスに加えて、本報告では、インタフェースロボットを紹介する。このロボットは、主に、オンライン情報をユーザに会話的に提供する。本来、書き言葉で表現された文書コンテンツを話し言葉的に変換し、さらに、ユーザの反応に応じて、さらなる情報を追加・変換していく作業もトランスコーディングの一種と言える。

本報告では、報告者のグループで研究・開発を行っている、トランスコーディング技術およびインタフェースロボット技術について紹介する。

### 1.5.8 法的推論システムの研究動向と新しい展開

本報告では、まず、法的推論の研究に関する基礎知識を簡単に説明する。次に、法的推論を実現するための主な技術課題として、法的知識の記述方法や、法的議論のモデルや、法的推論システムの構築方法論についてその研究動向を含めて解説する。さらに、今後、法的推論技術の有望な応用分野として、調停・仲裁、紛争予防、法学教育が有望であることを述べ、それらの分野で実システムを開発するには、法律の知識だけでなく、マルチモーダルインタフェース技術や、交渉技術や、教育戦略についての技術開発が必要であることを述べる。

### 1.5.9 音楽情報処理研究の新しい潮流

本報告では、近年の音楽情報処理に関する新しい研究テーマを概観する。その特徴は、3 点にまとめられる。それらは、音楽理論を援用し音楽の意味を考慮した処理を実現すること、応用システムをインターネット/Web 上に展開すること、実現対象とするタスクが、作曲、編曲、演奏という大粒度のものから検索、模倣という中粒度のものに変化したことである。まず音楽を対象とする形式的な処理を計算機上で実現するのに有望な音楽理論を

紹介し、次に、新しい研究テーマを標榜する研究事例や、関連する国際会議を幾つか取り上げる。

#### 1.5.10 インタラクティブグラフィックスの動向

CG 技術は、映画での特撮やアニメーション制作などでの利用以外に、Web ブラウザ上でのコンテンツ再生、PC や家庭用ゲームでのインタラクティブな CG 映像の生成などにも応用されている。本報告では、インターネット上でのグラフィックス技術、リアルタイム CG を支えるハードウェア技術とその具体的な展開例である家庭用ビデオゲーム機の現状、さらに、インタラクティブグラフィックスの応用として VR や AR、インスタレーションアートなどの事例を紹介し、インタラクティブグラフィックスの将来像を考察する。

#### 1.5.11 検索履歴を用いた情報検索サポート

Google などの既存の情報検索システムでは、膨大な Web ページからあらかじめ生成しておいたインデックスに基づいて検索を行う。このため、異なるユーザが異なる意図を持って検索を行ったとしても、同一の検索語で検索した場合、必ず同一の検索結果が返される。しかし、ユーザが何を意図して検索を行ったかがわかれば、ユーザごとに異なる検索結果を返すことが可能となる。本報告では、このようなユーザごとに特化した情報検索をサポートする仕組みとして、従来提案されている手法を紹介するとともに、新しい手法として、同一ユーザが検索を行う際の連続する検索語の出現パターンから、ユーザの検索意図を発見する手法を紹介する。

#### 1.5.12 Web マイニングとその将来像

Web の拡大傾向は衰えず現在 20 億とも 40 億ともページがあると言われる。そうした膨大な Web におけるさまざまな情報から、目的に応じた知識やその素を取り出すプロセスが Web マイニングであり、現在、大学および企業から多方面のアプローチがある。

Web マイニングでは、Web 文書のテキスト群だけでなく、ハイパーリンク関係、アクセスログ、サービスログなど Web に関するあらゆるデータが対象となる。それに応じて、統計処理、自然言語処理、リンク解析といった多彩な技術が Web 向けに拡張されてきている。

本報告では、こうした Web マイニング技術について、技術の概要と、実践事例として Web の自動構造化の試み、さらに、今後の展開として次世代 Web (SemanticWeb、Web

サービス) とのかかわりについて述べる。

### 1.5.13 3次元視覚システムの技術動向と展望

コンピュータに人間の眼と同じような視覚機能を持たせようとするコンピュータビジョンの研究は、1960年代の人工知能の研究とともに始まった。そして、研究の初期に開発された2次元の図形を認識する2次元視覚技術は、大量生産型の工程の自動化を典型例として多くの実績を挙げてきた。しかし、現在、製品ニーズの多様化に対応する多品種少量型の生産には対応できなくなっている。この問題を解決できるキーテクノロジーとして、立体を立体的に認識する3次元視覚技術があり、製造業だけでなく、多くの分野でその早期実用化が求められている。本報告では、コンピュータビジョンの歴史から最近の技術動向と課題を述べ、産業技術総合研究所で体系的に開発している高機能3次元視覚システムVVVを紹介する。最後に、視覚、人工知能、ロボットなどを統合する新世代情報処理システムへの展望を述べる。

## 2. 米国政府が支援する情報技術研究開発の動向

本章では、まず 2.1 節で、米国連邦政府の情報技術研究開発を取り巻く環境として、連邦政府の研究開発予算に見られる動向を述べる。続いて、2.2 節では、米国の情報技術研究開発政策の柱であるネットワーキングおよび情報技術研究開発 (NITRD) イニシアチブの概要を説明する。次に、2.3 節では、NITRD において本ワーキンググループに関連する、ヒューマンコンピュータインタフェースおよび情報管理 (HCI&IM: Human-Computer Interaction and Information Management) の研究領域のテーマを含む研究プログラムの例として、NSF (National Science Foundation、全米科学財団) の情報技術研究 (ITR: Information Technology Research) プログラムとデジタルライブラリ・イニシアチブ (DLI: Digital Library Initiative) を簡単に紹介する。

NITRD (およびその前身) と DLI の歴史的経緯やそのインパクトの詳細については、本年度の AITEC 報告書の別冊[1]にも報告があるので、そちらも参照頂きたい。本報告では、できるだけそれらとの重複は避ける。

### 2.1 はじめに

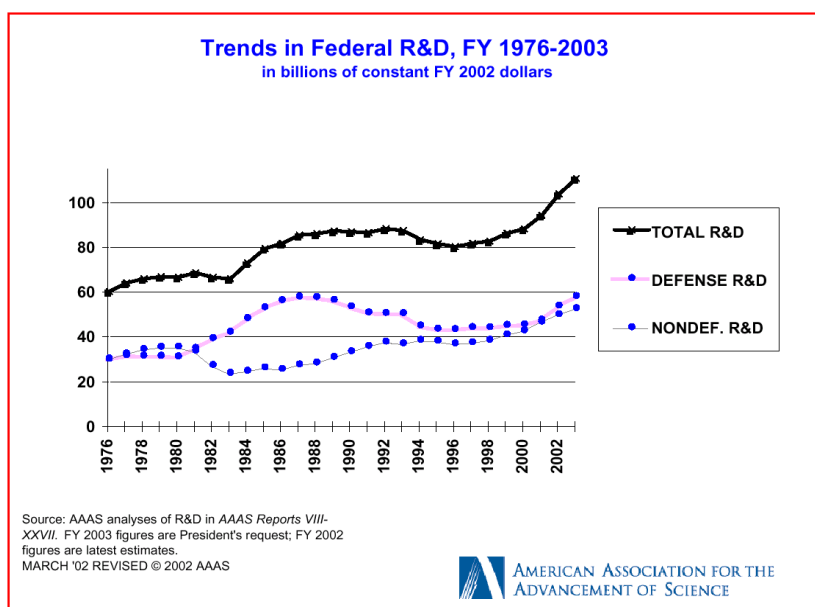
#### 2.1.1 米国連邦政府研究開発予算の動向

全米科学振興協会 (AAAS) がまとめたレポート[2]によれば、2002 年度の米国連邦政府の研究開発予算総額は、前年度比 12.7%増という大きな伸びを示し、1,032 億ドルとなった。初めて 1,000 億ドルの大台を超える大規模予算である。これは、同時多発テロに伴う緊急テロ対策支出金などの影響もあり、2001 年春の大統領予算教書による要求からかなり上積みされた額になっている。続いての 2003 年度は、2002 年 3 月 4 日に提出された大統領予算教書によれば、伸びの大きかった 2002 年度に対しても 8.6%増の 1,120 億ドルを要求しており、引き続き大きな伸びを示す予定である (図 2.1-1)。しかし、その内容は、防衛、健康、対テロ防衛を最優先として、それ以外のほとんどの国家ミッションに対する研究開発予算はむしろ減少する。

図 2.1-1 に見られるように、冷戦終結後、防衛予算は減少し、その後しばらくは横ばいが続いた結果、クリントン政権最後の 2001 年度には、防衛予算と非防衛予算がほぼ拮抗していた。

ところがブッシュ大統領最初の予算年度である 2002 年度は、選挙公約の「国防力の強化」の反映の上に、対テロ緊急支出で、防衛 R&D 支出は再び大きく増大した。近年の米国 R&D 政策は、ブッシュ政権となって、大きく「ミサイルと医薬(missiles and medicine)」の性格を強めている。2003 年度も再び、国家ミッションの中で防衛 R&D と健康 R&D に高い優先度を置く。DOD (Depart of Defense: 米国国防総省) と NIH (National Institute of Health: 国立衛生研究所) の 2 つが、ブッシュ大統領の最優先 R&D 投資先となっている (図 2.1-2)。

図 2.1-1 では、防衛 R&D も非防衛 R&D も伴に大きく伸びているが、非防衛 R&D 予算の大きな伸びは、NIH の伸びによるものである。NIH 予算を、1998 年度のレベルから 2003 年度までに倍増するという計画があり、これを達成するためにも最終年度の 2003 年度に大きな伸びが必要となっている。



R&D 予算は、2001、2002、2003 年度と大きな伸びが続いている。

図 2.1-1 米国連邦政府 R&D 予算の動向

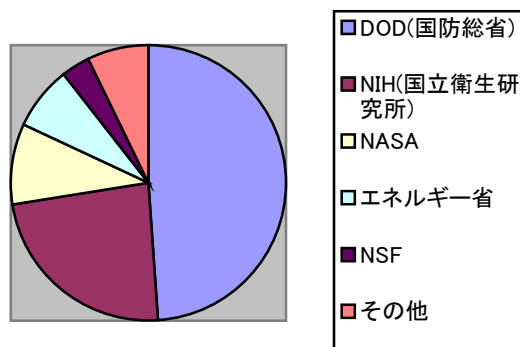


図 2.1-2 2003 年度予算教書における R&D 予算の主要エージェンシー比率

## 2 米国政府が支援する情報技術研究開発の動向

図 2.1-3 に示すように、非防衛 R&D 予算の中で、NIH の R&D は着実に伸びてきているが、それ以外の非防衛 R&D は伸び悩んでいる。2003 年度要求では、NIH を除いた非防衛 R&D 全体としては、0.2%減少し、NIH とそれ以外がほぼ同一レベルとなる。

結局、連邦政府 R&D 予算のおおまかな構成は、全体の約半分が防衛 R&D であり、残りのうちのまた半分が健康 R&D、その残りが一般の R&D 予算ということになる(図 2.1-4)。

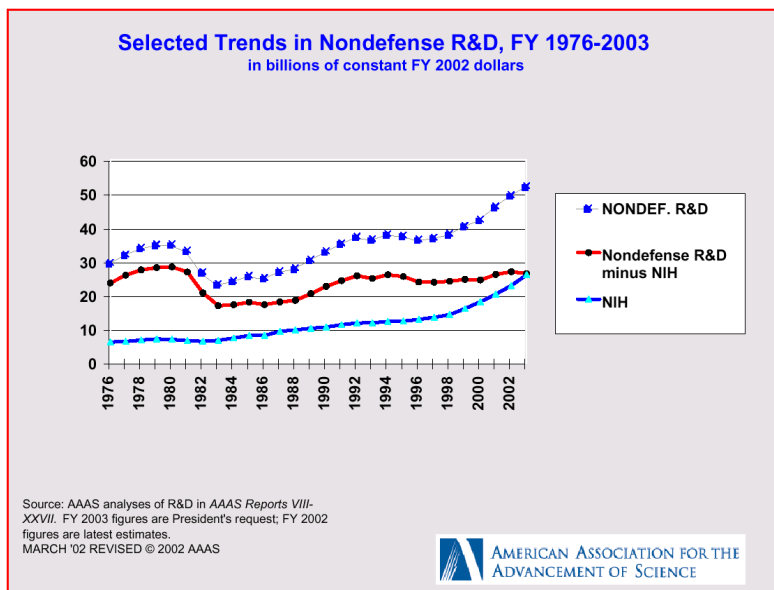


図 2.1-3 非防衛 R&D 予算の推移動向

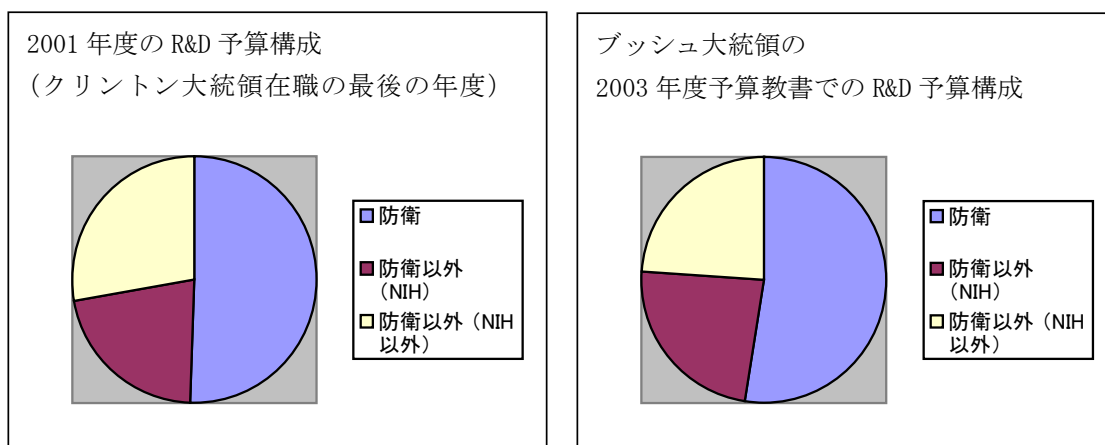


図 2.1-4 米国連邦政府 R&D 予算の構成比率 (防衛 R&D、非防衛 R&D)

表 2.1-1 省庁間連携研究開発イニシアチブの予算

|                                     | 2001 年度      | 2002 年度      | 2003 年度      | 2002-03 年度の推移 |       |
|-------------------------------------|--------------|--------------|--------------|---------------|-------|
|                                     | 実績<br>(百万ドル) | 推定<br>(百万ドル) | 要求<br>(百万ドル) | 金額<br>(百万ドル)  | 増率    |
| ナノテクノロジー<br>(Nanotechnology)        | 466          | 604          | 710          | 106           | 17.5% |
| ネットワークングおよび<br>情報技術研究開発 (NITRD)     | 1,768        | 1,844        | 1,890        | 46            | 2.5%  |
| 気候変動研究<br>(Climate Change Research) | 0            | 0            | 40           | 40            | N/A   |
| 米国地球変動研究<br>(USGCRP)                | 1,728        | 1,670        | 1,714        | 44            | 2.6%  |

## 2.1.2 研究開発投資効率の改善

上述のような、非防衛・非 NIH については全体としてフラットな連邦政府 R&D 予算の中で、省庁間にまたがるイニシアチブは重視されており、その R&D 予算は順調に増加している（表 2.1-1）。

これは、ブッシュ政権の、連邦政府 R&D 投資の効率改善策の一環でもある。この 1 年にわたり、OSTP (Office of Science and Technology Policy、科学技術政策局) と OMB (Office of Management and Budget、行政管理予算局) は、連邦政府機関や科学コミュニティとともに、連邦政府 R&D の最優先領域を識別する作業を行ってきた。その結果は、例えば情報技術のような、国家にとって重要な領域、ナノテクノロジーのように、多くの分野にわたって新たなブレークスルーを提供することが期待される新興の領域、その他には、(表 2.1-1 にはないが) 対テロリズム R&D のように、新たに認識されたニーズを扱うもの、などである。

これらの省庁間連携イニシアチブのうち、ナノテクノロジーと NITRD は、よく調整された戦略的な計画と実行を保証するために、NSF の支援の下に、個別の調整組織 (coordination office) を持っている。両イニシアチブは、2003 年に向けた詳細な研究提案を記述した統一的な計画を、それぞれ作ることになっている。対テロリズム R&D と気候変動 R&D については、新しい組織の設立と省庁間調整の強化を行いつつある。政府は、

## 2 米国政府が支援する情報技術研究開発の動向

よりよい省庁間連携によって将来役に立つ重要なニーズが得られるような領域選択の検討を、今後も続ける予定である。

また、2004年度予算割り当てでの適用を目標に、R&D投資判定基準を開発中である。多くのR&D、特に基礎研究では、結果が数年後あるいは数十年後にしか明白にならないかもしれないので、R&Dの効率の評価には特別の配慮を必要とするが、それにもかかわらず、政府は、プロジェクトのマイルストーン達成の判定基準の適用も含めて、R&D間での投資対象決定のための基盤の改善を試みる。

この判定基準はまだ開発中であるが、すでにいくつかの基本原則については、明らかにされ、一部適用されている。その結果、2003年度予算において、いくつかの研究プログラムの省庁間移動なども行われている。基本原則としては、例えば以下のような項目が挙げられている。

- ・ 連邦政府R&Dの優先順位は、大統領の識別する優先順位と一貫性があるべきである。
- ・ 連邦政府のR&Dプログラムは、国家の目標を達成するために連邦政府が行う必要がある活動に重点を置くべきである。公的資金が民間投資を排除することを避けるため、連邦政府のR&D資金援助は、主に民間セクターがR&Dの利益を得ることのできない分野に重点を置くべきである。
- ・ プログラムおよび提案は、研究の明確な目標と、適切な場合には、計画された終了点または終了の目やすを備えた十分な計画を持つべきである。
- ・ 研究プロセスの質と公的投資の効率を最大化するために、研究プログラムは、適切なならば、競争的で価値に基づく（merit-based）選考プロセスを用いるべきである。例外は十分に正当な理由がなければならない。
- ・ R&Dの指揮、選考、および管理において優れていると判断される省庁機関やプログラムは、識別され支援されるべきである。
- ・ あまりうまくいっていないプログラムは、改善を達成するために、うまくいっているモデルに従うべきである。あるいは、縮小されるか、もっと効率的に管理され得る省庁機関に移されるべきである。
- ・ 新しい重点R&Dのための資源は、ミッションを完了したプログラム、あるいは、冗長または時代遅れとなったプログラムに対する資金を減額、あるいは削除することによって、増額されるべきである。



## 2.2 ネットワーキングおよび情報技術研究開発計画

ブッシュ大統領は、2003 年度予算要求において、「1995 年以來の 80%の経済生産性増加の 3 分の 2 は、情報技術のおかげに帰することができる」と述べるなど、情報技術研究開発は、ブッシュ政権においても重要優先分野である[13]。

情報技術は、バイオやナノテクノロジーを初め、ほかのほとんどの分野の研究開発においても基盤技術として横串的に重要な役割を果たすが、情報技術そのものの発展を主眼とした研究開発として、前述の省庁間研究開発イニシアチブの 1 つであるネットワーキングおよび情報技術研究開発 (NITRD) イニシアチブが、米国の情報技術研究開発政策の柱となっている。

上院議員時代にゴアの提出した法案「National High-Performance Computing Act」を基に始まり、クリントン・ゴア政権において、HPCC や CIC、ITR&D といった名前で強力に推進されてきた情報技術研究開発イニシアチブが、ブッシュ政権下では、クリントン・ゴア政権での IT R&D の基本的な枠組みを受け継いで、NITRD (Network and Information Technology R&D) と呼ばれ継続している。

### 2.2.1 2002 年度の概要 —BlueBook 2002

クリントン・ゴア政権下では、議会に提出する予算教書の補助資料「BlueBook」において、IT R&D の有効性・重要性を示すため、毎年、多くの具体的研究開発事例を紹介してきた。一方、ブッシュ政権もそれを継承して、2001 年 8 月に NITRD に関する BlueBook 2002[3]を公表したが、この政権下での具体的な成果はこれからという理由からか研究事例の詳細はなく、一般的記述として、長期的研究開発主体の 10 個の研究チャレンジと、現実的な課題である 7 つの国家的グランドチャレンジアプリケーションを示している。

長期的研究開発主体の研究チャレンジは次の通り。

- (1) 次世代コンピューティングとデータ記憶技術。
- (2) シリコン CMOS における障壁の克服。
- (3) 21 世紀のための多目的、安全、拡張可能なネットワーク。
- (4) 科学と工学における米国の強みを維持するための先進的 IT。
- (5) 重要なシステムにおける信頼性と安全な運用の確保。
- (6) 現実世界のためのソフトウェア作成。
- (7) 人間の能力と万人の進歩に対する支援。
- (8) 知識世界の管理と実現。
- (9) 世界レベルの IT 要員の教育と育成の支援。

## 2 米国政府が支援する情報技術研究開発の動向

- (10) 利益を最大化するための IT の効果についての理解。

国家的グランドチャレンジアプリケーションは次の通り。

- (1) 次世代の国防と国家安全保障システム。
- (2) 全国民のための改良された健康管理システム。
- (3) 科学的に正確な人体の 3 次元機能モデルの制作。
- (4) 大規模環境モデリングとモニタリングのための IT ツール。
- (5) 生涯学習のための世界最良のインフラストラクチャの創出。
- (6) 危機管理のための統合 IT システム。
- (7) 先進的航空管理のための技術とシステム。

### 2.2.2 プログラムコンポーネント領域 (Program Component Area: PCA)

NITRD 計画の研究領域は、IT R&D を継承して、表 2.2-1 と表 2.2-2 に示す 7 つのプログラムコンポーネント領域 (PCA) からなっている。

なお表 2.2-1 で、NGI は、LSN で実現される予定の「テラビット/秒」ネットワーキングの目標を除くすべての目標を 2002 年度中に達成して、成功裏に終了した。関連する先進ネットワーキング研究は、LSN の下で継続されている[4]。

本報告書で扱っている人間主体の知的情報技術に関する研究開発項目は、ほぼ、HCI&IM (ヒューマンコンピュータインタフェースおよび情報管理) の PCA に対応している。

表 2.2-1 NITRD のプログラムコンポーネント領域 (PCA) (その 1)

|                                    | PCA<br>(プログラムコンポーネント領域)                  | 概要                                                                                                                                                                                                                                                                                                                     |
|------------------------------------|------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| HECC (ハイエンドコンピューティングおよびコンピューテーション) |                                          |                                                                                                                                                                                                                                                                                                                        |
| 1                                  | HEC I&A<br>(ハイエンドコンピューティング基盤およびアプリケーション) | ハイエンドコンピューティングのアプリケーション開発とコンピューティングインフラストラクチャの整備。                                                                                                                                                                                                                                                                      |
| 2                                  | HEC R&D<br>(ハイエンドコンピューティング研究開発)          | ハイエンドコンピューティングのハードウェア技術とシステムアーキテクチャの研究。                                                                                                                                                                                                                                                                                |
| 3                                  | LSN<br>(大規模ネットワーク技術)                     | <p>基本ネットワーク技術に関する長期的研究を実施。NGI および SII の 2 つのイニシアチブを含む。なお、下記の NGI は、目標をほぼ達成して、2002 年度中に終了した。</p> <p>NGI (次世代インターネット)：ネットワークのサービス品質、信頼性、堅牢性、安全性を向上させる次世代技術の研究開発および高速ネットワーク上での革新的アプリケーションの開発と実証。</p> <p>SII (拡張可能な情報基盤)：今後出現するであろう大規模なユーザやアプリケーションに対応してインターネットを円滑に成長させ、高度にネットワーク化されたシステムや「いつでもどこでも」接続できる環境を実現するための研究。</p> |
| 4                                  | HCI&IM<br>(ヒューマンコンピュータインタフェースおよび情報管理)    | <p>ヒューマンコンピュータインタフェースの向上と、コンピューティング装置や情報リソースを管理し活用する能力を高める技術の研究開発。</p> <p>研究テーマ：ロボット工学、遠隔/自律エージェント、コラボレーション技術、可視化技術、バーチャルリアリティ、知識データベース、デジタルライブラリ、情報エージェント、言語認識、インテリジェントシステム、多言語翻訳など。</p>                                                                                                                              |

## 2 米国政府が支援する情報技術研究開発の動向

表 2.2-2 NITRD のプログラムコンポーネント領域 (PCA) (その2)

|   | PCA<br>(プログラムコンポーネント領域)    | 概要                                                                                                                      |
|---|----------------------------|-------------------------------------------------------------------------------------------------------------------------|
| 5 | HCSS<br>(高信頼ソフトウェアおよびシステム) | コンピュータシステムに求められる信頼性、安全性、セキュリティ、可用性を実現するための技術の研究開発。<br>研究テーマ: ネットワークとデータのセキュリティ、暗号化、情報の生存性、システムの耐久性など。                   |
| 6 | SDP<br>(ソフトウェアの設計および生産性)   | ソフトウェア需要に対する開発能力の不足や、ソフトウェアの設計、テストおよび維持管理の困難化といった緊急の課題に対処するため、有用で安全性/信頼性の高いソフトウェアを効率良く開発するための理論的基礎研究、技術やツールに関する研究開発を実施。 |
| 7 | SEW<br>(IT の社会・経済への影響)     | IT が社会システムに与える影響、IT 技術者の需要増加に対処するための技術者育成、デジタルデバイド等に関する研究。                                                              |

### 2.2.3 NIITRD 計画の計画・執行組織

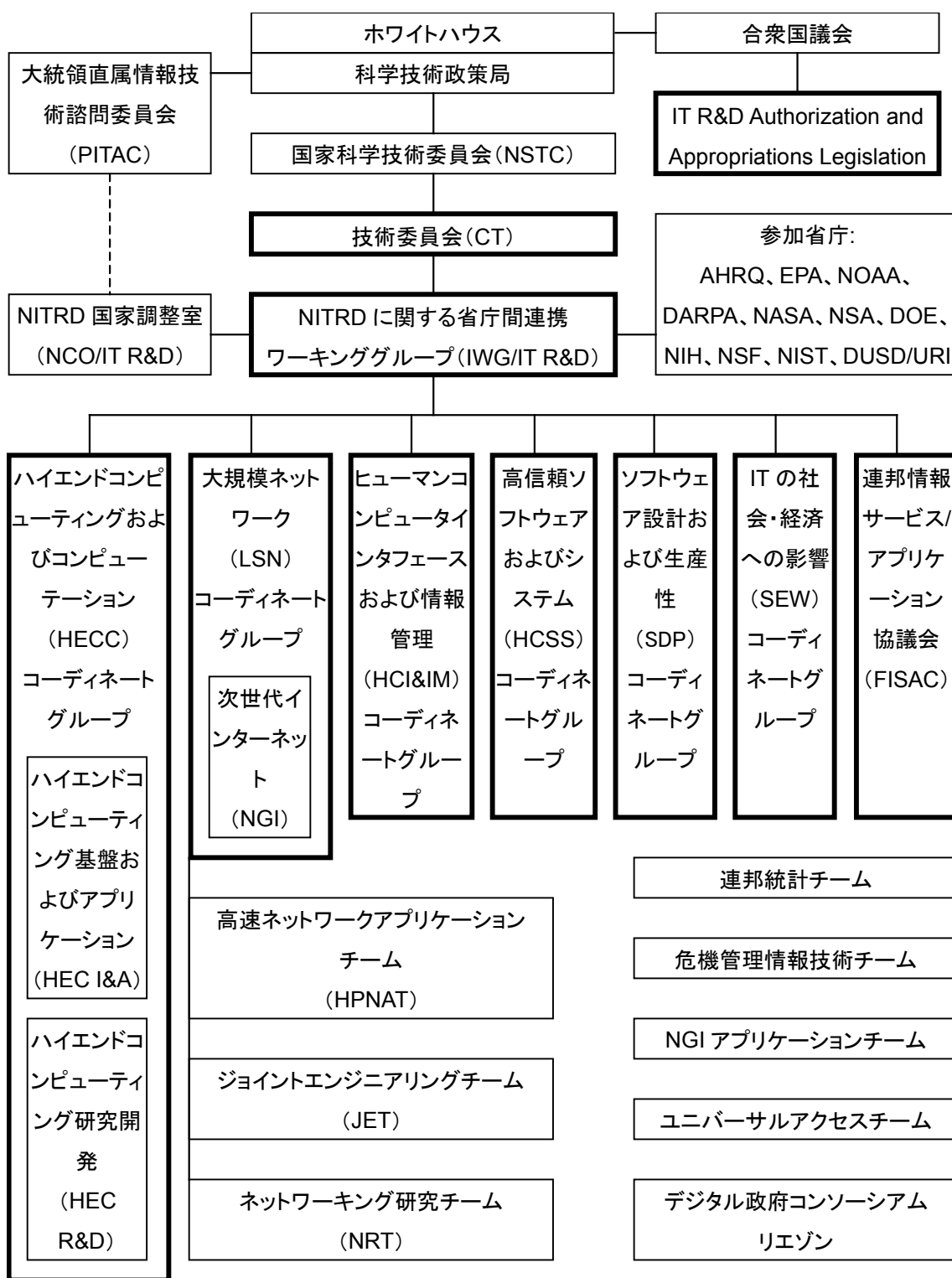


図 2.2-1 NIITRD の計画・執行組織[14]

## 2 米国政府が支援する情報技術研究開発の動向

NITRD の計画および執行に関係する組織について以下に説明する（図 2.2-1）。組織名  
称の中では、まだ NITRD ではなく IT R&D の呼称が使われている。

### (1) 科学技術政策局（Office of Science and Technology Policy: OSTP）

ホワイトハウス内に置かれ、米国政府の R&D 投資に関して大統領への助言を行うとともに、大統領直轄の国家科学技術委員会の活動を監督する。大統領科学技術担当補佐官が OSTP の長官を兼ねる。

### (2) 国家科学技術委員会（National Science and Technology Council: NSTC）

大統領が議長を務め、委員は各省の長官レベルで構成される。NSTC は連邦政府の科学技術投資のための国家目標の設定およびエージェンシー間にわたる R&D 戦略の確立を目的としている。

### (3) 技術委員会（Committee on Technology: CT）

技術委員会は、NSTC と協力して R&D プログラムの優先順位の設定支援と、米国政府の NITRD ポートフォリオと関係機関のミッションおよび能力のバランス調整を行い、進捗を管理する。

ここは、OSTP 長官が議長を務め、メンバーは NITRD に資金援助を行っているエージェンシーの上級官吏で構成される。

### (4) NITRD に関する省庁間連携ワーキンググループ（Interagency Working Group on IT R&D: IWG/IT R&D）

IWG/IT R&D は、実施エージェンシーの政策、計画、予算に関する助言を行う。

このワーキンググループは計画、予算策定、実施およびレビューなどを含む複数のエージェンシーが関係する IT 研究活動の調整を行い、連邦議会、OMB（行政管理予算局）、OSTP、学会、および産業界との間の調整を行うほか、PITAC への技術的支援を行う。

### (5) コーディネートグループ（Coordinating Groups）

それぞれの PCA に対応したコーディネートグループがあり、IWG/IT R&D の管轄下に置かれている。ただし、HEC I&A（ハイエンドコンピューティング基盤）と HEC R&D（ハイエンドコンピューティング研究開発）は相互に密接に関係するため、ともに HECC コーディネートグループが担当する。この 6 つのコーディネートグループは各々の専門領域における R&D の目標と実施に関する調整を行う。LSN コーディネートグループ（LSNCG）はさらに 3 つのサブグループ（チームと呼ばれる）に分けられる。

### (6) 連邦政府情報サービス/アプリケーション協議会（Federal Information Services and Applications Council: FISAC）

FISAC は NITRD 計画の成果として得られた技術をエージェンシーの業務や米国政府のシステムに適用するための支援を行う。

### (7) NITRD 国家調整室（National Coordination Office for Information Technology R&D: NCO/IT R&D）

NCO/IT R&D は関連エージェンシーの計画、予算策定、評価資材の準備、R&D 情報および文書の集中管理など、NITRD 計画全体に対する支援を行う。

#### 2.2.4 2003 年度の重点課題

ブッシュ大統領の 2003 年度予算教書では、NITRD に 19 億ドルを要求している。これは前年度予算に対して 3%の増加となる。2003 年度は、以下の項目に重点を置く予定である[5]。

- ・ ネットワークの信頼性（セキュリティ、信頼性、およびプライバシー）
- ・ 高信頼度（high assurance）ソフトウェアおよびシステム
- ・ マイクロおよび埋め込みセンサ技術
- ・ ハイエンドコンピューティングプラットフォームのコスト、サイズ、パワー要件を引き下げるための革新的なアーキテクチャ
- ・ 情報技術の社会的および経済的インパクト

達成目標項目としては、以下のものを含む。

- ・ end-to-end の光ファイバーネットワークの開発。これにより、研究および商業アプリケーションのために、帯域幅とネットワークセキュリティの大幅な向上を提供する。
- ・ クラスタあるいは「グリッド」コンピューティングを可能にする新技術。これにより、初めて全国の科学研究者のために、ハイパフォーマンスコンピューティングへのアクセスを提供する。
- ・ 侵入検出や危険性と脆弱性の分析のようなネットワークセキュリティ保護技術。
- ・ 大規模な情報リポジトリ（情報保管庫）の利用、管理、アーカイブ（大量の文書保管）技術。あるいは、デジタルライブラリ（電子図書館）。

本ワーキンググループのテーマに関連する「ヒューマンコンピュータインタフェースと情報管理」の研究領域にあたるものとしては、最後の「デジタルライブラリ（電子図書館）」が、達成目標項目に挙げられている。これは、従来の図書館の電子化の枠を超えて、インターネット上の膨大なデジタル情報資源の利用と管理の全般にわたる技術であり、今後のインターネットを中核に据えた情報社会の基盤技術である。

## 2.3 ヒューマンコンピュータインタフェースおよび情報管理領域

NITRD のヒューマンコンピュータインタフェースおよび情報管理 (HCI&IM: Human-Computer Interaction and Information Management) プログラムコンポーネント領域 (PCA) では、ヒューマンコンピュータインタフェース技術を発展させ、コンピューティング装置や情報資源を管理し活用する能力を高めるための高度な技術の研究開発を行う。HCI & IM では、幅広い研究テーマを扱うことにより、生体臨床医学、商業、危機管理、教育、法執行、図書館学、製造、国防、学問、気候分析など多くの分野で新たな IT の可能性を生み出している。

NITRD の省庁別・研究分野 (PCA) 別の予算配分を見ると、HCI&IM の研究開発予算は、NSF が最も多く (37.5%)、次が NIH で (26.7%)、この 2 つの機関で HCI&IM 全体の予算の 6 割を超える (計 64.2%) (表 2.3-1)。

2001 年度までの BlueBook では、各プログラムコンポーネント領域の具体的な研究開発事例のトピックが紹介されていたが、ブッシュ政権下となって初めての 2002 年度の BlueBook では、具体的事例の紹介はなく、一般的解説のみとなっている。

以下の節では、HCI&IM に関連する代表的なプログラムとして、NSF の ITR (Information Technology Research) と、NSF がリーダーシップを取る省庁間連携プログラムであるデジタルライブラリ・イニシアチブ (Digital Library Initiative) について紹介する。

他の省庁では、HCI&IM の研究もその省庁のミッションと何らかの関係のある形で行われることが多い。例えば NIH 関連の HCI&IM 関連研究としては、以下のようなプロジェクトが、過去に紹介されている (BlueBook 2000)。

- ・ 医療手順をシミュレートするためのバーチャルリアリティ技術 (NIH/NCRR)。
- ・ 生体臨床医学共同実験室 (NCRR)。
- ・ 統一医学用語システム (NLM)。
- ・ Medline Plus (NLM)。
- ・ Visible Human プロジェクト (NLM)。

ここで NCRR、NLM は、NIH 傘下の研究機関である。

- ・ NCRR: National Center for Research Resources、国立研究資源センター。
- ・ NLM: National Library of Medicine、国立医学図書館。



表 2.3-1 2002 年度 NITRD の省庁別・研究分野別予算要求額

| 省庁             | HEC<br>I&A | HEC<br>R&D | HCI&IM<br>(HCI&IM 内<br>比率) |         | LSN   | SDP   | HCSS  | SEW  | 合計      | 総額<br>比率 |
|----------------|------------|------------|----------------------------|---------|-------|-------|-------|------|---------|----------|
| NSF            | 249.7      | 65.1       | 104.8                      | (37.5%) | 98    | 39.7  | 46.1  | 39.1 | 642.5   | 32.6%    |
| DARPA          | 55.5       | 42.7       | 38.2                       | (13.7%) | 49.2  | 44.6  | 32.9  |      | 263.1   | 13.4%    |
| NIH            | 55.1       | 13.7       | 74.6                       | (26.7%) | 81.1  | 6     | 10.1  | 11.4 | 252     | 12.8%    |
| NASA           | 36.1       | 26         | 27.8                       | (9.94%) | 14.4  | 22.4  | 47.1  | 7    | 180.8   | 9.2%     |
| DOE/<br>OS     | 98.3       | 31.5       | 16.4                       | (5.9%)  | 25.9  |       |       | 4    | 176.1   | 8.9%     |
| NSA            | 33.6       |            |                            |         | 1.9   |       | 46.6  |      | 82.1    | 4.2%     |
| NIST           | 3.5        | 6.2        | 3.2                        | (1.1%)  | 2     | 7.5   |       |      | 22.4    | 1.1%     |
| NOAA           | 13.3       | 1.8        | 0.5                        | (0.2%)  | 2.7   | 1.5   |       |      | 19.8    | 1.0%     |
| AHRQ           | 9.2        | 6.7        |                            |         |       |       |       |      | 15.9    |          |
| ODUSD<br>(S&T) |            | 2          | 2                          | (0.7%)  | 4.2   | 1     | 1     |      | 10.2    | 0.5%     |
| EPA            | 1.8        | 1.8        |                            |         |       |       |       |      |         | 0.0%     |
| 小計             | 513.3      | 216.4      | 279.7                      | (100%)  | 287.3 | 115.7 | 192.8 | 61.5 | 1,666.7 | 84.6%    |
|                |            |            |                            |         |       |       |       |      |         |          |
| DOE/<br>NNSA   | 133.8      | 37         |                            |         | 35.5  | 41.1  |       | 56.5 | 303.9   | 15.4%    |
| 合計             | 647.1      | 253.4      | 279.7                      | (100%)  | 322.8 | 156.8 | 192.8 | 118  | 1,970.6 | 100.0%   |
|                |            |            |                            |         |       |       |       |      |         |          |
| 比率             | 32.8%      | 12.9%      | 14.2%                      |         | 16.4% | 8.0%  | 9.8%  | 6.0% | 100.0%  |          |

(単位: 百万ドル) [2]

## 2.3.1 ITR (Information Technology Research) プログラム

表 2.3-2 NSF Funding by Priority Area

| NSF の 2003 年度重点領域<br>(Priority Area)                        | 2002 年度計画 | 2003 年度要求 | 増加率    |
|-------------------------------------------------------------|-----------|-----------|--------|
| 環境における生物複雑性<br>(Biocomplexity in the Environment)           | 58.10     | 79.20     | 36.3%  |
| 情報技術研究<br>(ITR: Information Technology Research)            | 277.52    | 285.83    | 3.0%   |
| ナノスケール科学技術<br>(Nanoscale Science and Engineering)           | 198.71    | 221.25    | 11.3%  |
| 21 世紀労働力のための学習<br>(Learning for the 21st Century Workforce) | 144.82    | 184.69    | 27.5%  |
| 数理科学<br>(Mathematical Sciences)                             | 30.00     | 60.09     | 100.3% |
| 社会・行動・経済学的科学<br>(Social, Behavioral and Economic Sciences)  | 0.00      | 10.00     | NA     |
| 合計                                                          | 709.15    | 841.06    | 18.6%  |

(単位: 百万ドル) [9]

NSF の ITR プログラム[6]は、1999 年 2 月の PITAC レポートによる勧告[7]と、それを受けた「21 世紀情報技術 (IT<sup>2</sup>)」イニシアチブ[8]の直接の流れを汲むプログラムである。

PITAC は、1999 年のレポートにおいて、米国は情報技術研究において「深刻に投資が不足している」と断言した。今日の情報経済の繁栄を導いたようなハイリスクの長期研究投資を、収益確立前に企業が引き受けることはないであろうから「長期のハイリスク研究」に対する新たなファンディング無しには、IT における米国の優位が脅かされるだろうと委員会は指摘した。それを受けて政府は、2000 年に「21 世紀に向けての情報技術」IT<sup>2</sup> 計画として、従来からの研究とは別枠で、基礎研究に 366 百万ドルの研究開発投資を追加した。その目玉として開始されたのが、NSF のこの ITR プログラムである。

現在、ITR は、NSF で最も大きな重点プログラムとなっている (表 2.3-2)。

ITR プログラムは、ハイリスクの長期的基礎研究への助成を目的としており、NITRD のすべての PCA (Program Component Area) の範囲をカバーしている。ITR のメインゴールは、国家の IT 知識ベースの拡大と、IT 労働力の拡大である。

ITR は初年度 (2000 年度) は、IT の基礎研究と教育に焦点を置いて、2000 年 9 月に

210 件の助成プロジェクトが発表された。2 年度目（2001 年度）はそれに加えて、IT を科学および工学のチャレンジ全般に適用しようとする研究と教育に焦点を当て、309 件の新規プロジェクトが採択されている。

現在公募・採択作業が進行中の 2002 年度プロジェクトでは、IT と他の分野との間のインタフェース領域の発展に着目し、学際的領域における研究と教育へ焦点が当てられる。

このプログラムでは、IT を解明、拡張、そして探求するような、研究・教育における、革新的プロジェクトを支援する。

表 2.3-3 ITR の 2001 年度の採択プロジェクト

| プロジェクト規模 | 件数    | 総額                | 年額       |
|----------|-------|-------------------|----------|
| 大規模      | 8 件   | 5 百万ドル超～15 百万ドル以下 | 3 百万ドル未満 |
| 中規模      | 113 件 | 500 千ドル超～5 百万ドル以下 | 1 百万ドル未満 |
| 小規模      | 188 件 | 500 千ドル以下         | ---      |

### 2.3.1.1 2001 年度 ITR のトピック

#### 2.3.1.1.1 概要

NSF は 2001 年 9 月 25 日に、第 2 年度目にあたる 2001 年度の採択プログラムを発表した。2000 件を超えるプロジェクト提案の中から、309 件が採択されており、規模の内訳は表 2.3-3 の通りである[10]。

NSF 長官の Rita Colwell は、PITAC 会合での ITR 採択プロジェクト報告に際して次のように述べている。

NSF は、これらの大胆な ITR プロジェクトの推進役であることを誇りに思う。長期のハイリスク研究を通して、我々は、全体として国民の利益になるような、広範な積極的結果を期待している。我々の目標は、科学者や技術者が最終的に産業に応用されるような種類の発見を助けるようなソフトウェアと IT サービスの開発を支援することである。

### 2.3.1.1.2 2001 年度の ITR 研究推進領域

ITR プログラムの主要な目標は、国家の IT 知識ベースの拡大と、IT 労働力の強化である。2001 年度の重点領域としては、以下の 5 つを挙げている。

- (1) システムの設計と実装（ヒューマンコンピュータインタフェースを含む）(SY)。
- (2) IT と相互に影響し合う人々および社会グループ  
（経済的および労働人口的意味を含む）(PE)。
- (3) 情報管理（コンテンツ/データの分析と情報処理学を含む）(IM)。
- (4) 科学および工学アプリケーション  
（シミュレーションおよび先端的計算を含む）(AP)。
- (5) スケーラブル情報基盤  
（セキュリティ、テザーフリーコンピューティング<sup>1</sup>、遠隔没入型仮想現実を含む）(SI)。

各分類の先頭に示してある英文名頭文字から取った 2 文字 (SY、PE、IM、AP、および SI) は、各技術領域を識別するために使われ、2001 年度のプロポーザルのタイトルは、この 2 文字を付けて、"ITR/XX:"の形式で始まらなければならない。技術領域にまたがる場合は、例えば、次のようなタイトル形式となる。

"ITR/SY+PE:"

### 2.3.1.1.3 2001 年度の ITR プロジェクトのトピック

NSF では、2001 年度の ITR 採択プロジェクトの発表にあたり、いくつかの事例[10]を紹介している。

- (1) カリフォルニア大学バークレー校の ITR プロジェクトは、エネルギー、災害対策、および教育等に関連した複雑な問題を解決することを目的とした、「社会的規模」の情報システムを開発する。
- (2) カンザス大学では、なぜ海水のレベルが過去 100 年間上昇し続けてきているのかを突き止める助けとするために、極地域に、アイスシート（氷板）、大洋、および大気間の相互作用に関するリアルタイムデータの収集と解析のための、レーダーセンサーを設置する。
- (3) カーネギーメロン大学、ライス大学、およびオールドドミニオン大学のコンピュータ科学やその他の分野の研究者たちは、航空力学、エネルギー、環境、地球物理学、医学、およびその他の応用に関連した物理プロセスからのデータをたえず収集活用するために、オンラインシミュレーションのためのソフトウェアを開発する。

---

<sup>1</sup> "tether free" computing: 無線など、線につながれていない環境でのコンピューティング。

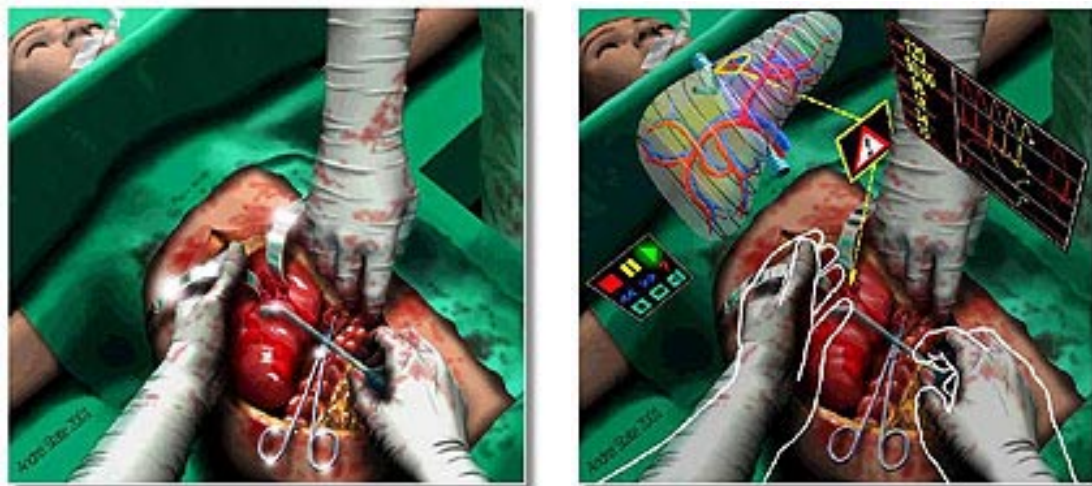


図 2.3-1 3次元グラフィックスによる手術の再現

(4) Survivors of the Shoah Visual History Foundation は、ホロコースト（ナチスのユダヤ人大虐殺）の生存者と目撃者の 116,000 時間を超えるインタビューのデジタル化ビデオを保管している非営利組織である。ITR 助成が、このマルチメディアコンテンツの目録を作るための音声認識ソフトウェアに当てられる。その多言語性の局面は特に研究的チャレンジとなる。言語間にまたがる新たな検索能力は、「データに関するデータ」として定義されるメタデータに対して広範な意味を持つことになる。

(5) ニューヨークのクラークソン大学は、個体物理学研究に対して先端 IT 技術を適用する。それには、シリコンチップに取って代わるために原子レベルのプロセスを利用する超高速量子コンピュータ開発を手助けする若手科学者の教育も含まれる。

(6) オregon 州立大は、海洋と大気に関する大量のデータを処理するための研究活動を行う。6 つの大学、1 つの国研、および民間企業からの、コンピュータ科学者、海洋学者、および土木技術者を含むチームが、天気予報、海洋事象、地下水、および地球物理学を支援するため、モジュラーソフトウェア（規格部品化されたソフトウェア）を開発する。

(7) ブラウン大学とノースカロライナ大学チャペルヒル校によって進められているプロジェクトでは、外科医をトレーニングするために、「テレイマージョン（遠隔没入）」技術が開発される。実物大の 3 次元グラフィックスで高忠実度の手術を再現し、訓練生が学習のために作業を中断したり、前に戻って作業を反復したりすることを可能にする。3 次元グラフィックスによる手術の再現の様子を図 2.3-1 に示す。

## 2 米国政府が支援する情報技術研究開発の動向

ここで述べた 2001 年度採択 ITR プロジェクト例の情報を表 2.3-4 と表 2.3-5 に示す。

表 2.3-4 2001 年度採択 ITR プロジェクト（本節で紹介したもののみ）（その 1）

|     | Award Number<br>Title                                                                                                                                                                                                                                                        | 開始<br>終了             | 助成タイプ<br>総額（予想）                                       | 研究機関                                   |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|-------------------------------------------------------|----------------------------------------|
| (1) | 0122599 ITR/SI:<br>Societal Scale Information Systems: Technologies,<br>Design and Applications<br>社会的規模の情報システム: 技術、設計、および<br>アプリケーション                                                                                                                                       | 10-01-01<br>10-31-06 | Continuing<br>Grant<br>\$7,127,606<br>LARGE           | カリフォルニア<br>大学バー<br>クレイ校                |
| (2) | 122520 ITR/SI+AP:<br>A Mobile Sensor Web for Polar Ice Sheet<br>Measurements<br>極地域のアイスシート（氷板）計測のための<br>モバイルセンサー網                                                                                                                                                            | 09-30-01<br>09-29-06 | Continuing<br>Grant<br>\$5,490,000<br>LARGE           | カンザス<br>大学                             |
| (3) | 121667 ITR/AP COLLABORATIVE RESEARCH:<br>Real Time Optimization for Data Assimilation and<br>Control of Large Scale Dynamic Simulations<br>大規模な動的シミュレーションにおけるデータ処理<br>同化とコントロールのための実時間最適化<br>ほかの参加大学:<br>ライス大学 (NSF#121360) \$550,000、<br>オールドドミニオン大学 (NSF#121207) \$805,000 | 09-30-01<br>08-30-06 | Standard<br>Grant<br>\$1,145,000<br>MEDIUM<br>(GROUP) | カーネ<br>ギーメロ<br>ン大学                     |
| (4) | 122466 ITR Large:<br>MALACH: Multilingual Access to Large spoken<br>ArChives<br>MALACH: 大規模な話し言葉アーカイブへの多言語<br>アクセス                                                                                                                                                           | 09-30-01<br>08-30-06 | Continuing<br>Grant<br>\$7,500,000<br>LARGE           | Suv<br>of the<br>Shoah<br>Vis<br>His F |

表 2.3-5 2001 年度採択 ITR プロジェクト（本節で紹介したもののみ）（その 2）

|     | Award Number<br>Title                                                                                                                                                                                                                                                                                                                           | 開始<br>終了             | 助成タイプ<br>総額（予想）                                           | 研究機関         |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|-----------------------------------------------------------|--------------|
| (5) | 121146 ITR/SY:<br>Center for Modeling of Quantum Dynamics,<br>Relaxation and Decoherence in Solid-State<br>Physics for Information Technology<br>Applications<br>IT 適用に向けた、固体物理学における量子力学、<br>緩和、および非コヒーレント性のモデル化のための<br>センター                                                                                                                    | 08-31-01<br>09-29-08 | Continuing<br>Grant<br>\$1,600,000<br>MATERIALS<br>THEORY | クラーク<br>ソン大学 |
| (6) | 121542 ITR/AP:<br>Collaborative Research: Modular Ocean Data<br>Assimilation<br>モジュラー海洋データアシミレーション（海洋<br>データの同化処理のソフトウェア部品による構築）<br>ほかの参加大学:<br>ラトガー大学（NSF#121506）\$762,500<br>カリフォルニア大学サンディエゴ校<br>（NSF#121332）\$240,000<br>イリノイ大学アーバナシャンペーン校<br>（NSF#121449）\$500,000<br>アリゾナ州立大学（NSF#121315）\$845,000<br>コロラド大学ボルダー校<br>（NSF#121176）\$149,717 | 01-01-02<br>01-31-07 | Continuing<br>Grant<br>\$2,370,339<br>MEDIUM<br>(GROUP)   | オレゴン<br>州立大学 |
| (7) | 0121657 ITR/SY:<br>Electronic Books for the Tele-immersion Age: A<br>New Paradigm for Teaching Surgical Procedures<br>遠隔没入（型仮想現実）時代の電子ブック:<br>手術手順教育のための新しいパラダイム                                                                                                                                                                                | 09-01-01<br>08-31-04 | Standard<br>Grant<br>\$1,750,000<br>MEDIUM<br>(GROUP)     | ブラウン<br>大学   |

## 2 米国政府が支援する情報技術研究開発の動向

### 2.3.1.2 ITR におけるデジタルライブラリ関連プロジェクト

表 2.3-6 ITR プログラムのデジタルライブラリ関連プロジェクト一覧（その 1）

|    | Award Number<br>Title                                                                                                                              | 開始<br>終了                            | 助成タイプ<br>総額（予想）                             | 研究機関                         |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|---------------------------------------------|------------------------------|
| 01 | 0114011 ITR/IM (IDM):<br>Developing A Collaborative Information and<br>Knowledge Management Infrastructure<br>協同的な情報・知識管理インフラストラクチャ                | 09-01-01<br>06-30-04<br>2年<br>10か月  | Standard<br>Grant<br>\$224,998<br>'01-S     | アリゾナ<br>大学                   |
| 02 | ITR:(NSF データベース中にエントリ発見不可)<br>A Multilingual Gazetteer System for Integrating<br>Spatial and Cultural Resources<br>空間的および文化的資源のための他言語ガゼッター<br>システム | 09-01-01<br>08-31-02<br>1年          | 助成タイプ不明<br>\$99,346<br>'01-S                | カリフォ<br>ルニア大<br>学バーク<br>レー校  |
| 03 | 0114435 ITR/IM:<br>Digitizing the Burney Collection of Early English<br>Newspapers<br>英国の初期の新聞の Burney コレクションを<br>デジタル化する                          | 08-15-01<br>06-30-02<br>10.5か月      | Standard<br>Grant<br>\$400,000<br>'01-S     | カリフォ<br>ルニア大<br>学リバー<br>サイド校 |
| 04 | 0114436 ITR/IM:<br>Creating a Latin American Union Catalogue and<br>Bibliography to 1850<br>1850年までのラテンアメリカの連合国のカタログと<br>参考図書一覧を作成する               | 10-01-01<br>06-30-02<br>9か月         | Standard<br>Grant<br>\$99,751<br>'01-S      | カリフォ<br>ルニア大<br>学リバー<br>サイド校 |
| 05 | 0121631 ITR/PE:<br>AVENUE: Adaptable Voice Translation for Minority<br>Languages<br>AVENUE: 少数民族言語のための適応可能な音声翻<br>訳                                | 09-15-01<br>09-30-06<br>5年<br>0.5か月 | Continuing<br>Grant<br>\$2,500,000<br>'01-M | カーネ<br>ギーメロ<br>ン大学           |
| 06 | 0121641 ITR/IM:<br>Capturing, Coordinating and Remembering<br>Human<br>Experience<br>人間の経験を捕捉、調整し記憶する                                              | 10-01-01<br>08-31-05<br>3年<br>11か月  | Continuing<br>Grant<br>\$2,000,000<br>'01-M | カーネ<br>ギーメロ<br>ン大学           |



表 2.3-7 ITR プログラムのデジタルライブラリ関連プロジェクト一覧（その2）

|    | Award Number<br>Title                                                                                                                                            | 開始<br>終了                               | 助成タイプ<br>総額（予想）                             | 研究機関                |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|---------------------------------------------|---------------------|
| 07 | 0113878 ITR/IM:<br>The Million Book Project<br>100 万冊の本プロジェクト                                                                                                    | 08-15-01<br>06-30-02<br>10.5 か月        | Standard<br>Grant<br>\$500,000<br>'01-S     | カーネ<br>ギーメロ<br>ン大学  |
| 08 | 0121239 ITR/AP+IM:<br>Computational Tools for Modeling, Visualizing and<br>Analyzing Historic and Archaeological Sites<br>歴史的、考古学的遺跡のモデル化、可視化、および<br>分析のための計算ツール | 09-15-01<br>09-30-05<br>4 年<br>0.5 か月  | Continuing<br>Grant<br>\$2,000,544<br>'01-M | コロンビ<br>ア大学         |
| 09 | 0121764 ITR/PE:<br>A Live Performance Simulation System: Virtual<br>Vaudeville<br>生演奏（ライブパフォーマンス）シミュレーション<br>システム: 仮想ボードビル<br>※ボードビル: 歌・踊り・コメディ等からなる<br>ショー      | 09-15-01<br>08-31-04<br>2 年<br>11.5 か月 | Continuing<br>Grant<br>\$900,000<br>'01-M   | ジョージ<br>ア大学         |
| 10 | 0081944 ITR:<br>An Active, Personalized, Adaptive, Multi-format<br>Biological Information Delivery System<br>能動的な、個人化された、適応的な、複数様式の<br>生物学情報提供システム               | 09-01-00<br>09-30-03<br>3 年<br>1 か月    | Continuing<br>Grant<br>\$494,297<br>'00-S   | インディ<br>アナ大学        |
| 11 | 0121285 ITR/IM+PE+SY:<br>Summer Workshops on Human Language<br>Technology: Integrating Research and Education<br>言語技術に関する夏期ワークショップ（複数）: 研究と<br>教育を統合する           | 10-01-01<br>09-30-06<br>5 年            | Continuing<br>Grant<br>\$1,750,000<br>'01-M | ジョンズ<br>ホプキン<br>ス大学 |

## 2 米国政府が支援する情報技術研究開発の動向

表 2.3-8 ITR プログラムのデジタルライブラリ関連プロジェクト一覧（その3）

|    | Award Number<br>Title                                                                                                                                            | 開始<br>終了                             | 助成タイプ<br>総額（予想）                             | 研究機関                                                            |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------|---------------------------------------------|-----------------------------------------------------------------|
| 12 | 0114391 ITR/SY (CISE):<br>Geometrical Image Representation<br>幾何学的なイメージ表現<br>※イメージ圧縮方式研究。イメージ検索効率向上                                                              | 09-01-01<br>09-30-04<br>3年<br>1か月    | Continuing<br>Grant<br>\$449,973<br>'01-S   | ニュー<br>ヨーク大<br>学                                                |
| 13 | 0121446 ITR/IM:<br>3D Shape-Based Retrieval and Its Applications<br>3次元形状に基づく検索とその応用                                                                             | 09-15-01<br>07-31-04<br>2年<br>10.5か月 | Standard<br>Grant<br>\$500,000<br>'01-S     | プリン<br>ストン大<br>学                                                |
| 14 | 0113427 ITR:<br>Solving the Puzzle of the Forma Urbis Romae<br>古代ローマの巨大な大理石の地図「the Forma Urbis<br>Romae」のパズルを解く<br>※残りの断片を組み入れるのに、コンピュータに<br>よる形状マッチングアルゴリズムを用いる | 10-01-01<br>09-30-04<br>3年           | Continuing<br>Grant<br>\$449,995<br>'01-S   | スタン<br>フォード<br>大学                                               |
| 15 | 0122466 ITR Large:<br>MALACH: Multilingual Access to Large<br>spoken ArChives<br>MALACH: 「話し言葉」の大規模アーカイブへの<br>多言語アクセス                                            | 10-01-01<br>08-31-06<br>4年<br>11か月   | Continuing<br>Grant<br>\$7,500,000<br>'01-L | Survivors<br>of the<br>Shoah<br>Visual<br>History<br>Foundation |
| 16 | 0113007 ITR/AP(CISE):<br>Capturing and Modeling Physics from Images<br>イメージから物理を捕捉しモデル化する<br>※写真やビデオを解析し実世界の現象の物理的<br>特性を捕捉・モデル化する                               | 09-15-01<br>09-30-04<br>3年<br>0.5か月  | Continuing<br>Grant<br>\$463,532<br>'01-S   | ワシント<br>ン大<br>学                                                 |

ITR プログラムで、デジタルライブラリ関連のプロジェクトが採択されるようになった。DLI2 (Digital Library Initiative, Phase 2) のホームページから「Special Project ITR」を開くと、そのプロジェクト一覧を見ることができる(表 2.3-6、表 2.3-7、表 2.3-8) [11]。内訳は、2000 年度採択のプロジェクトが 1 件、残りは 2001 年度採択のプロジェクトで 15 件ある。プロジェクト概要を見ると、従来 DLI2 で扱われていた領域に該当しており、これは DLI2 (1988-2003) の完了時期が近づいたので、今後のデジタルライブラリ関連研究は ITR で扱う方針となったものと思われる。2001 年度には、DLI2 プロジェクトの新規募集は行われていない。

### 2.3.1.3 2002 年度 ITR プログラム

ITR の 2002 年度プロジェクトは、現在、公募・審査中であり、今回も 9 月に採択プロジェクトが発表される予定である[12]。今回の重点領域は、学際領域、いわゆる複数分野の融合領域である。以下の 3 つの領域について募集を行っている。これらの複数にまたがるものも推奨される。

- (1) ソフトウェアとハードウェアのシステム (software and hardware systems)。
- (2) 個人の能力の拡大と社会の変革  
(augmenting individuals and transforming society)。
- (3) IT による科学のフロンティアの前進  
(advancement of the frontiers of science via information technology)。

以下に、各領域の要件概要を簡単に説明する。詳細については参考文献[12]を参照されたい。

#### 2.3.1.3.1 ソフトウェアとハードウェアのシステム

この領域の提案は、研究および教育のための新しい複合 (複雑) システムを創出するのを助けるものでなければならない。例えば、

- ・ ハードウェアとソフトウェアのバランス。計算と通信のバランス。
- ・ 超大規模なシステムまでのスケールアップ (コンポーネント数、地理的分散、参加人数、装置数など)。
- ・ 巨大システムの信頼性、パフォーマンス予測。
- ・ 建物、設備、場所への組み込みデバイスと通信。人やグループとの各種共有。
- ・ 物理的システム内での情報技術を使った実験。

## 2 米国政府が支援する情報技術研究開発の動向

これらの問題には、グリッドコンピューティング、仮想現実感/テレプレゼンス、ハイブリッドおよび組み込みシステム、ハイパフォーマンスネットワークとミドルウェアアプリケーションなどが含まれる。

### 2.3.1.3.2 個人の能力の拡大と社会の変革

この重点領域は、どうすれば、より多くの人により多くの仕事に IT を使うことができるようになるかを研究する。

さまざまな活動、職業、および社会的コンテキストにおいて、人間のパフォーマンスと体験を増大させるために、人間の肉体的、精神的、および感覚的な能力を拡大するなど、一般に、この分野への提案は、新しい方法で情報技術を効果的に使って、我々自身や社会の能力を向上させる研究である。

例えば、次のような研究が該当する。

- ・ 人間能力を拡大する新しい物理デバイスやシステム技術。
- ・ 社会的用途のための情報利用。これはデジタルライブラリや多言語システム等も含む。
- ・ 教育、学習、共同作業、および教育環境創出を支援する技術。
- ・ ビジネス、遠隔通勤、モバイルや分散ワークの実行、組織の生産性、e-コマースや電子マーケットの新技术等における変化。
- ・ 市民参加、政府施設と構成員間の対話機会の向上。
- ・ 倫理、価値、プライバシーを配慮した情報設計。知的財産の保護と侵害。
- ・ IT 設計を社会的、経済的結果に結びつける理論と手法。
- ・ 情報技術によって引き起こされる科学、工学、人文科学での変化。

### 2.3.1.3.3 IT による科学のフロンティアの前進

この重点領域は、IT と科学・工学の接点における研究教育活動を支援する。プロジェクトは理論でも実験に基づくものでもよく、以下のことが推奨される。

- ・ 新しい機能を持った新しい装置を創作する。
- ・ 新しい科学・技術の概念を探求するために計算を使う。
- ・ コンピューティングの新しいパラダイムを開発する（例：量子コンピューティング）。

プロジェクトの例には、以下のようなものが含まれる。

- ・ 大規模な予測モデル。さまざまな規模の現象の分析、モデル化、最適化。
- ・ 複雑系システムの計算手法。人間対話の確率モデル。幾何学的新アルゴリズム。
- ・ 大規模不均質データセットからの知識の発見、可視化、解釈。

- ・ 多重情報源のデータ統合。革新的データベース構造と情報検索の新方法。
- ・ 高信頼ソフトウェアとシステムのためのツール。
- ・ 科学・工学におけるインターネットの革新的利用。

例えば、以下のような科学・工学的課題は、この領域に該当する。

- ・ 生物システム情報学。生物学コンテキストでのデータマイニング。
- ・ 地球システムにおける広範な現象予測。複数観測点からの異種環境データセットのモデルへの統合。
- ・ 宇宙物理、数理物理学と生物・環境科学との接点、量子情報科学、ナノスケール科学、新計算的手法、例えば、大量データセットの収集と分散型分析のためのコンピューテーショナルデータグリッド。
- ・ 次世代 IT 製造技術を加速するハイエンドコンピューティングツール。シミュレーション。微細構造に基づく材料設計。分散型仮想製造・サプライチェーン・配送システムの計画・制御。
- ・ IT と心理学、言語科学が連携した言語情報技術。

### 2.3.2 デジタルライブラリ・イニシアチブ (電子図書館構想、Digital Library Initiative)

この世は今や、まさに情報社会、インターネット社会となりつつある。インターネット上を行き交い、あるいはインターネット上に蓄えられるデジタル化された電子情報は、多様化をきわめるとともに、ますますその重要性が増大しつつある。デジタルライブラリの研究は、従来の図書館の枠を超えて、インターネット上のデジタル情報資源の利用と管理の全般にわたる基盤技術として、きわめて重要である。

1994年にスタートした電子図書館構想(DLI: Digital Library Initiative)は、だれでもがアクセス可能な人類の知識の電子貯蔵庫の夢を実現する上で、直面せざるを得ない概念、構造、および計算機能に対する課題に取り組んでいる。当初のDLI計画(DLI-1: Digital Library Initiative, Phase 1, 1994-1998)は、6つの大学の6つのプロジェクトからなる、コンピュータサイエンス指向のこぢんまりとしたものであったにもかかわらず、参加した研究者たちの共同研究の成果は、Lycos および Google の検索エンジンや Go2Net のような商用の副産物を生み出し、多くの研究者の注目を浴び、デジタル情報資源の大いなる可能性を際立たせた。

1999年度に開始されたDLIフェーズ2(DLI-2, 1998-2003)では、DLI-1の成功を受け、プロジェクト数も大幅に増大するとともに、より広範な研究活動へと広がりを見せて

## 2 米国政府が支援する情報技術研究開発の動向

いる。これらの研究では、今日のますます増大する計算能力と帯域幅容量を使って、大規模な分散型の電子コレクションを、広域的な知識ネットワークを通して、アクセス・相互運用可能にし、かつ役立つものにするという目標の実現をはかっている。

このプロジェクトも最終段階となり、この1年間、新規プロジェクトはまったく追加されていない。2001年度には、2.3.1.2節で述べたように、代わりに ITR プログラムの方で、デジタルライブラリ関連のプロジェクトが採択されている。現在公募・選考作業中の2002年度 ITR プロジェクトにおいても、重点領域の項目の中で「デジタルライブラリ」に言及されており、今後、デジタルライブラリ関連の研究は ITR の中で行われていくものと思われる。

デジタルライブラリ・イニシアチブの歴史的経緯や、成果のインパクト等については、本年度 AITEC 報告書の別冊[1]に報告があるので、詳細はそちらを参照頂きたい。

## 2.4 おわりに

本節では、情報技術に関連する米国政府支援の研究開発動向の概要を簡単に説明した。

ブッシュ大統領の 2003 年度予算要求では、防衛 R&D、健康 R&D の予算は大幅に増大し、連邦政府研究予算の 4 分の 3 以上を占める。一方、それ以外の国家ミッションのほとんどに対する予算はむしろ減少する。そのような中でも情報技術の研究開発は、重点分野の 1 つとして、なんとか着実に伸びていると言える。重点分野である省庁間連携イニシアチブを見ても、NITRD は、比較的新しい NNI（国家ナノテクノロジー・イニシアチブ）などに比べると、伸び率は小さいが、投入規模では最大を保っている。

ブッシュ政権の政策方針として、「減税」、「小さな政府」があり、情報技術の研究開発においても、産業界でできることは産業界にまかせ、政府は産業界では困難な基礎研究を支援するという方向は、さらに進むものと思われる。産業・科学技術はもちろん、あらゆる面でリーダーであり続けたい米国にとって、他国より一歩先んじるための基礎研究および優れた技術者の育成は、きわめて重要である。その点では、NSF の 2003 年度計画に向けた方針説明は、その辺のところを良く説明しているので、以下に引用しておく [9]。



NSF の支援で開発されたこのロボットは、9 月 11 日に続くワールドトレードセンターでの捜索と救助活動の、重要メンバーだった。救助隊員や犬たちが近寄れないつぶれた建物の場所に入ることができた（写真提供：南フロリダ大学）。

図 2.4-1 靴箱サイズのロボット

## 2 米国政府が支援する情報技術研究開発の動向

1990年代における生産性の成長は、包括的な基礎研究と教育の活動の経済的恩恵を、十分に明らかにした。この成長は、新しい知識によって力を与えられ、技術革新によって後押しされたものである。新製品、新プロセス、および新産業などすべてが、研究の急速な進歩に依存している。高度に競争的な世界経済の中では、科学技術の継続的進歩は、米国の継続的なリーダーシップに通じる最も大切な道である。

9月11日の事件（同時多発テロ）とそれに続く炭疽菌攻撃は、科学技術に強い国家であれば、危機や国家を取り巻く環境変化に対して、迅速かつ効率的に対応することができることを証明した。

科学技術分野の全範囲にわたる基盤的研究は、研究と革新を可能とする高度に技術を身につけた労働力と相まって、例外的ニーズが生じたときに、国家が引き出すことのできる知識資本を提供するものである。研究の最前線に重点を置いた知識の蓄積の成長は、そのような対応の可能な選択肢を増大させる。才能のある高度の技術を身につけた科学技術労働力は、不測のニーズを満足させるための新たな技術開発を加速する。

### 参考文献

- [1] 調査資料「重点分野に関する調査 ―米国における IT と応用分野の融合領域の研究開発に関する調査 I―（電子図書館、バイオテクノロジー、インフラストラクチャ）」、AITEC H13 年度報告書。
- [2] A Preview of AAAS Report XXVII: Research and Development FY 2003, <http://www.aaas.org/spp/dspp/rd/prev03p.htm>
- [3] BlueBook 2002: "Networking and Information Technology Research and Development", <http://www.itrd.gov/pubs/blue02/index.html>, BlueBook2002 AITEC 編集日本語版「ネットワークング及び情報技術研究開発」, <http://www.icot.or.jp/FTS/Ronbun/BlueBook2002-J.pdf>
- [4] NGI ホームページ"About NGI", <http://www.ngi.gov/>
- [5] Analytical Perspectives, "8 Research and Development" (pdf), <http://www.ostp.gov/html/ap08.pdf>
- [6] ITR - Information Technology Research, <http://www.itr.nsf.gov/>



- [7] Information Technology Research: Investing in Our Future,  
<http://www.itrd.gov/ac/report/>, PITAC レポート 情報テクノロジーに関する調査報告:  
未来への投資, AITEC H10 年度報告書「わが国が行う情報技術の研究開発のあり方に関  
する調査研究 (その 3)」 付属資料 8.
- [8] Information Technology for the Twenty-first Century: A Bold Investment in  
America's Future, working draft, <http://www.itrd.gov/archive/it2/initiative.pdf> (Jan.  
1999), IT<sup>2</sup> 構想 21 世紀のための情報技術: 米国の将来に対する大胆な投資, AITEC  
H10 年度報告書「わが国が行う情報技術の研究開発のあり方に関する調査研究 (その  
3)」 付属資料 9.
- [9] NSF Summary of FY2003 Budget Request to Congress,  
<http://www.nsf.gov/bfa/bud/fy2003/start.htm>
- [10] NSF Announces \$156 Million in Awards for Information Technology Research,  
NSF Press Release, <http://www.nsf.gov/od/lpa/news/press/01/pr0174.htm> (Sep.  
2001).
- [11] DLI2 Home - Information Technology Research (ITR) Projects,  
<http://www.dli2.nsf.gov/itrprojects.html>
- [12] Information Technology Research (ITR) Program Solicitation,  
<http://www.nsf.gov/cgi-bin/getpub?nsf01149>
- [13] AAAS Report XXVII: Research & Development FY 2003,  
<http://www.aaas.org/spp/dspp/rd/03pch23.htm>
- [14] Interagency Working Group on Information Technology R&D Organization Map,  
<http://www.itrd.gov/iwg/orgchart.html>

### 3. 研究開発の新しい展開と内外の動向

#### 3.1 ロボットの耳は2つで十分か

報告者: 奥乃博主査

聴覚は人間にとって最も重要な感覚である。言語によるコミュニケーションが聴覚によって成立することは容易に理解されるが、「ヒトは聴覚によってのみ言語を獲得し、そこに文化が生まれ、継承される。書かれた言語は目によって伝承されるが、話す言葉は耳からしか得られない。話し言葉があつて書く言葉が生まれる」ことを、多くの人が理解していないのは残念なことである。

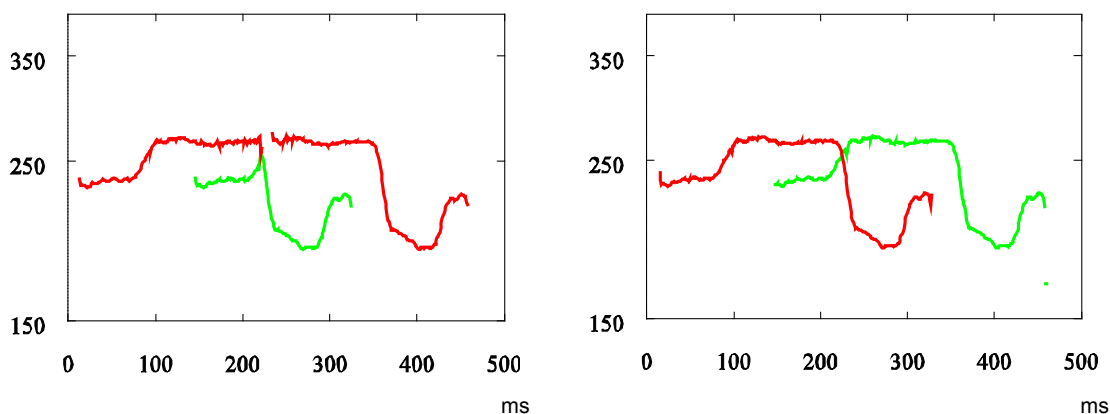
鈴木淳一、小林武夫共著『耳科学 - 難聴に挑む』（中公新書、1598、2001）

##### 3.1.1 はじめに

鉄腕アトムの誕生の年 2003 年が近づき、ヒューマノイドと呼ばれる人間型ロボットが数多く開発されるようになった。新千年期までは、早稲田大学で 30 年間以上にわたるヒューマノイド研究で開発された 2 足歩行ロボット Wabian や人間との協調作業を行う Hadaly など、あるいは、ホンダで 10 年間以上にわたり開発されてきた P3 など、一連のヒューマノイド群ぐらいであった。

今日では、ホンダの ASIMO や科学技術振興事業団北野共生システムプロジェクトの PINO はテレビ放送の CM にも登場している。科学技術振興事業団川人プロジェクトの油圧式制御による DB は、脳神経科学の成果を生かした複雑な運動制御ができ、見る人を驚かせる。映画スターウォーズに登場するロボット R2D2 の形をした NEC の Papero や ATR の Robovie などのほかに、動物型のロボットも次々と登場してきている。

これらのロボットの多くは、マイクフォンを装備している。その本数は、2 本、ないし、3 本のものが多いが、マイクフォンアレイを搭載しているものもある。本節では、「耳が 2 つある」ことがロボット聴覚において本質的であるかを検討する。なお、ヒトの聴覚については生理学や心理物理学などで数多くの知見[1]が得られているので、そのような研究は他の解説に譲る。



(a) マイクロフォン 1 本での間違った分離 (b) マイクロフォン 2 本での正しい分離

濃い線と薄い線が分離された音声を示す。

図 3.1-1 2 話者同時発話分離でのあいまい性

### 3.1.2 我々が日常聞く音は混合音

一般に、我々が日常生活で耳にする音は単一音源からの音ではなく、複数の音源からの音が混じった混合音である[11]。2 つの音から構成される混合音の分離を 1 本のマイクロフォンで行うと、分離結果にあいまい性が含まれる恐れがある。

例えば、時刻ごとに調波構造を抽出し、基本周波数の時間的連続性を基にグルーピングすることによって、調波構造をもつ音響ストリーム（以下、調波構造ストリームと呼ぶ）を分離するとしよう。このシステムで、同じ話者の「あいうえお」という発話をずらして録音したモノラル音を分離させると、一方が「あいうーえお」、もう一方が「あいえお」と誤分離される場合がある（図 3.1-1）[6]。

もし、2 つの音が別々の方向から到達する場合には、そのような音を 2 つのマイクロフォンで聞くことができれば、方向情報が利用できることで、調波構造ストリーム分離でのあいまい性は解消できる。図 3.1-1 に、同じ発話を、時間を 150ms ずらして、左右 30 度から行ったときの分離結果を示す。図からわかるように、マイクロフォン 1 本の際には、調波構造のグルーピングが間違ってしまうのに対して、マイクロフォン 2 本を使用した分離では、方向情報が利用できることで、調波構造だけを手がかりとした分離で遭遇するあいまい性は解消されている。

### 3.1.2.1 混合音分離の手法

混合音の分離には、マイクロフォンアレイが一般的に使用される。マイクロフォンアレイで得られたマルチチャネルのデータから特定方向の音だけを抽出するには、ビームフォーミングや独立成分解析 (ICA: Independent Component Analysis) といった技法がよく使われる。ビームフォーミングの原理は、「 $N+1$  本のマイクロフォンで、 $N$  個の音響的な死角が作成できる」というものである。ビームフォーミング技法で音源分離によく使用されるのは、遅延型加算 (delayed sum) である。

すべての音源が情報論的に互いに独立であるとすると、独立成分解析を使えば、 $N$  本のマイクロフォンで  $N$  個の音源を理論的に分離することができる[5]。実際、2 本のマイクロフォンによる 2 話者同時発話分離では、独立成分解析の方が、調波構造と方向情報を手がかりにした分離よりも、分離音の音声認識結果がよいという結果が得られている[9]。Wang からも複数のマイクロフォンを使用することによって、音響ストリーム分離の精度の向上を達成している[13]。

つまり、マイクロフォンの数を増やせば、音源分離の性能が向上することは、理論的にも、実験的にも明らかである。では、マイクロフォンは何本用意すればよいのか? 「想定される音源の数より多ければよい」が正解なのであろうか。

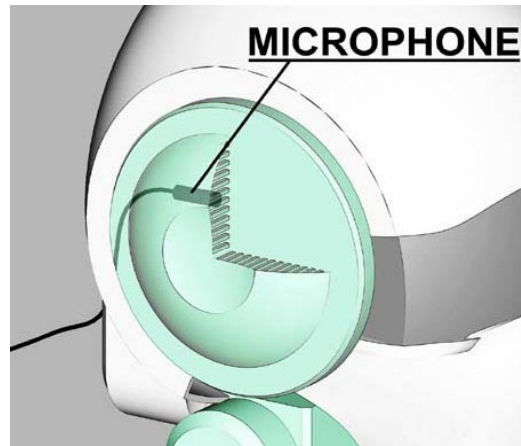
一般環境では、マイクロフォンの位置が変化したり、音源が移動したりするので、必ずしも理論通りに行かない。体の動くロボットに装着したマイクロフォンは、頻繁に動き、ロボットの体に何十本というマイクロフォンを装着することも現実的ではない。もちろん、このような状況にも適用可能なマイクロフォンアレイを開発する研究プロジェクトも始まっている[12]。

報告者らは、このような問題を「マイクロフォンの数が音源の数よりも少ないときに、音源分離を行うにはどのようにしたらよいのか」ととらえ、研究を進めてきた。次のように問題は整理できよう。

- (1) 体が固定されたときに、耳がいくつ必要か。
- (2) 逆に、自由に体の動きが許されるときに、耳がいくつ必要か。

前者の問題については、方向情報を得るためには、最低 2 本のマイクロフォンは必要であると考えている。後者の問題については、最少解は 1 本のマイクロフォンであろう。しかし、上述したような理由から、実用的な意味での最少解は、2 本のマイクロフォンではないかと予想している。3 本のマイクロフォンを装備しているロボットもあるが、処理の基本は 2 本のマイクロフォンであり、2 本のマイクロフォンでは難しい「前後問題」を残り 1 本のマイクロフォンで解決している[2]。

以下、2 つの耳で十分かを検討をする。



(a) 上半身ヒューマノイド SIG の全体像      (b) 耳の中のマイクロフォンの位置

図 3.1-2 上半身ヒューマノイド SIG

### 3.1.2.2 2本のマイクロフォンによる音源定位

人間の耳の聴覚機能として、音源定位のモデルとして知られている Jeffress の cross-correlator モデルは、2本のマイクロフォンから得られた入力音を遅延させながら相関をとり、最大の相関を与える遅延から両耳間時間差 (Interaural Time Difference: ITD) を求めるものである[3]。

Jeffress のモデルは、ITD しか使用していないが、のちの研究では両耳間強度差 (Interaural Intensity Difference: IID) も音源定位に寄与するように修正されている。ITD と IID は周波数帯域によってその貢献度が異なる。一般のヒトの場合には、1500Hz 付近を境に、それ以下では ITD が、それ以上では IID が優位であることが知られている[4]。

ITD の代わりに両耳間位相差 (Interaural Phase Difference: IPD) を使用すると、ITD と IID の役割が変わる 1500Hz という値を決める 1つの要因が得られる。両耳間の顔の表面上の距離がおおよそ 23cm とすると、IPD が初めて 2ss になる、つまり、定位の際に、1 周期まわり込みによって簡単にあいまい性が解消できなくなるのが、1500Hz である。ただし、ss となる 750Hz から、前後がわからなくなるというあいまい性が生じる。報告者らは図 3.1-2 に示す SIG と呼ぶ上半身ロボットのための聴覚機構を開発中である。SIG

の両耳間距離は 18cm、頭部形状を考慮し、外装に沿った両耳間距離が約 23cm となり、ほぼヒトと同じような値で、IPD と IID の役割が変わる。

2本のマイクロフォンを使用して音源定位を行うためには、IPD や IID の方向情報との関係を示す頭部伝達関数 (Head-Related Transfer Function: HRTF) が必要である。

中谷と奥乃らのグループは調波構造に注目した音源分離システムを構築し、無響室環境で 2 話者同時発話の音声分離を行っている[6]。このシステムは、調波構造を抽出し、左右の耳で同じ音源から来る調波構造を持った音のペアを抽出し、その音の各倍音の IPD と IID を計算し、それらの値から HRTF を基に方向情報を求める。このようにして求めた調波構造断片を方向情報でグルーピングし、調波構造ストリームを抽出する。この方法は、基本周波数が重ならない限り、理論的には音源数がいくつあっても音源定位とそれに基づいた音源分離が可能である。実際には、混合音中の調波構造抽出の精度は音源数が増えるにしたがって低下するので、音源数が増えると精度のよい分離は難しくなる。例えば、2 話者同時発話の上位 10 位の単語認識率 75%が、3 話者同時発話認識では 30%程度と大幅に分離性能が劣化する[10]。

このような問題点を回避するため、著者らは、特定の音源方向から来る音だけを分離する方向通過型フィルタ (Direction-Pass Filter: DPF) を設計した。DPF における方向推定のアイデアは、IPD と IID に関して、実際の値と所与の方向情報から得られる HRTF からの値との間の距離から確信度を計算し、IPD と IID の確信度を Dempster-Shafer 理論で統合するという仮説推論にある。

具体的には、各サブバンド (離散フーリエ変換、DFT の各点) で、IPD と IID を求める。次に、それらと所与の方向情報の HRTF から求めた IPD と IID との距離を計算し、確率密度関数を用いて、それらを確信度に変換する。最後に、2 つの確信度を Dempster-Shafer 理論で統合し、統合確信度の高いサブバンドだけを集めて、逆離散フーリエ変換で分離音を再構成する。無響室環境における DPF による 3 話者同時発話認識では、ほぼ単一音源並みの単語認識率が達成されている[10]。

HRTF は無響室で、スピーカの位置を変え (例えば 5 度刻み)、インパルス応答を測定することによって求める。そのために、無響室以外では、実際の環境での空間伝達特性を畳み込まないと、得られる方向情報の精度が低下する。マイクロフォンの位置が変化するようなシステムの場合には、HRTF の測定と同様にスピーカ位置を変化させるだけでなく、マイクロフォンの位置も変化させて、空間伝達特性を測定する必要がある。また、測定点数の制限から離散的な方向情報しか使えないので、移動音源への対処は難しくなる。

ここまでの議論では、静的な環境での音源分離であった。それに対して、実環境で動き回るロボット、あるいは、頭部を動かす機能があるロボット、あるいは、行動と知覚が結びついたアクティブパーセプションでは、マイクロフォン自体が動き、それに伴って、モータ雑音や機械音が発生する。このような音は、例え小さくてもマイクロフォンに近いので、相対的に大きな雑音となり、外部からの音の信号雑音比が低下する。

ロボットやシステムが発生する内部雑音を軽減する最も簡便な方法は、動作を中断してから、聞くことである。実際、このような"stop-perceive-act"法を、マイクロフォンを搭載した大部分のロボットが採用している。

アクティブオーディションでは、内部雑音や自己生成音の抑制がきわめて重要である。特に、ロボットヒューマンインタラクションでは、自分の発話を削除し、相手の発話の信号雑音比を向上させることも必要である。また、よく聞こうとして動いたところ、自分の出す音が災いして、かえって聞こえ難くなるということも想定される。

### 3.1.3 複数話者実時間追跡システム

本節では、2つの耳でどのような機能が実現可能かを示すために、図 3.1-2 に示した SIG の上に実現されている複数話者実時間追跡システムを紹介する。このシステムは、繁華街の大通りに面したマンションの一室 (3m×3m) に置かれている。

#### 3.1.3.1 2本のマイクロフォンによる音響と画像を統合した実時間複数話者追跡

本システムのアイデアは、以下の通りである[7]。

(1) HRTF 使用上の問題点である、部屋の伝達特性が必要であり、離散的な点の HRTF しか使えない、という 2 点を解決するために、聴覚エピポラ幾何学を提案し、HRTF に相当する部分を代用する。これは、ステレオ画像処理でのエピポラ幾何学の焼き直しであり、ステレオカメラとマイクロフォンの位置が一定であることを利用して、画像処理から得られる 3 次元位置を使って、IPD と IID を求める。音源が無限遠にあるときには、聴覚エピポラ幾何学で求めた値と、幾何学的に求めた IPD や IID の値[4]とは一致する。

(2) SIG では、体内に有する 1 対のマイクロフォンから取得されるモータ音や機械音についての簡単なモデルを持っており、モータが稼働中でモデルに合うような音が発生すると、ヒューリスティクスを用いて、破壊されているサブバンドを推定し、破棄する。ヒューリスティクスによる方法を採用したのは、FIR フィルタを応用したアクティブノイズキャンセレーションでは、IPD を計算するために必要な左右の位相特性の線形性が実データでは成立しないからである。

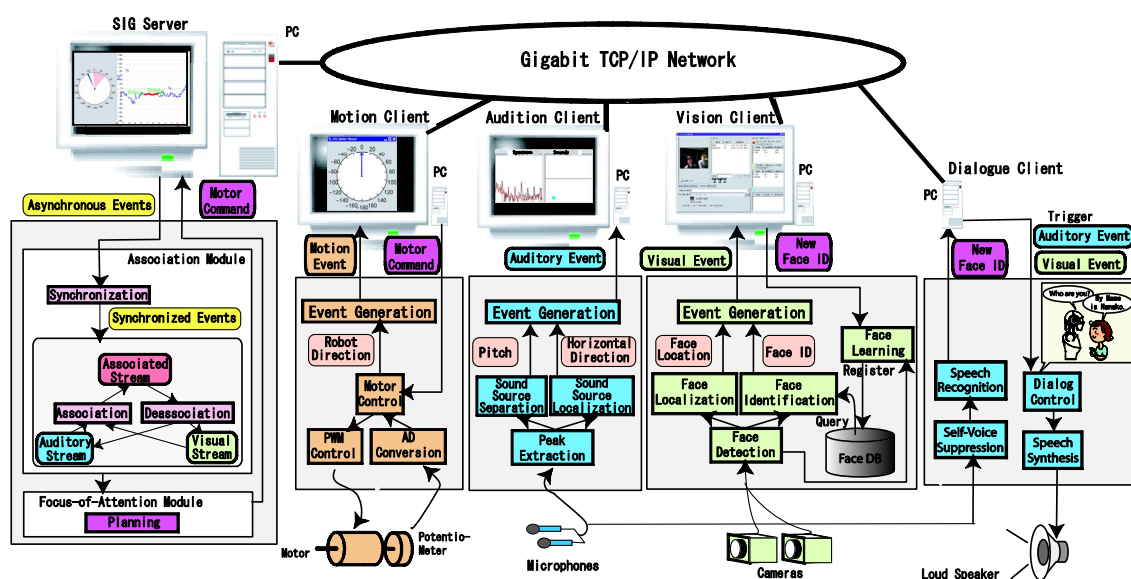


図 3.1-3 視聴覚統合による実時間話者追跡システム

システム全体の構成を図 3.1-3 に示す[8]。システムは、音響処理部、画像処理部、モータ駆動部、アソシエーション部、対話管理部、注意制御部、および、サーバから構成されている。これらは 5 台の PC 上に配置されており、Gigabit Ethernet で接続されている。音響処理部は DPF と同じ考え方で、音源方向を抽出する。ただし、画像からの音源方向が得られない場合には、すべての方向について仮説を生成することによって、音源方向を抽出する。抽出した方向情報は、確信度とともにアソシエーション部に送られる。画像処理部は、肌色抽出により顔を発見し、ステレオ画像処理により、3 次元情報を取得する。さらに、抽出した顔ごとに顔認識を行い、3 次元情報と顔 ID 情報をアソシエーション部に送る。モータ駆動部からは、現在の SIG の体の向きについての情報をアソシエーション部に送る。

アソシエーション部では、各モジュールから得られる情報（方向や顔）を同期させ、音響ストリームと画像ストリームを構築する。次に、音響ストリームと画像ストリームを、時間的継続性や距離的近さを基にグルーピングをし、アソシエーションストリームを構成する。一定時間、音響ストリームや画像ストリームが消失すると、アソシエーションは解除される。



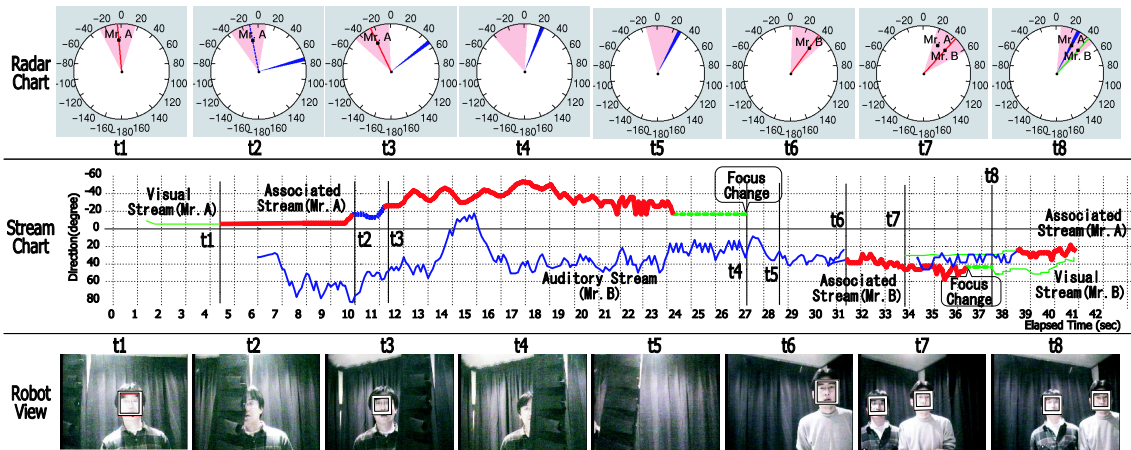


図 3.1-4 2 話者追跡の時間経過

実際の 2 話者追跡の時間経過を図 3.1-4 に示す。時刻 t1 で左側の人 (Mr. A) がしゃべり始め、画像ストリーム (細い線) と音響ストリームがアソシエートされ、アソシエーションストリーム (太い線) が構築される。Mr. A はしゃべりながら移動し、t2 で物影に隠れ始め、t3 で再び現われる。この間、アソシエーションは一時的に解除される。Mr. A はさらに移動し、しゃべるのを止め、t4 で再び物影に隠れ始める。システムは、Mr. A のストリームのアソシエーションを解除し、画像ストリームへ変更し、さらに、t4 で消滅させる。

一方、Mr. B は Mr. A から少し遅れてしゃべり出し、音響ストリームが構築される。図 3.1-4 から、音響ストリームの音源定位の精度は画像よりも格段に劣ることがわかる。システムは、t4 で Mr. B の音響ストリームに注意を向け、回転をし、t5 では何も見えず、t6 で Mr. B の顔を発見し、画像ストリームと音響ストリームをアソシエートする。t7 で Mr. A が現れ、また、しゃべり出し、同様にアソシエーションストリームが構築される。顔認識により、システムは t4 と t7 の 2 名が同一人物であることがわかる。t8 で Mr. B がしゃべり終わり、注意が Mr. A に移る。以上のように、システムは、内部的には複数の話者や人物の位置を常時把握している。

### 3.1.3.2 注意制御部

注意制御部は、システムが保持する話者や人物情報をもとに、どの対象に注意を向け、



(a) 「こんにちは」と声をかけると、声のする方に振り向き、顔認識をする。



(b) 知らない人だと「どちらさまですか」と応答し、話者の返答から音声認識によって名前を取得し、顔と名前の対応づけをすることにより来客者を登録する。

図 3.1-5 受付嬢としての SIG の振る舞い

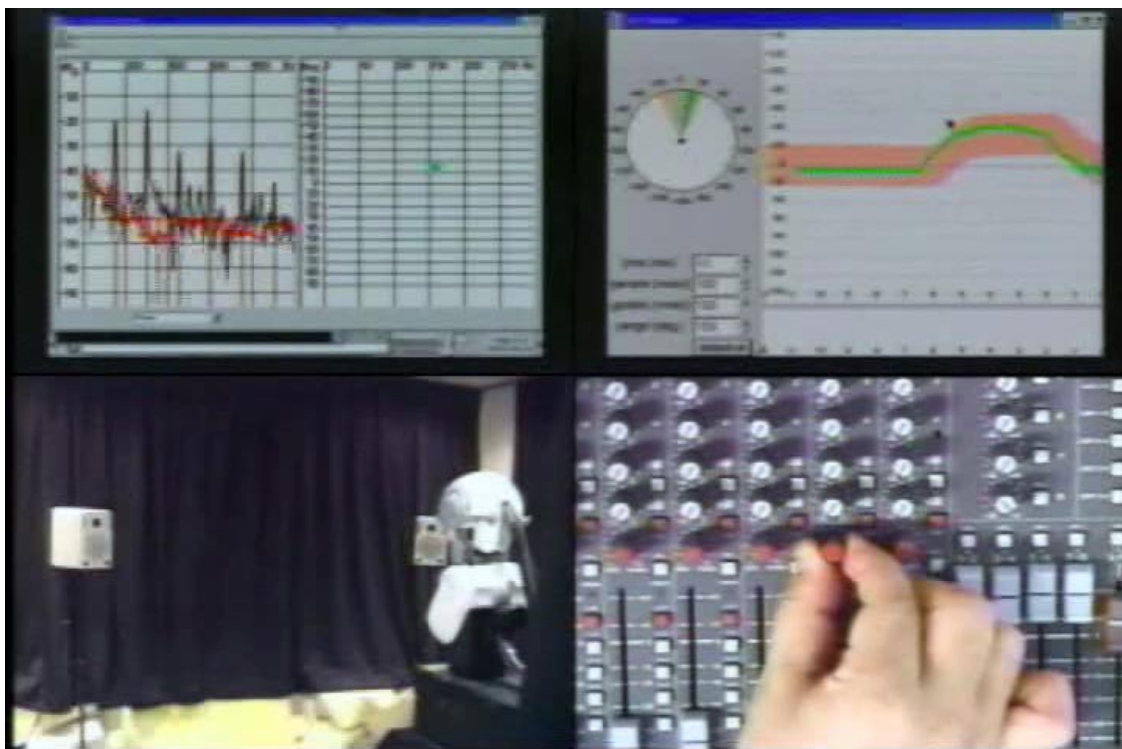
体を回転し、正対するのかを制御する。この部分は完全にプログラム可能であるので、いくつかのシナリオで、SIG の挙動を紹介する。

受付嬢: 話している人に注意を向けるのが第一目的であり、そうでなければ、音のする方向に振り向く。図 3.1-4 に示したように、注意を置くべきストリームの優先順位は、(1) アソシエーションストリーム、(2) 音響ストリーム、(3) 画像ストリームである。

具体的な挙動を図 3.1-5 に示す。来客が知らない人だと、「どちらさまですか」という応答をし、名前と顔との照合データを登録する。既知の人の場合には、「こんにちは、XX さんでいらっしゃいますか」とその人の名前を呼び、確認を行う。このように、対話管理部は顔認識・音声認識・音声合成を含んでおり、自己生成音の抑制と音声の分離を行っている。音声認識はフリーのディクテーションソフトウェアである Julius を使用し、音声合成は市販のソフトウェアを使用している。



(a) 4人の話者で声のする方を向く。



(b) バランスコントロールの音源定位に従従する。

図 3.1-6 SIG の音源追跡のさまざまな評価実験

コンパニオン: 「聞き耳を立てる」ために、新たな音のした方に注意を払うように、(1) 音響ストリーム、(2) アソシエーションストリーム、(3) 画像ストリームという優先順位を設定する。4人の中で話し声に顔を向ける実験風景を図 3.1-6 (a) に示す。

仮想音源: 追跡するものは、人物だけではなく、仮想的な音源であってもよい。基本周波数が 100Hz の調波構造を持つ音を左右のスピーカから流し、そのバランスコントロールを変化させることで、仮想的に音源を左右に振る。SIG は仮想的な音源を正しく追跡する (図 3.1-6 (b))。

以上の例から、2本のマイクロフォンでも、画像処理と統合すると、実時間で音源定位を行い、話者追跡を行えることがわかった。2本のマイクロフォンによる音源定位の前後問題のあいまい性は、画像を使えば解消はできよう。また、アクティブオーディションにより、体の回転に対して、音源がどちらに動くかという情報からも、前後問題のあいまい性は解消できよう。

#### 3.1.4 おわりに

本節では、ロボットには耳 (マイクロフォン) は2つで十分かという問いに対して、現状で実現できた聴覚機能について、著者の考え方も含めて解説をした。体を動かして聞くというアクティブオーディション、音響処理と画像処理 (3次元位置情報と顔認識) やモータ処理との情報統合が重要である。このために、方向通過型フィルタや聴覚エピポーラ幾何学、実時間処理方法を開発してきた。今後の課題としては、複数話者の話者認識や音声認識が挙げられる。このためには、話者認識や音声認識のフロントエンドとしての精度のよい音源分離、および、雑音の混入により失われたデータをうまく回避できる話者認識や音声認識の双方からの研究が必要であろう。

冒頭に引用した一節からも、聴覚機能というのは人間生活にとってきわめて重要であり、そのような原点に戻り、音響処理の新しい枠組みを打ち立てることが、音声認識システムが一般的となってきた今求められてことではないであろうか。本節が、そのような新しい潮流の一助となれば幸いである。

#### 参考文献

- [15] Bregman, A. S.: Auditory Scene Analysis - the Perceptual Organization of Sound, The MIT Press (1990).
- [16] Huang, J.: Spatial Sound Processing for a Hearing Robot, Enabling Society with

Information Technology, LNCS, Springer-Verlag (2001).

- [17] Jeffress, L.A.: A Place Theory of Sound Localization, *Journal of Comparative Physiology*, Vol. 41, pp. 35-39 (1948).
- [18] ムーア, B. C. J., 大串ほか (訳): 聴覚心理学概論, 誠信書房 (1994).
- [19] Murata, N. and Ikeda, S.: An On-line Algorithm for Blind Source Separation on Speech Signals, *Proceedings of 1998 International Symposium on Nonlinear Theory and its Applications*, pp. 923-927 (1998).
- [20] 中谷智広, 奥乃博, 川端豪: 音環境理解のためのマルチエージェントによる調波構造ストリームの分離, *人工知能学会誌*, Vol. 10, No. 2 pp. 232-241 (Mar. 1995).
- [21] Nakadai, K., Lourens, T., Okuno, H.G., and Kitano, H.: Active Audition for Humanoid, *Proceedings of AAI-2000*, pp. 832-839 (Aug. 2000).
- [22] Nakadai, K., Hidai, K., Mizoguchi, H., Okuno, H., and Kitano, H.: Real-time Auditory and Visual Multiple-object Tracking for Robots, *Proceedings of IJCAI-01*, pp. 1424-1432 (2001).
- [23] 奥乃博: 音環境理解 - 混合音の認識を目指して, *情報処理*, Vol. 40, No. 10, pp. 1096-1101 (Oct. 2000).
- [24] Okuno, H.G., Nakadai, K., Lourens, T., and Kitano, H.: Separating Three Simultaneous Speeches with Two Microphones by Integrating Auditory and Visual Processing, *Proceedings of Eurospeech 2001*, pp. 2643-2646, ESCA (Sep. 2001).
- [25] Rosenthal, D. and Okuno, H. G. (Eds.): *Computational Auditory Scene Analysis*, Lawrence Erlbaum Associates (1998).
- [26] 田中穂積: 言語理解と行動制御に関する研究 (課題番号 12NP9201) 平成 12 年度科学研究費補助金 (創成的基礎研究) (Mar. 2001).
- [27] Wang, F., Takeuchi, Y., Ohnishi, N., and Sugie, N.: A Mobile Robot with Active Localization and Discrimination of a Sound Source, *Journal of Robotic Society of Japan*, Vol. 15(2), pp. 61-67 (1997).

## 3.2 エージェントの研究動向

報告者: 大須賀昭彦委員

### 3.2.1 エージェントとは

現状、「エージェント」という言葉はさまざまな定義で用いられている。ここでは、プログラミング言語としてのエージェントをオブジェクト指向言語の次の技術と位置づけて考える[1][2]。

これまで、プログラミングパラダイムは主にアセンブラ言語、手続き型言語、オブジェクト指向言語という流れで変遷してきた。アセンブラは計算機のアーキテクチャを強く意識して計算機の挙動を記述するものであったが、手続き型言語では処理を主体にプログラムが書けるようになり、オブジェクト指向言語では計算対象 (object) を主体とする記述が可能となった。これは、手続き型言語においては計算機のアーキテクチャを隠蔽し、オブジェクト指向言語においてはデータ抽象の考え方によってデータ実装の詳細を隠蔽した抽象化の歴史とみることができる。この抽象化は、計算の世界と現実の世界を近づける方向で続けられてきた。オブジェクトの次に位置づけられるエージェントもこの抽象化の延長上にあると考えられる。さらなる抽象化によって、計算の世界と現実の世界をより近づけるのがエージェント指向パラダイムのねらいである (図 3.2-1)。

ここでは、オブジェクト指向にどのような特性が加えられ、それによって何が抽象化されたかという観点でエージェントを分類し、分類ごとの研究動向や事例を概観する。

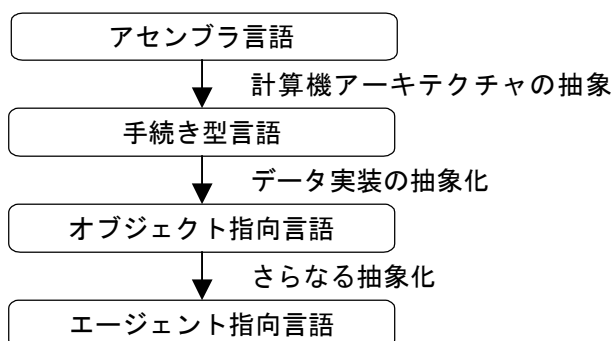


図 3.2-1 プログラミングパラダイムの変遷

表 3.2-1 代表的なモバイルエージェントシステム

|                                     |
|-------------------------------------|
| Telescript、Odyssey (General Magic)  |
| Aglets (日本アイ・ビー・エム)                 |
| Voyager (Object Space)              |
| MOA (The Open Group)                |
| Mole (Stuttgart 大学)                 |
| Kafka、April (富士通)                   |
| Concordia (三菱電機)                    |
| Mobidget (日本電気)                     |
| Flage (情報処理振興事業協会)                  |
| Grasshopper (IKV++)                 |
| Ara (Kaiserslautern 大学)             |
| Agent-Tel (Dartmouth 大学)            |
| Safe-Tel (Sun)                      |
| Mobile Space、Agent Space (お茶の水女子大学) |
| Planet (筑波大学)                       |
| 共生・寄生エージェントモデル (慶應義塾大学)             |
| Plangent、Bee-gent (東芝)              |

### 3.2.2 エージェントの移動性

オブジェクトに加えられた特性として、最初にエージェントの移動性を挙げる。ここでは、移動性を持つエージェントをモバイルエージェントと呼ぶ。モバイルエージェントにおいては、分散処理における通信の概念が抽象化されたとみることができる。モバイルエージェントとして多数のシステムが実装されており、Stuttgart 大学教授の Hohl が作成した The Mobile Agent List[3]にも 70 以上のシステムが登録されている。代表的なモバイルエージェントシステムを表 3.2-1 に挙げる。

モバイルエージェントの実装には、独自のスクリプト言語を定義してインタプリタと共に動作させるもの、ライブラリを拡張して既存の言語 (多くは Java) に移動性を持たせたもの、言語処理系 (Java の場合は JVM) に手を加えて移動性を持たせたものなどがある。

モバイルエージェントの利点には、ネットワークバンド幅の有効活用が挙げられる。モ

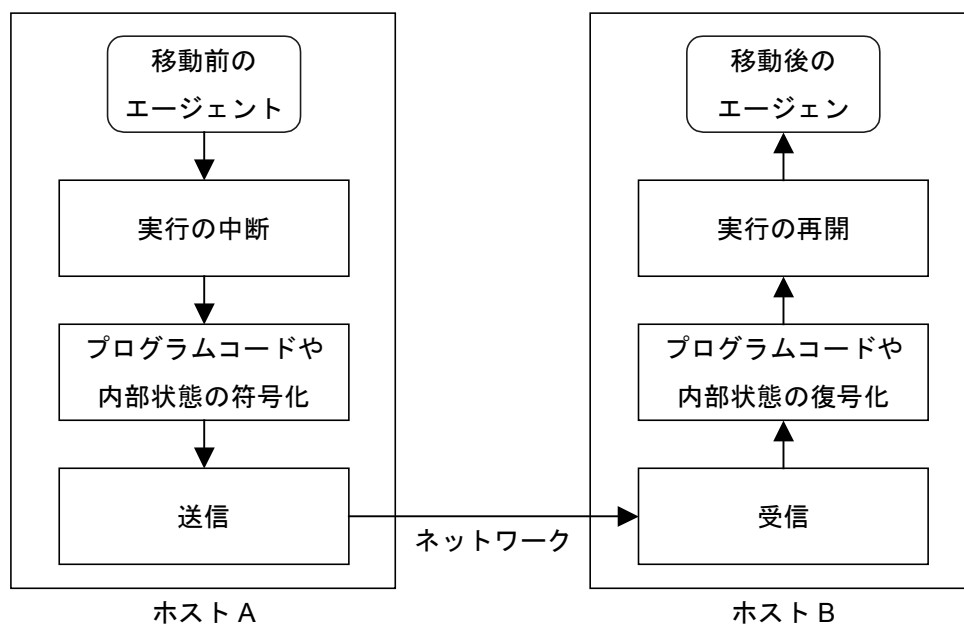


図 3.2-2 モバイルエージェントの移動の仕組み

モバイルエージェントは移動した先の資源にローカルにアクセスできるので、大量のデータに対して高速にローカルアクセスを行ったあと、得られた結果のみを移動して持ち帰るといった動作が可能である。必要に応じて移動先のマシン環境の情報を活用したり、負荷の小さいマシンに移動して負荷分散を行ったりもできる。無線アクセスのように通信が切れやすい環境において、通信切断後も処理を継続できる点では特にモバイルエージェントが有効と言える。

モバイルエージェントの移動の仕組みを図 3.2-2 に示す。この図は、モバイルエージェントがホスト A からホスト B に移動する流れを示している。まず、ホスト A で移動前のエージェントの実行が中断される。そして、エージェントのプログラムコードや内部状態が符号化（シリアライズ）される。符号化の結果は通常の方法と同じようにしてホスト A からホスト B に送信される。ホスト B では、受信したデータに対してホスト A と対称の操作を行う。つまり、復号化（デシリアライズ）によってプログラムコードと内部状態を復元し、エージェントの実行を再開させる。この一連の流れが完了すると、ホスト A にいたモバイルエージェントがホスト B に移動し、処理を続けることになる。

モバイルエージェントの実行の中断/再開、プログラムコードや内部状態の符号化/復号化、送信/受信などは、モバイルエージェントのプラットフォームと呼ばれるミドルウェアによって行われる。内部状態の符号化/復号化において、プログラムコードとインスタンス変数のみを対象とする場合を、「弱い移動」と言う。これに対して、プログラムコード、インスタンス変数に加えてプログラムカウンタまで対象とするものを「強い移動」という。ここで強調すべき点は、モバイルエージェントのコーディングの簡単さである。例えば、



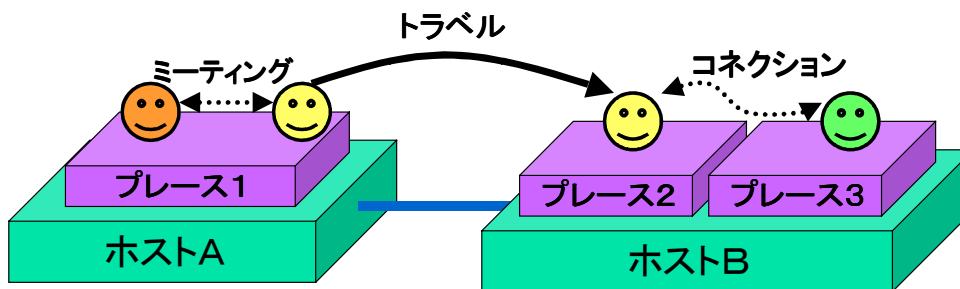


図 3.2-3 Telescript

ホスト A で処理 1、2 を行って、次にホスト B に移動して処理 3、4 を行うとする。このときのコーディングは以下のような簡単な記述で済む。

```

処理 1;
処理 2;
go(ホスト B);
処理 3;
処理 4

```

このため、さまざまな非同期分散処理が必要となる場面でもモバイルエージェントを活用すれば、簡明な記述の集合によって全体の処理を定義できる可能性がある。

モバイルエージェントの例として、General Magic が開発した Telescript を紹介する(図 3.2-3)。Telescript においては、移動するプロセスをエージェント、移動しないプロセスをプレイスと呼ぶ。エージェントがプレイス間を渡り歩き、同じプレイス内にいるエージェントと出会って(ミーティングして)相手のプロシージャを呼び出したり、他のプレイスにいるエージェントと通信(コネクション)したりするモデルが、独自のスクリプト言語上に実装されている。エージェントの移動には、強い移動が採用されている。各エージェントは権限を持っており、これにより資源へのアクセス制御も可能である。今日のモバイルエージェントで用いられる概念の多くは、Telescript において実現されたものである[4]。

### 3.2.3 エージェントの協調性

オブジェクトに加えられた特性として、次にエージェントの協調性を挙げる。ここでは、協調性を持つエージェントを協調エージェントと呼ぶ。協調性の導入は、分散処理における異種性、位置、通信手順などの抽象化につながるものである。代表的な協調エージェントシステムを表 3.2-2 に示す。

協調エージェント技術の利点は、エージェント間のコミュニケーションの意味や意図のレベルまで踏み込んだサービスを提供して、エージェント同士の動的で緩やかな連携を可能にしたこと（疎結合モデル）や、あらかじめ連携することを意図して設計していなかった異種システムを連携動作させられるようにしたこと（相互運用性）などである。協調エージェントの研究は分散オブジェクトの研究と深くかかわっている。分散オブジェクトには、もともと分散している（非オブジェクト）システムのそれぞれをオブジェクト化する垂直分散アプローチと、非分散のオブジェクトを（分散化を意識させずに）分散させる水平分散のアプローチとがあるが、どちらにおいても透過性の概念が重要である。分散オブジェクトにかかわる透過性には、位置透過性、異種透過性、障害透過性、規模透過性、移動透過性などが挙げられる。透過性とはある種の抽象化であり、その実現には数々の協調エージェントの技術やコンセプトが応用されている。例えば、分散オブジェクトにおける位置透過性の追求には、ネーミングサービス、トレーディングサービス、メディエータなどの技術が使われるが、これにはファシリテータの考え方がかかわっている。UDDI (Universal Description, Discovery and Integration) [5]における Web サービスのためのディレクトリ機能は、次に述べるファシリテータによるリコメンドパターンへの応用と見ることができる。

表 3.2-2 代表的な協調エージェントシステム

|                                                                                                                          |
|--------------------------------------------------------------------------------------------------------------------------|
| AOP、AGENT0 (Shoham)<br>Open Agent Architecture (SRI International)<br>AgentPro (富士通)<br>Linda (Yale 大学)<br>Bee-gent (東芝) |
|--------------------------------------------------------------------------------------------------------------------------|

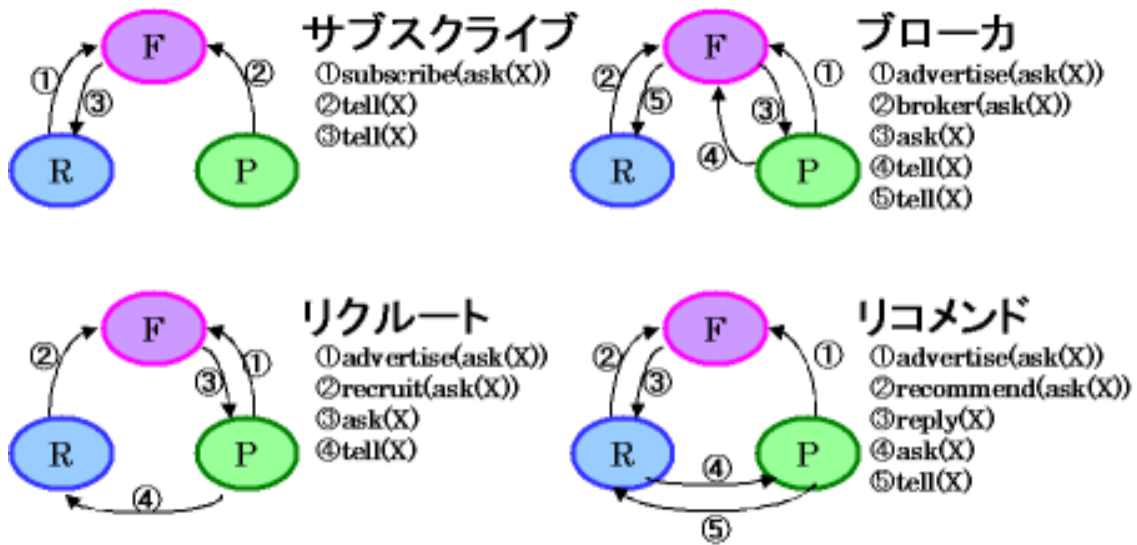


図 3.2-4 ファシリテータによるサービス利用パターン

ファシリテータは、エージェント間の協調動作させるための仲介機能をもったエージェントである。ファシリテータによるサービス利用の代表的なパターンを図 3.2-4 に示す。F はファシリテータを、P はサービスプロバイダを、R はサービス利用者を表す。障害透過性の追求には、オブジェクトの多重化や冗長待機などの手法が用いられるが、共有データ空間、タプルスペース、黒板モデルなどを介してオブジェクト同士が間接的なコミュニケーションを行う場面にも、協調エージェントの技術が応用され、効果を発揮している[6]。

協調エージェントの例として、東芝の Bee-gent を紹介する[7]。Bee-gent は、既存アプリケーションをエージェント化するエージェントラッパーと、アプリケーション間の連携手続きを管理するモバイル仲介エージェントから構成される。これらの仕組みにより、デ

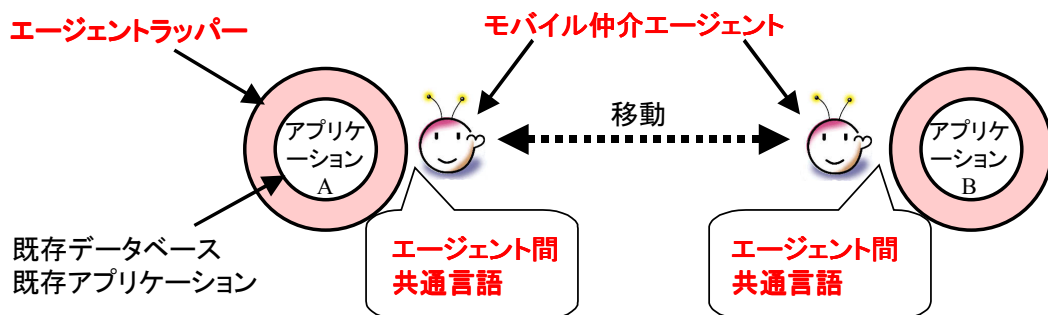


図 3.2-5 Bee-gent

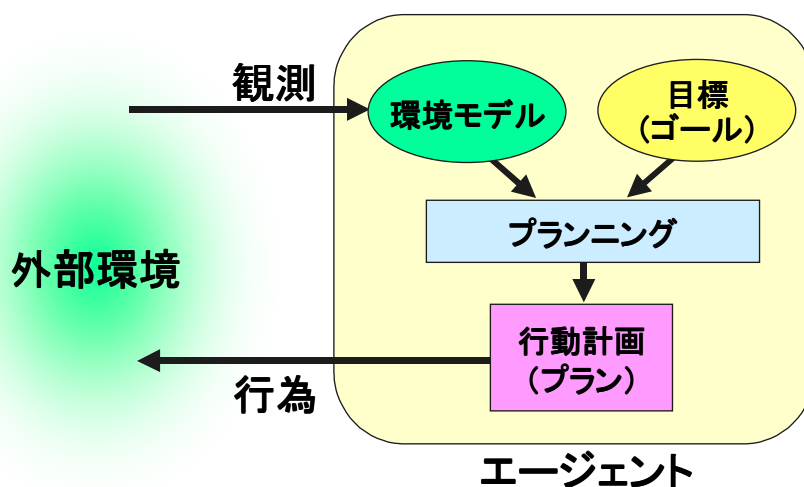


図 3.2-6 プランニングエージェントの構造

データベースや Web サービス、レガシーシステムなどの各種ソフトウェアを柔軟に接続して活用することを目指している。Bee-gent の概念図を図 3.2-5 に示す。モバイル仲介エージェントは、接続対象のアプリケーションが存在する場所へ移動し、エージェントラッパーと対話しながら要求を伝える。エージェントラッパーはアプリケーションの状態を管理し、必要に応じてアプリケーション処理を起動することによって、モバイル仲介エージェントからの要求にこたえる。エージェント間の対話には、FIPA 準拠の ACL[8]が用いられる。ACL は、会話の意図を表す通信行為 (Communicative Act) を記述することが特徴であり、これにより受け手側は意図を解釈した上で自分なりの処理手順を決定できる。Bee-gent では、FIPA が策定する ACL を XML 形式で表現した XML/ACL を採用している。XML 形式の採用によってインターネットとの親和性が格段に向上し、企業間での XML/ACL の交換も可能となっている。モバイル仲介エージェントやエージェントラッパーの動作を定義するための専用のビジュアル開発環境も提供される[9][10][11]。

### 3.2.4 エージェントの自律性

オブジェクトに加えられた特性として、次にエージェントの自律性を挙げる。ここでは、自律性を持つエージェントを自律エージェントと呼ぶ。自律性の導入は、処理の手順や手続きの抽象化にねらいがある。

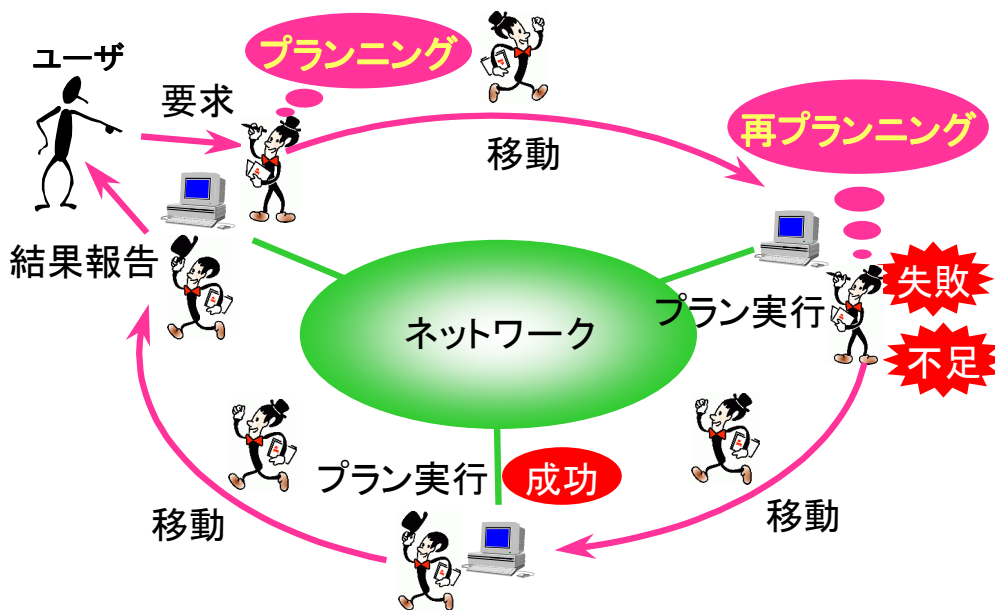


図 3.2-7 Plangent

一般にオブジェクトは受動的であり、エージェントは能動的と言われる。しかし、分散オブジェクトなどにおいては作りこみによってオブジェクトの自律性・能動性を実現している例が多く、外部から観察しただけでは両者の違いは明確ではない。自律エージェントのための言語においては、自律性が言語処理系において支援される。これを実現するために、多くの自律エージェントの処理系ではプランニング技術を採用している。一般的なプランニングエージェントの構造を図 3.2-6 に示す。プランニングには熟考プランニング、反応プランニング、ハイブリッドプランニングなどがある[12]。

自律エージェントの利点は、従来の手続き型言語やオブジェクト指向言語では書ききれないような処理、特に柔軟性を必要とする処理が書きやすい点にある。1対1ビジネスにおけるパーソナライズ処理や、状況に応じてダイナミックに処理内容を変更する状況依存処理などに対するニーズは高く、これらの問題に自律エージェントを応用することが期待されている。また、予期せぬ事態に対して柔軟に対処できる点も自律エージェントの利点である。例えばネットワーク上の処理が何らかの原因で中断された場合にも、自律エージェントは自分で代替手段をさがして処理を継続する。

自律エージェントの例として、東芝の Plangent を紹介する[13]。Plangent は、プランニング機能を持つモバイルエージェントが、目標達成の手段を自分で考えながら、ネットワーク上を自律的に動き回って処理を行う点に特徴がある。ユーザからの要求を受けとった Plangent エージェントは以下のように行動する (図 3.2-7)。

表 3.2-3 その他のエージェントのシステム

|                                                                                                                                                                                                                                                                                                                                                                                   |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>・ 環境適応性・順応性</li> </ul> <p>Flage (情報処理振興事業協会)</p> <ul style="list-style-type: none"> <li>・ 有機的プログラミング</li> </ul> <p>Gaea (産業技術総合研究所)</p> <ul style="list-style-type: none"> <li>・ 成長するエージェント</li> </ul> <p>Flage、Music、Morphe</p> <ul style="list-style-type: none"> <li>・ 計算の場 (フィールド)、セル</li> </ul> <p>Flage、Gaea、Cellua、Kamui、Kemari</p> |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

- (1) どこで何をするかといった行動計画をプランニングによって立案する。
- (2) この行動計画に基づいて必要な情報やサービスのある場所まで移動する。
- (3) 移動先の情報やサービスを活用して処理を実行する。
- (4) 予期せぬ事態などによって計画の実行が失敗した場合、再プランニングによって状況に合った行動計画を作りなおす。
- (5) 新たな行動計画に基づいて移動と実行を行う。
- (6) 目標が達成されるまでこれを繰り返し、最後にユーザのところへ戻って結果を報告する。

このような Plangent には多くの利点がある。例えば、ユーザが細かい作業手順を指示する必要がないこと、予期せぬネットワーク環境の変化に対してエージェントが柔軟に対処すること、個人の好みや過去の事例を学習してエージェントが成長することなどは大きな利点である。また、エージェントの移動時以外は通信回線が切断されてもよいため、今後増加する携帯機器向けソフトウェアとしても適した特性を持つ[13][14][15]。

### 3.2.5 その他の特性

これまでに、移動性、協調性、自律性の3特性をオブジェクトがエージェントとなる際に加えられた主要なものとして挙げた。ここではその他の試みを簡単に紹介する(表3.2-3)。エージェントに環境適応性や順応性を持たせる試みに Flage (情報処理振興事業協会) の提案がある。また、環境の変化を陽に扱う有機的プログラミング言語として Gaea

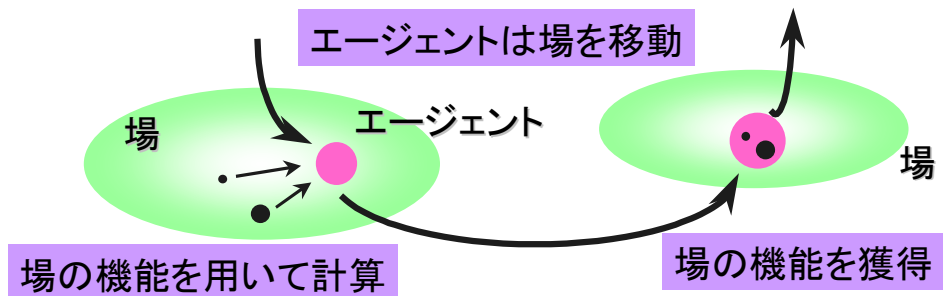


図 3.2-8 Flage

(産業技術総合研究所) が提案されている。エージェントの成長をコンセプトとする言語には、Flage のほかに Music、Morphe があり、計算の場(フィールド) やセルを導入した言語には、Flage、Gaea のほかに Cellua、Kamui、Kemari などがある。ほかにはオブジェクトの通信方式を拡張したり、オブジェクトに永続性を持たせたり、リフレクションを導入してエージェント化をはかった言語の提案がある。

ここではこれらの言語の中で Flage を紹介する。Flage は環境変化に対する適応、エージェントの成長、場の計算といったコンセプトの実現を目的としたエージェント言語である(図 3.2-8)。Flage のモバイルエージェントが場に入ると、そこに置かれている機能を使って計算を行うことが可能となる。これにより、エージェントは移動先の環境に合った計算方法を獲得し、その場で実行する。また、エージェントは場に置かれている計算方法を自分のものとして取り込み、場を去った以降も持ち歩くことができる。この機能をうまく活用すれば、エージェントが場を渡り歩きながら次々と新機能を獲得し、エージェントが成長するモデルを実現できる[1][17]。

#### 参考文献

- [1] 本位田真一, 飯島正, 大須賀昭彦: エージェント技術, 共立出版 (1999).
- [2] 本位田真一, 大須賀昭彦: オブジェクト指向からエージェント指向へ, ソフトバンク (2001)
- [3] <http://www.informatik.uni-stuttgart.de/ipvr/vs/projekte/mole/mal/mal.html>
- [4] 山崎重一郎, 津田宏 (編訳): Telescript 言語入門, アスキー出版局 (1996).
- [5] <http://www.uddi.org/>
- [6] 河込和宏, 中村秀男, 大野邦夫, 飯島正: 分散オブジェクトコンピューティング, 共立出版 (1999).

- [7] <http://www2.toshiba.co.jp/beegent/>
- [8] <http://www.fipa.org>
- [9] 川村隆浩, 長谷川哲夫, 大須賀昭彦, 本位田真一: Bee-gent: 移動型仲介エージェントによる既存システムの柔軟な活用を目的としたマルチエージェントフレームワーク, 信学論, Vol. J82-D-I, No. 9, pp. 1165-1180 (1999).
- [10] Tahara Y., Ohsuga A., and Honiden S.: Agent System Development Method Based on Agent Patterns, Proceedings of 20th International Conference on Software Engineering, pp. 356-367 (1999).
- [11] 田原康之, 大須賀昭彦, 本位田真一: ビヘイビアパターンに基づくモバイルエージェントシステム開発手法, 情処論文誌, Vol. 40, No. 12, pp. 4319-4332 (1999).
- [12] Russell, S. and Norving, P., 古川康一 (監訳): エージェントアプローチ - 人工知能, 共立出版 (1997).
- [13] <http://www2.toshiba.co.jp/plangent/>
- [14] Ohsuga A., Nagai Y., Irie Y., Hattori M., and Honiden S.: PLANGENT: An Approach to Making Mobile Agents Intelligent, IEEE Internet Computing, Vol. 1, No. 4, pp. 50-57 (1997).
- [15] 大須賀昭彦: プランニングモバイルエージェント, bit, Vol. 31, No. 3, pp. 88-95 (1999).
- [16] 大須賀昭彦ほか: 知的モバイルエージェント Plangent プログラミング, bit, Vol. 31, No. 8-10 (1999).
- [17] Kumeno F., Ohsuga A., and Honiden S.: Flage: A Programming Language for Adaptive Software, IEICE Trans. Inf. and Syst., Vol. E81-D, No. 12, pp. 1394-1403 (1998).



### 3.3 人間中心の情報環境 —感性的相互作用の視点から—

報告者: 加藤俊一委員

#### 3.3.1 はじめに —情報システムにおける「人間」の位置づけ—

ごく普通の日常生活の環境を情報化して「人間中心」「人にやさしい」情報環境を作るためには、人間をどのようにモデル化し、人間に対してどのような接し方をするインタラクション環境を構築する必要があるのだろうか？

人間の知能の知的な側面を工学的に実現することを目標として、人工知能・パターン認識理解技術の研究は 1960 年代から始まっている。今日では我々は、データベースシステム技術、エキスパートシステム技術、画像・物体認識技術の大きな成果の恩恵を受けており、これらの技術なしには高度情報社会を営むことは不可能となっている。しかし、人間そのものとかかわりという点では、どうだろうか？ 電子図書館、発電所運転管理システム、文字・物体認識システムなどに代表されるデータベース、人工知能、パターン認識理解技術は、人間（人間の知能）を代行する機器・システムを目指したもので、必ずしも「生きた人間」との共存を目的とはしていない。

1990 年代から、人間とロボットなどの情報機器がともに働く環境の実現に向けて、景観画像からの人間の認識、顔の認識などの技術開発も進んできた。しかし、人間の位置や姿勢、行動を認識し、また、個人の識別に結びつけるこれらの技術も、人間とかかわりという点では、人間をモノ（物理的な実在）として計測し、認識し、働き掛けるという考えにとどまっている。

2000 年代に入って、新しいセンサ技術やネットワーク技術と人工知能・パターン認識理解・ロボット技術を結びつけて、人間の体温や呼吸、筋肉の緊張の度合いや、皮膚感覚の快適性を計測する技術開発が活発になってきている（例：経済産業省や文部科学省関連のデジタルヒューマンプロジェクトなど）。人間の状態を計測し、人間に快適な環境を実現しようとする試みは、「人間中心」で「人にやさしい」情報システムの実現にも大きく貢献するものと期待される。しかしながら実際の研究開発は、人間を生物（生理的な実在）として計測し、認識し、働き掛けるという範囲にとどまっているようである（デジタルヒューマンの基本的なコンセプトは、1990 年代に実施された通商産業省工業技術院の人間感覚計測プロジェクトで行われた人間の生理的な状態を計測する技術に、最新のセンサ技術、ネットワーク技術、ロボット技術を融合させたものともいえよう）。

人間を情報技術の面から「感情を持った生物」としてとらえて、その生理的・心理的・認知的な特性をモデル化する試みとしては、1990 年代後半に実施された通商産業省工業技

術院のヒューマンメディアプロジェクトを挙げることができる。感性のモデル化や感性データベースなどの感性情報処理技術と、バーチャルリアリティを含むヒューマンインタフェース技術を融合させるコンセプトと試みは、今なお、今日的な技術課題であると思われる。

本報告では、「人間の感性」という切り口から、「人間中心」「人にやさしい」情報システムの実現に重要な役割を果たす、人間と情報機器・情報システムとのインタラクションに焦点を絞って、技術課題の分析と最近の研究開発の話題を紹介する。

#### 3.3.2 インタラクションの形態

情報提供サービスにおける利用者とデータベースとのインタラクションのモデル化を行うために、ここでは、インタラクションの形態と、インタラクションを構成する要素、システムが提供すべき機能を、次の 7 つの視点から分析してみよう：主体（Who）、対象（What）、状況（When）、場所（Where）、理由（Why）、目的（What for）、行為（How）。形態や要素を表 3.3-1 に要約して示す。

##### (a) Who（だれが）

インタラクションの主体となるもの。情報を評価する規準を与えるもの。

今日までのセールスプロモーション（情報提供サービスを含む）では、「20 歳代の女性」や「JPOP の好きな人」のような利用者グループ（消費者グループ）を定めると、市場調査（マーケティング）などを利用して、マスとしての平均的な趣味・嗜好・ライフスタイルを推定し、これにあわせた商品企画・開発・提供を行ってきた。

しかし、利用者一人ひとりに、より適合した高品質な商品（情報提供サービス）を提供するためには、利用者一人ひとりの知識・趣味・嗜好などの主観的な特性（感性）を的確にモデル化（感性のモデル化）するとともに、これにマッチした個別的な商品（情報提供サービス）を提供する必要がある。また、「特定のだれ」がインタラクションの主体であるかを識別し、それに合わせてインタラクションを制御することが重要である。

このような感性のモデル化を利用すると、「20 歳代の女性」や「JPOP の好きな人」のように、主観的な特性のいくつかが共通した利用者グループの感性モデルを集約化することができれば、これらの利用者グループに共通的・平均的な感性のモデルを精度良く求められる。さらには、時々刻々変化する利用者の趣味・嗜好・ライフスタイルを、的確に市場調査（マーケティング）することも可能となる。

また、街に設置された情報キオスクや、インターネット上のサーチエンジンのように、不特定多数の利用者への情報提供サービスを行う場合は、現在利用している利用者の主観的な特性がどの利用者グループに近いかを判定するとともに、インタラクションを通じてその利用者の主観的な特性により近づけていくことも必要となる。

表 3.3-1 インタラクションの形態とインタラクションを構成する要素

| 視点                  | 要素               | 必要となる要素の説明                         |
|---------------------|------------------|------------------------------------|
| Who                 | 主体               | 特定個人の利用者（例：加藤）                     |
|                     |                  | 共通点のある利用者群（例：20歳代の女性）              |
|                     |                  | 不特定多数（例：だれでも）                      |
|                     |                  | 未知の利用者（例：新規の顧客）                    |
| Why                 | 判断理由             | 知覚感性の階層的モデル                        |
|                     | 判断規準             | 物理的レベル（例：ほとんど同じ）                   |
|                     |                  | 生理的レベル（例：同じに見える）                   |
|                     |                  | 心理的レベル（例：なんとなく似ている）                |
| 認知的レベル（例：「さわやか」な印象） |                  |                                    |
| What                | コンテンツ            | マルチメディアコンテンツ（例：画像）                 |
|                     | 対象               | コンテンツの物理的特徴量（例：物理・生理レベル）           |
|                     |                  | コンテンツの主観的特徴量（例：心理レベル）              |
|                     |                  | コンテンツの統合的特徴量（例：認知レベル）              |
|                     |                  | 教師データ                              |
|                     | 操作の履歴（例：加藤の操作ログ） |                                    |
| When                | 時間               | 日時・時刻                              |
|                     | 状況               | 繰り返しの周期・頻度                         |
|                     |                  | 偶然・ランダム                            |
| Where               | 場所               | 物理的空間（例：現在地、VRによる仮想空間）             |
|                     | 状況               | 情空間（例：類似度空間、オントロジー）                |
|                     |                  | ユビキタス空間（例：モバイル + 情報家電ネット）          |
| What for            | タスク              | 正確な検索（例：このデータの所在）                  |
|                     | 目的               | あいまいな検索（例：こんな感じのデータはあるか）           |
|                     |                  | インタラクションそのもの（例：アミューズメント）           |
| How                 | タスク<br>方法・手段     | 探す、選ぶ、比較する、場所を変える、可視化する、抽象化する、提案する |

(b) Why（理由で）

情報を取捨選択する際の判断の理由、規準。

マルチメディアコンテンツの物理的特徴から始まって、これを生理的、心理的、認知的なレベルで知覚していく過程で、知覚にかかる感性を工学的にモデル化することを考える。

感性のモデル化では、次の4つの階層を考えることができる。

(i) 物理的レベル

コンテンツ自身の構造的・統計的な性質などの客観的な情報として特徴が記述される。パターンマッチング的な例示検索は、物理的レベルの記述の一致度を類似度の判定規準として実現できる。

(ii) 生理的レベル

コンテンツが感覚器で受容されてから、網膜から一次視覚野などまでの特徴抽出機構で得られる客観的な（個人差の少ない）特徴が記述される。より柔軟な例示検索は、生理的レベルの記述で構成される多次元ベクトル空間上での距離を類似度の判定規準として実現できる。

(iii) 心理的レベル

一人ひとりの利用者が教示するグループ分けなどの主観的な類似度の規準を、統計的な教示学習などにより構成する。これは、生理的にはほぼ同じ構造の感覚器（眼など）から得られる生理的な特徴量に対して、利用者ごとに異なった重みづけで評価しているためと考えられる。類似検索は、心理的レベルの記述で構成される多次元ベクトル空間上での距離を類似度の判定規準として実現できる。

(iv) 認知的レベル

主観的なグループ分けと主観的なイメージ語による表現が結びつけられていると考えられる。感性検索は、マルチメディアコンテンツの主観的特徴と、認知的レベルの言葉による記述を統合した多次元ベクトル空間上での距離を、類似度の判定規準として実現できる。

(c) What（何を）

情報処理やインタラクションの対象（コンテンツ）。  
マルチメディアコンテンツを対象とした検索では、

- (i) 例示検索: 利用者が実際例を提示して、よく似た情報を検索する。
- (ii) 類似検索: 一人ひとりで異なる主観的な基準に照らして、類似した情報を検索する。
- (iii) 感性検索: 主観的で漠然としたイメージ語による表現から、それにふさわしい情報を検索する。

が基本的なサービスの形態となる。

マルチメディアコンテンツを対象に、例示検索のような情報提供サービスを受けるためには、コンテンツそのものだけでなく、コンテンツを客観的に計測して得られる物理的な特徴量をインタラクションの過程で扱う必要がある。

また、類似検索や感性検索のような、主観的な規準に基づく情報提供サービスを受けるためには、個々の利用者がコンテンツをどのように主観的に解釈するかを（客観的に）

観測した主観的な特徴量をも扱う必要がある。このような観測は、ある程度の時間をかけて、データ量を集めて、教師あり学習の形で統計的に行われることが多い。個々の利用者が、どのようなコンテンツを対象に、どのような操作を行い、また、どのような判断をしたかの操作履歴は、主観的な規準をモデル化する上で有用な情報である。

(d) When (いつ)

日時・時刻、期間、周期・頻度など、時間的な状況。

インタラクションが行われる日時・時刻や、繰り返しの周期・頻度などは、インタラクションの動的な特徴を表す。長くは、1年や季節、1か月をサイクルとしたインタラクション、短くは、1時間、1分を単位としたインタラクションもある。

一方、偶然性に影響されて、ランダムに発生するインタラクションも考えられる。

(e) Where (どこで)

インタラクションがどこでどのような状況で行われるかは、インタラクションの空間的な特徴を表す。今日的なネットワーク環境では、次の3種類の空間を考えることができる。

(i) 物理的空間

現実世界、あるいは、現実世界に成り立つ種々の物理法則などを忠実に計算機上でシミュレーション・再現した仮想空間をいう。現実世界と同様に、近いものとのインタラクションは強く、離れたものとのインタラクションは弱い。対象が微小に変化しても全体に与える影響は少ない。

また、物理的空間は、現実空間の代わり（シミュレーション、メタファ）として構成されることが多いため、仮想空間にリアリティを持たせることが重要となる。

(ii) 情報的空間

意味ネットワークとその上でのノード、World Wide Web とその上でのデータ、対象や規則の論理的な記述群などのように、抽象的に記述された情報・データ相互間の関係に意味のある空間を言う。物理的空間とは異なり、対象に関するわずかの記述の変化で、情報的空間の規則や関係が大きく変化する場合もある。

情報的空間は、情報・データ間の関係を理解するために構成されることが多いため、見かけのリアリティよりは、関係構造を明示する手法に工夫が必要となる。

(iii) ユビキタス的空間

現実世界に情報処理の要素が遍在し、ネットワーク化されている空間をいう。例えば、室内（= 現実世界）の機器や家具（= 情報家電）が相互にネットワーク化されているとともに、インターネット（= 情報世界）などと人間とが相互にイ

インタラクションできるポイントが遍在（あまねく存在）しているような空間をいう。

#### (f) What for（何のために）

情報処理やインタラクション（タスク）の目的。

何のために情報提供サービスを利用するのかによって、望ましいインタラクションのあり方も大きく変わる。

利用者が欲しいと希望している情報が、存在することが保証されていてこれを検索する場合は、場所を探し当てる過程が容易に行えるようにインタラクションを構成する必要がある。これに対して、希望している情報のイメージがあいまいで、適切な「解」が存在するかどうかさえ不明である場合には、許容される「近似解」を推定・判定しながら、同時に、場所を探し当てる過程が容易に行えるようにインタラクションを構成する必要がある。このような場合には、インタラクションの過程での操作の仕方や提示された情報のわかりやすさと共に、正確さ・効率がインタラクションのしやすさの重要な指標となる。

一方、情報を得ることが目的なのではなく、インタラクションそのものが目的となる場合もある。インターネット上のチャットや携帯電話のショートメールなどでは、利用者は、新しい情報を得ることを目的とするのではなく、会話そのもの、インタラクションそのものに、楽しさ・面白みを感じていると考えられる。このような場合には、操作の仕方の容易さよりも、インタラクションを楽しくするための演出の良し悪しが、インタラクションのしやすさの重要な指標となる。

#### (g) How（どのように）

情報処理やインタラクション（タスク）の方法・手段。

利用者と情報機器あるいはインターネット上で運用される情報提供サービスとが、どのようなスタイルのインタラクションを通じて、目的とする情報処理を進めていくのかは、ヒューマンインタフェースの具体的な設計指針と直結している。

例えば、インターネット上の未整理で巨大なデータベースから、漠然とした条件でしか記述できない情報を発見して検索する場合、以下のような操作が必要となる。

#### (i) 探す

インターネット上のさまざまなサイトに構築されたデータベースから検索する。検索の方法としては、旧来のキーワードによる検索の他、例示検索・類似検索・感性検索などの高度な検索方式も、単一のインタフェースから統合的に利用することが望ましい。

(ii) 選ぶ

検索された候補の中から選択する。候補の数が非常に多い場合、それをそのまま表示するだけでは、利用者の知覚能力を超えるため選べない。内容に則してグループ化して表示するなど、利用者に選ばせるための工夫が必要となる。

(iii) 比較する

検索された候補を相互に比較して、漠然とした条件を具体化・詳細化させるなど、次の検索の戦略を立てる必要もある。この場合も、候補の数が非常に多い場合、それをそのまま表示するだけでは、利用者の知覚能力を超えるため、適切に比較できない。内容に則してグループ化して表示するなど、利用者に比較させるための工夫が必要となる。

(iv) 場所を変える

インターネット上で検索する場所（サイト、サーバ）を切り替えることも必要となる。利用者の指示を受けて切り替える場合のほかに、利用者にこのような操作を意識させずに自動的にサイトを切り替えて検索し続けるなどのサービスが必要となる場合もある。

(v) 可視化・不可視化する

検索された候補の数が非常に多い場合、候補をそれらの間の関係・類似度などによりグループ化する必要がある。このような検索結果だけからは得られないデータ間の関係を、利用者が把握しやすいように明示的に示すなどの工夫も必要である。逆に、詳細すぎる情報・複雑すぎる関係を簡略化して表示するなどの工夫も必要である。

(vi) 抽象化・具象化する

大規模なデータベース内のデータや検索された大量のデータをそのままの形で表示するのではなく、意味的な関係、感性的な類似関係など、さまざまな情報から関係を抽象化して利用者に提示する必要がある。同様に、利用者が漠然と与えた条件や事例群の間の隠れた関係を発見して、これを明示・具体化して、より効果的な検索に役立てるなどの工夫も必要である。

(vii) 提案する

人間中心の情報環境・人にやさしい情報環境とは、「情報処理のすべてを自動化して、利用者は命令するだけ」という形態を意味するわけではない。利用者の発想やアイデアを刺激し、より高度な情報処理や情報検索を実現するために、システムの側から提案する仕組みも必要である。ただし、その提案が、利用者の独創性や発想の広がりを制限するものとならないような工夫も必要である。

#### 3.3.3 インタラクションのモデル化

表 3.3-1 を整理すると、以下に述べる 4 つの面（感性、コンテンツ、状況、タスク）からインタラクションの特徴をモデル化することができる。

##### 3.3.3.1 感性モデル

###### 3.3.3.1.1 感性のモデルとその役割

感性モデルは、インタラクションの過程で、だれが (Who)、あるいは、だれのために (For whom)、どのような判断基準 (Why) で情報の取捨選択を行うかを定めるとともに、その人の感性的な状態（内部状態）を記述するために必要となるモデルである。

本報告の範囲で考える「感性」とは、一人ひとりの利用者、あるいは、利用者のグループがマルチメディア情報を解釈（あるいは表出）する過程で示す主観的な特徴をいう。これを客観的に計測し、マルチメディアコンテンツの客観的・物理的な特徴との対応関係・相関関係を数理的な手法でモデル化して表現すれば、このモデルは、工学的にシミュレーション可能な感性のモデルと考えることができる。

一人ひとりの利用者は異なった感性を持っているので、情報機器やシステムから人間への情報の提示では、利用者の年齢や嗜好などに合わせて情報の提示の仕方や GUI などを工夫しなければならない。

また、利用者から情報機器やシステムへの命令においても、利用者の年齢や嗜好などの感性の違いに応じて、命令の解釈・操作の内容・具体的な情報処理を切り替えなければならない場合がある。例えば、表現上、たまたま同じ命令になったからといって、例えば、大人と幼児が同じ内容の処理を希望しているとは限らない。

例えば、通商産業省工業技術院のヒューマンメディアプロジェクトでは、利用者一人ひとりの感性のモデルを利用することで、漠然としたイメージ語をキーとしても、その利用者の感性にマッチした印象を与える画像（絵画）を検索できる、電子美術館 ART MUSEUM などの感性検索システムを試作した。電子美術館では、「さわやかな印象の絵画」という主観的な表現に対して、それぞれの利用者ごとに異なった画像の集合を候補として提示できる仕組みを実現している。システムは、利用者ごとに、情報検索の主観的な基準を統計的に学習して感性モデルを構築する。検索の際には、利用者がだれであるかを識別して、その利用者の感性モデルに基づいて情報の取捨選択の基準を選択し、適切な候補を検索する（図 3.3-1）。





絵をクリックすると大きな画像で見られます。

あなたの印象を学習させたい絵画については、チェックボタンをクリックした後、画面の下にある「感性を学習させる」ボタンをクリックして下さい。

特に学習させる必要がない場合は、Backボタンで戻り他の印象語での検索をお楽しみ下さい。

### Kaoriさんの感性モデルによる検索結果

ややナチュラル とてもさわやか



第1候補

monet.adresse.jpg  
( 1.9639 )

自信を持ってお答えします!



第2候補

monet.martin.jpg  
( 2.4234 )

少しだけ自信があります。



第3候補

pissarro.village-path.jpg  
( 2.4303 )

少しだけ自信があります。



第4候補

monet.wl-clouds.jpg  
( 4.1802 )

あまり自信がありません。



絵をクリックすると大きな画像で見られます。

あなたの印象を学習させたい絵画については、チェックボタンをクリックした後、画面の下にある「感性を学習させる」ボタンをクリックして下さい。

特に学習させる必要がない場合は、Backボタンで戻り他の印象語での検索をお楽しみ下さい。

### Toshiさんの感性モデルによる検索結果

ややナチュラル とてもさわやか



第1候補

monet.martin.jpg  
( 1.6390 )

自信を持ってお答えします!



第2候補

monet.wl-clouds.jpg  
( 2.0767 )

少しだけ自信があります。



第3候補

monet.coquelicots.jpg  
( 2.8467 )

少しだけ自信があります。



第4候補

r-airship.jpg  
( 4.3000 )

あまり自信がありません。

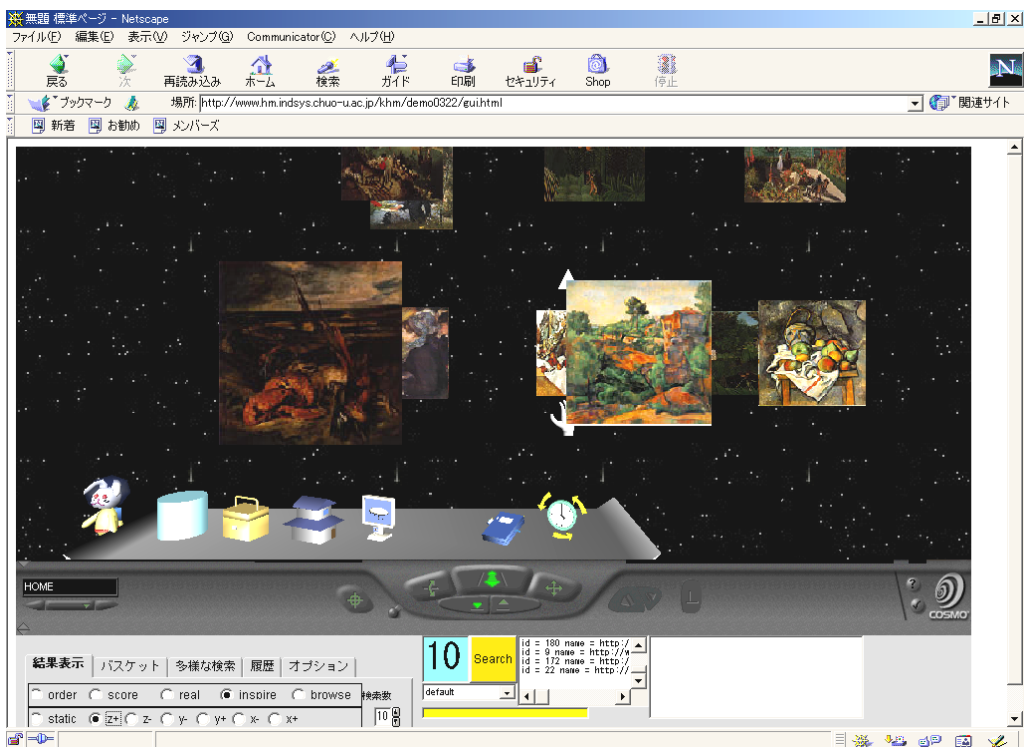
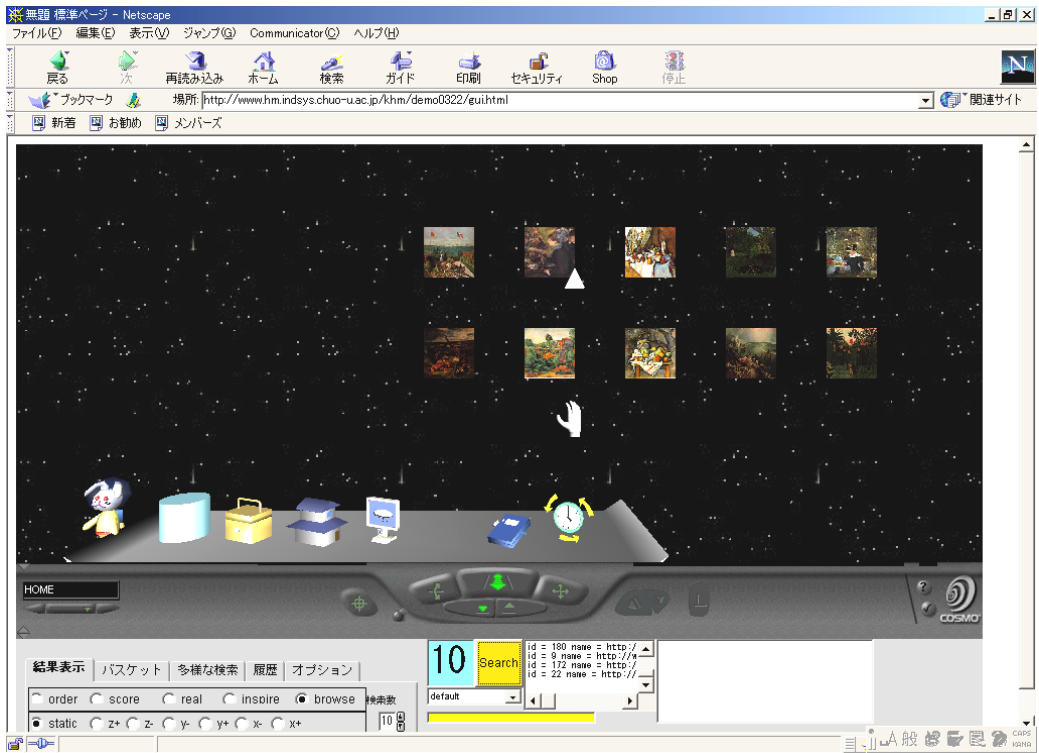


異なった利用者 (Kaori、Toshi) の感性のモデルを利用して検索

図 3.3-1 電子美術館における感性検索

### 3 研究開発の新しい展開と内外の動向

また、同プロジェクトで試作された工業デザイン支援システム **Kansei Factory** では、検索結果の表示法に関して、一人ひとりの利用者の技能・嗜好に応じて、検索結果を多様なスタイルで提示することができる（例：一覧表の形式での表示、類似度順に仮想空間上に一直線上に表示、類似度と仮想空間上の距離を対応させて 3 次元表示など）。このように、インターフェースでの表示法の制御にも感性モデルを利用することにより、利用者に直感的に検索結果を理解させることができる（図 3.3-2）。



利用者の好みに合わせて、一覧表示や類似度空間上での表示が可能

図 3.3-2 工業デザイン支援システム Kansei Factory での検索結果の表示法



前述のヒューマンメディアプロジェクトで開発された感性の工学的なモデルを図 3.3-3 に示す。今日では、感性のモデルの枠組みとして、大部分のアルゴリズム研究やシステムの試作・開発が、この枠組みに則して研究開発されている。

感性の構造に階層性を導入するのは、物理的な刺激（マルチメディア情報）に対応した感性から、抽象的主観的な概念（言語）に対応した感性までを、客観的に計測して結びつけるためである。

客観的に感性を計測し表現する基点として、客観的に存在するマルチメディア情報そのものを利用し、これと個人個人で解釈の幅がある抽象的主観的な概念（言語）を結びつけるようなモデル化を行う。このような階層的なモデル化によれば、例えば、文化圏の違いにより抽象的な概念語の解釈に大きな開きがあったとしても、客観的な存在であるマルチメディア情報を媒介に、人間の感性的な特性の違いを客観的に評価することが可能となる。また、新たに与えられた未知のマルチメディア情報であっても、モデルに基づいて解釈できる。

従来のキーワード・索引語による情報検索サービスでは、データベース作成者の主観によりマルチメディアコンテンツが分類整理され、必ずしも利用者の判断基準とは一致しない。また、一度作成された索引は、利用者の興味の変遷などには追従できない。

感性データベースシステムは、このような主観的な評価基準に自動的に対応しつつ、利用者からの感性的な検索要求に適合して適切なマルチメディア情報を提供するシステムをいう。

感性の生理学的・心理学的なメカニズムは複雑であるが、例えば視覚に関する感性の場合、次のように単純化してモデル化して、データベースシステムに応用することが可能である。

#### (i) 生理的レベルでの感性

人間の視覚系（初期視覚）にはエッジ・コントラスト・自己相関・空間周波数など種々の特徴抽出機構があることが知られている。これは視覚的情報がパラメータ化されて知覚されることに相当する。

#### (ii) 心理的レベルでの感性

人間は解剖学的には同一の目の構造を持ち、特徴抽出機構の働きも同様であるにもかかわらず、時として人により異なった判断をする。対象とする画像・映像を類別したり、画像間の類似度を与えたりする際などに、その差は顕著に現れる。これは、種々の特徴量の評価の重みが人によって異なるためと考えられる。

#### (iii) 認知的レベルでの感性

あるカテゴリーの視覚的情報をどう解釈するか、すなわち、どのような概念のラベル（印象語、イメージ語）を割り当てるかも時として人により異なった判断をする。個々人によって、種々の特徴量と概念の対応関係が異なると考えられる。

#### 3.3.3.2 コンテンツモデル

##### 3.3.3.2.1 マルチメディアコンテンツのモデルとその役割

コンテンツモデルは、インタラクションの過程で、何を（What）情報処理や検索の対象にしているのかを記述するために必要となるモデルである。利用者の感性によって情報処理の仕方や取捨選択の基準が変わるが、一方、利用者の感性の違いに関係なく、コンテンツ自身の性質から情報処理の仕方や取捨選択の基準が定まる場合もある。

マルチメディアデータベースでは、さまざまな種類のコンテンツが格納されている。したがって、インタラクションの対象となるコンテンツのメディア（映像、音など）の種類が多様であるため、それぞれのメディアに適した入出力手段や GUI を提供する必要がある。例えば、映像コンテンツの場合、一方的に再生するだけではなく、巻き戻しや早送り、一時停止などの手段が必要となり、絵画コンテンツでは拡大や縮小、サムネイル画像（縮小画像）の手段などが必要となる。

一方、同種のメディアであっても、検索結果のコンテンツが大量に存在する場合、検索結果が一度に表示されると人間の認知能力を超えてしまう。そのため、利用者は混乱してしまう。したがって、検索されたコンテンツの数や、コンテンツ間の関係に基づいて、検索結果を自動的に分類、組織化して、利用者にわかりやすく提示する必要がある。

例えば、中央大学の試作したガーデンコーディネータでは、開花時期や花色などの条件で多数検索された植物群を、データベースに格納された植物に関する専門的知識を多角的に利用することにより（注：利用者がわざわざ指示しているわけではない）、植物の生息条件に合わせて自動的にグループ化して提示し、利用者による庭の設計を支援している。

##### 3.3.3.2.2 コンテンツと操作履歴

高度なインタラクションを実現するためには、同じ利用者が、同種のコンテンツを対象に、同じような操作を行っても、操作の履歴と組み合わせて解釈すると、異なった処理方式やサービスを提供しなければならない場合がある。

例えば、デザインなどの仕上げの段階で、オブジェクトの形状や位置を微調整する場合など、操作の進行に応じて定型的な操作の既定値を変えるなどの調整も必要となる。具体的には、入力手段にマウスを用いて、工業デザインなどで形状の設計を行っている場合を考えてみよう。デザインの初期の段階では大雑把な移動距離の調節で充分だが、仕上げの段階では、微妙な位置調節の操作は難しい。そこで、適切なスケール変換を行うことにすれば、人間がマウスを動かしている距離は物理的には同じでも、デザイン支援システム上では異なった移動距離となり、操作性を格段に良くすることができる。

これは操作性の良いインタラクションを実現するのに重要である。

### 3.3.3.3 状況モデル

#### 3.3.3.3.1 状況モデルとその役割

情報システムが、適切な情報処理や人間とのインタラクションを進めていくためには、自律的に外界を観測する必要がある。このとき、観測の対象となる環境は、(a) 人間のいないモノだけからなる環境、(b) 生物としての人間の含まれた環境、(c) 人間が営む知的な活動を含んだ環境が考えられる。

状況モデルは、ある空間的な状況 (Where)、時間的な状況 (When) で、情報処理の対象や、利用者とシステムとのインタラクションがどのような状況で進行しているかを記述する。

同じ利用者によっても、利用者の置かれた状況が違えば、同一の操作などに対して異なったインタラクションの形態を提供する必要がある。

物理的空間・情動的空間・ユビキタス的空間を特徴づける性質としては、以下のような属性が考えられる。

##### (a) 物理的空間の状況

現実世界 (あるいはそれに等価な仮想世界) における事情の生じる時刻、周期、偶然、ランダムなどの時間的な属性や、位置、移動方向、ランダムなどの場所に関する属性がある。また、仮想空間を用いた GUI によって物理的世界を対象とする場合には、現実世界で人間が知覚するのと同じくらいの事象と、事象に関する物理法則、リアリティが求められる。

これらの情報を利用すると、例えば、現在の時刻に利用者のいる場所に基づいて、適切な道案内が行える。

##### (b) 情動的空間の状況

メディアの種類、情報空間上の位置、アクセスの数や頻度、データの量、情報の有効期限などに関する属性がある。情動的空間は、情報・データ間の関係を理解するために構成されることが多いため、論理的・抽象的な関係構造を明示する手法にわかりやすさが求められる。

この情報を利用すると、例えば、同じ商品を扱っている人気サイトの情報を紹介することができる。

##### (c) ユビキタス的空間の状況

時刻や位置などの実際に存在する空間の状況と、その空間に関連したモノ・事象に関する情報、アクセス数や頻度などの情動的空間の状況などが組み合わされた状況。

これらのさまざまな空間状況の情報をを用いることによって利用者が置かれている状況がわかり、利用者の状況に応じたインタラクションを進めていくことができる。

#### 3.3.3.3.2 ユビキタス空間と情報提供サービスの例

最近では携帯電話などのパーソナル情報機器の進歩とモバイルネットワーク、Bluetooth や IC メモリなどのチップを利用した近接したエリアにおける局所的なネットワーク、GPS などを利用した位置の計測技術などを組み合わせることにより、情報機器や情報サービスが実空間に遍在（あまねく存在）して埋め込まれているユビキタス空間が現実のものとなってきた。

このようなユビキタス的空間では、利用者やシステムの場所にとらわれない、サービスやインタラクションを提供することができる。

例えば、新宿近辺で現在の時刻で間に合う、人気商品のタイムセール情報を紹介することができる。実際に用いられている例としては、DoCoMo の PHS を利用した位置情報案内サービス（いまだこサービス）が挙げられる。これは、現在利用者がいる場所から、例えば、近いラーメン屋さんの情報を検索し、道案内してくれるサービスである。

#### 3.3.3.4 タスクモデル

利用者と情報機器あるいはインターネット上で運用される情報提供サービスとが、どのような文脈で、どのようなサービスを提供するかを定義するモデルである。

応用に依存する点が多いだけに、一般的な記述の枠組みを考えることは難しいが、さまざまな分野における「用語」「知識」を統合的に管理する技術に関しては、オントロジーベースなどの先行研究がある。

利用者の職業や知識的な背景に基づいて、インターネット上で検索する場所（サイト、サーバ）を切り替えることも必要となる。例えば、ある利用者がその業務に関して「水の利用法」に関する質問を発した場合、「業務の分野」「利用者の専門知識」「水をどのような目的に利用するのか」と無関係に「水の利用法」を検索すると、利用者に必要な情報は膨大なノイズの中に埋没してしまう。これは例えば、製薬分野、プラント、防災、造園、料理のようにまったく異なる分野においても「水」は共通に用いられる重要なキーワードだからである。それぞれの分野では役立つ情報であっても、当該の利用者の業務分野と異なる情報であれば、ノイズに等しいからである。したがって、システムは、当該の利用者の検索質問を「業務の分野」「利用者の専門知識」「水をどのような目的に利用するのか」という文脈のもとに理解し、適切なサーバのみから検索するなどの検索要求に自動的に再構成して処理する必要がある。

#### 3.3.4 事例紹介 インテリアコーディネート支援システム

以上のような4つの視点から、インタラクションの特徴をモデル化して記述することができる。利用者とシステムとのインタラクションのログを蓄積するとともに、現在のイン



タラクションの状態と比較することで、利用者に対し、次にどのようなインタラクションの環境を提供すれば有効であるかわかるようになる。これにより、インタラクションをスムーズに行うための戦略が立てられるようになる。

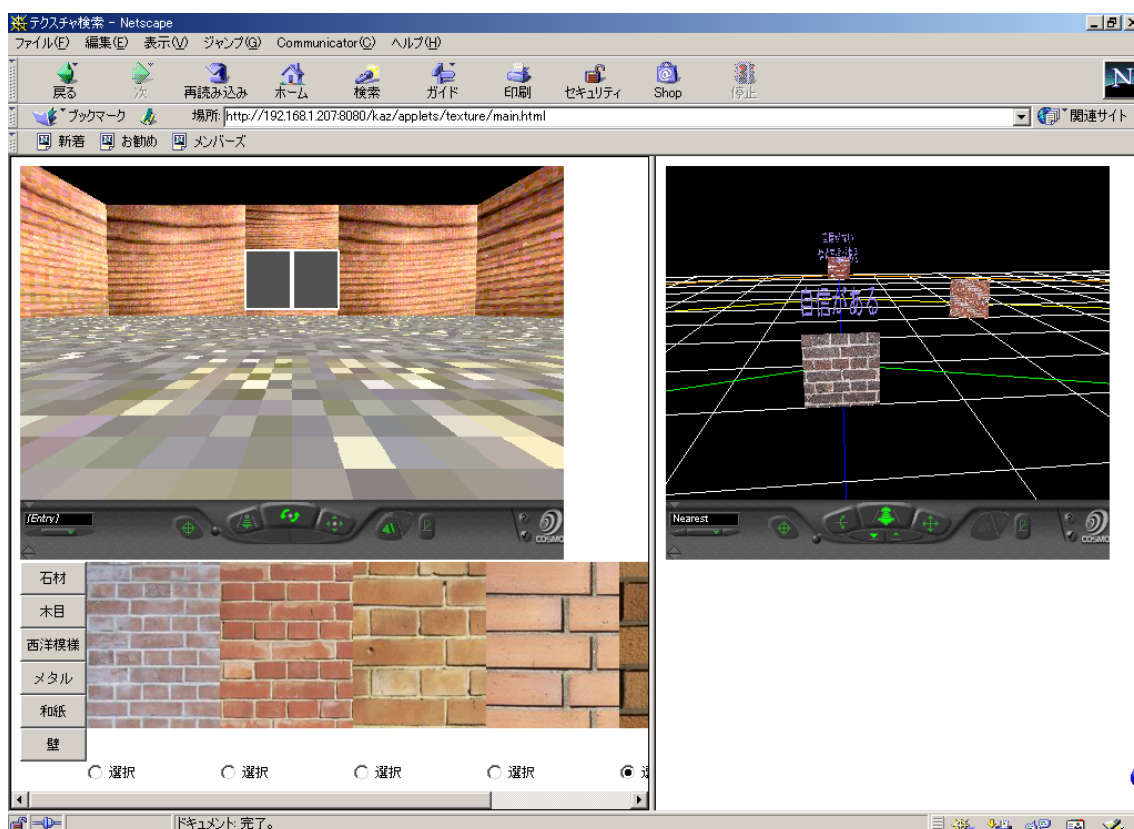
4つのモデルを意識してシステムを設計すると、1つのモデルだけでは対応できないさまざまなインタラクションに対応できるようになる。すなわち、それぞれのモデル特有の技術を組み合わせることで、単独の技術だけでは対応できないことができるようになる。

ヒューマンメディアプロジェクトで試作された工業デザイン支援システムや、これを発展させた中央大学のインテリアコーディネート支援システムでは、画像をキーとした例示検索や類似検索機能を提供するとともに、検索結果を工業デザインやインテリアコーディネートに利用するための作業空間をも提供している。

想定されている利用者は、工業デザイナー、インテリアデザイナーと顧客であり、それぞれの利用者の感性モデルを構築し管理している。また、個人プロフィールの共通する人々の感性モデルを集約化した平均感性モデルも利用可能である。対象とするコンテンツはインテリアのコーディネートに必要な家具（3次元物体モデル）や壁材・床材（テクスチャ画像）である。

データベースからの検索結果を利用してデザインやコーディネートを試行錯誤的に進めるための（仮想空間による）物理空間を提供している。利用者は、検索したテクスチャ画像（例：床材や壁材のパターン）を仮想空間上で床や壁に貼りつけて、また、仮想空間をウォークスルーして、コーディネートを試したり出来映えをさまざまな角度から評価できる。

これらのシステムはまた、データベースから検索されたデータを、利用者の感性に合った順に、また、データ間の類似度にしたがってこれらデータを3次元空間上に配置することが可能である。このように、データ間の関係を直感的に理解できるようにするために（仮想空間による）情報空間を構築し、利用者にはアプリケーションシステムへのインタフェースとして提供している。



(左: 物理空間、右: 情報空間)

図 3.3-4 テクスチャ検索システム上の空間例

これら2つの空間は連動しており、情報空間で指定したテクスチャが、直ちに、物理空間上で壁材や床材として扱われる。テクスチャ検索システム上の空間例を図 3.3-4 に示す。

典型的な利用の仕方としては、デザイナーが顧客の意向などを聞き、漠然とした要求を順次具体化・詳細化しながら、デザイナーと顧客が共同作業の形でコーディネートを進めていくことを考えている。したがって、単に検索を行ったりコーディネートの試行錯誤を行ったりするだけでなく、両者の意見の相互理解を進めるために、情報空間を用いて、感性モデルの可視化なども行われる。

### 3.3.5 人間中心型の情報技術の研究開発動向

#### 3.3.5.1 わが国の関連プロジェクトとその位置づけ

1990年代から、例えば通商産業省のリアルワールドコンピューティングプロジェクトでは、人間とロボットなどの情報機器がともに働く環境の実現に向けて、景観画像からの人間の認識、顔の認識などの技術開発も進んできた。しかし、人間の位置や姿勢、行動を認

識し、また、個人の識別に結びつけるこれらの技術も、人間とのかかわりという点では、人間をモノ（物理的な実在）として計測し、認識し、働き掛けるという考えにとどまっている。

2000年代に入って、経済産業省や文部科学省関連のデジタルヒューマンプロジェクトのように、新しいセンサ技術やネットワーク技術と人工知能・パターン認識理解・ロボット技術を結びつけて、人間の体温や呼吸、筋肉の緊張の度合いや、皮膚感覚の快適性を計測する技術開発が活発になってきている。

産業技術総合研究所のデジタルヒューマンラボでは、人間を運動機械的なモデル、生理解剖学的なモデル、認知心理的なモデルからとらえ、人間に住みよい快適な環境を情報技術により実現することを目指している。人間の状態を計測し、人間に快適な環境を実現しようとする試みは、「人間中心」「人にやさしい」情報システムの実現にも大きく貢献するものと期待される。

デジタルヒューマンプロジェクトでは、

- (a) システムが人間を観察し、人間を支援するように環境を制御する「人を見守るデジタルヒューマン」
- (b) デジタル空間の中で人間と環境の親和性を評価し、人間と調和がとれるよう実環境を設計する「人に合わせるデジタルヒューマン」
- (c) 人間機能の統合的モデリングをめざす「人を知るデジタルヒューマン」
- (d) これらの技術環境を与える「デジタルヒューマンプラットフォーム」

の研究開発が精力的に進められている。しかしながら現在の研究開発は、人間を生物（生理的な実在）として計測し、認識し、働き掛けるという範囲にとどまっているようである。

デジタルヒューマンの基本的なコンセプトは、1990年代に実施された通商産業省工業技術院の人間感覚計測プロジェクトで行われた人間の生理的な状態を計測する技術に、最新のセンサ技術、ネットワーク技術、ロボット技術を融合させたものともいえよう。

同じく、2000年ごろから、モバイル情報環境とユビキタス情報環境が融合した新しい情報環境を指向した研究開発としてサイバーアシストプロジェクトが経済産業省、文部科学省などで進められ始めている。これは環境が人間を観察することを通して、その時々、その場所で、適切なサポートを人間に提供できるようにするためのものである。主たる研究目標は、

- (a) 高齢者にも安心して使える人間中心の情報サポートシステム
- (b) デバイス設計から情報センサ網、知的エージェントまでを有機的に結合

などである。

環境自身が人間と人間の知的な活動を観察し、その状態を理解し、人間にたいして必要な支援を行うという発想は、人間の主体性を尊重しながら「人にやさしい」環境を実現するものといえるだろう。

人間中心の情報システムとして、感性情報処理の仕組みに正面から取り組んだのは1990年代後半に実施されたヒューマンメディアプロジェクトである。これは、人間を情報技術の面から「感情を持った生物」としてとらえて、その生理的・心理的・認知的な特性をモデル化する試みである。感性のモデル化や感性データベースなどの感性情報処理技術と、バーチャルリアリティを含むヒューマンインタフェース技術を融合させるコンセプトと試みは、今なお、今日的な技術課題であると思われる。

#### 3.3.5.2 英国における関連研究開発の動向

ヨーロッパでは「ユーザフレンドリ情報社会」の実現を目指した研究開発プログラム（IST プログラム）が進行中である。ここでは、その一環として英国で進められているPACCIT（People@the Centre of Communication and Information Technologies）プロジェクトの概要を紹介する。

PACCITは、英国の教育省下の工学系の学術会議、社会科学系の学術会議、および産業省の大学と産業界を結びつけるLINKプログラムが合同して進められている。全体的な研究課題・技術課題は「人間、コンピュータ、組織の間の複雑な相互作用を理解し、また、どのように改良するか」である。

現在は、次の13テーマの研究開発が進められている。

##### (1) MAGIC: Multi-modality and Graphics in Interactive Communication

Human Computer Interactionにおける人間・人間、人間・計算機間のインタラクションツールをマルチモーダル化する。具体例として、建築設計での共同作業、打ち合わせを題材に、マルチモーダルなコミュニケーション、特に、図形・図示を利用することの有用性を解明する。

##### (2) E-space: Dynamic Information Displays for Transaction-Based Activities

人間に情報を提示する上での基本的な方法論を探る。具体例として、製品の設計過程や、消費者・エージェント間のコミュニケーションで必要となる種々の情報をどのようにわかりやすく提示・表示するかを解明する。

##### (3) Design for Interaction and Collaboration

博物館・美術館で利用するような携帯型の情報端末を開発するとともに、これを利用するインタラクティブな展示方法と展示物の設計を試みる。芸術家、システム開発者、キュレーターが一体となった、インタラクティブ展示の設置を行う。

##### (4) ‘The Way We Were’: Situational Shifts in Collaborative Remembering

組織の中で知識と経験の共有を効果的にするための情報環境のあり方を探る。特に、人

間の心理的・認知的な側面に注目しており、"collaborative inhibition effect"などの「社会的」な現象を解明する。

(5) **Socio-technical Systems Design: Knowledge Capture and Management Using Patterns**

技術開発とマーケット（ビジネス）の間にあるギャップをどのように知覚し、克服するかを研究している。「知識」の形態のミスマッチ、暗黙知識の共有、組織としての知識の欠如（不足）などの問題を研究している。

(6) **Representational Design Principles to Humanise Automated Scheduling Systems**

ビジネスにおけるスケジューリングの手法を「人間的」にする方法を研究している。

(7) **Collaborative Case-based Learning in Distributed and Synchronous Environments**

事例からの協調的な学習のメカニズムを開発する。仮想的な訓練センターを試作し、異なった知識分野でも有効な学習法を開発し検証する。

(8) **Effective Decision Support in the Neonatal Intensive Care Unit (NEONATE)**

新生児の集中治療において医者、看護婦らが、多種多様な情報を扱いながら、意思決定をするタスクを分析し、適切な支援を行うためのメカニズムを開発する。特に、患者と医療スタッフの用語の違いに注目して研究を進めている。

(9) **Decision Support for Risk management Planning**

意思決定の過程をサポートするシステム REACT (Risk, Events, Actions and their Consequences over Time) の開発を行う。特に、意思決定や計画立案に必要な情報の提示法を研究する。

(10) **Investigating the Impact of Tailored Reports on Anxiety**

個人向けの「健康に関するレポート」の作成法。レポートが患者にどのように利用されるか、どのような影響を与えるかを探る。また、このような情報の提供により、患者やその家族の間でそのような社会的な相互作用行われるかを探る。

(11) **Spoken and Written Language in Adaptable Multimedia Documents**

非ネイティブな人、高齢者など、リテラシー上の「弱者」を対象に、マルチメディア文書（発話&文字）の効果的な執筆法を開発する。

(12) **Multimedia Communication as a Recreational Activity**

レクリエーションの中での人と人とのマルチメディアコミュニケーションを調査する。よいレクリエーションの経験とは何かを探る。

(13) **Exploring and Mapping Interactivity with Digital Toy Technology**

幼児と電子おもちゃの間のインタラクションを調べる。幼児はおもちゃに対してどのように考え、感じるかを調査する。

以上に見られるように、産業界とリンクさせることが期待されているプロジェクトスキームであるにもかかわらず、実際的なシステム開発から、現象の深い調査・原理の探求までカバーされているのは、英国らしいと言えるだろう。

## 3.4 新世代グループウェアのためのアウェアネス研究の最新動向

報告者: 國藤進委員

### 3.4.1 はじめに

グループウェア研究が一段落した 1990 年代になって、同期・対面環境ではあたりまえの存在感、実在感や臨場感が双方向通信環境では欠落していることに気づき、存在感、実在感や臨場感などのアウェアネスを補完する研究[14]が勃興してきた。その草分けは Dourish らの「だれがだれと話し、だれが話し手や聞き手の周辺にいるか、彼らはどのような行為をしているか」[1]といった日常の同期・対面作業ではあたりまえの情報が双方向通信環境では欠けているという認識から出発した。そこで報告者らは存在感や実在感のアウェアネスを伝達する双方向通信環境構築を踏まえて、最終的には対面環境以上の臨場感のアウェアネスを伝達する双方向通信高品位 A-V 環境を構築するという研究目標を立てた。

実際、分散協調作業の進捗を支援するにはアイコンタクトのできる環境を必要とするという視点から、石井はすでにゲイズアウェアネス[9]の提供できる環境を構築していた。仮想空間でのオフィスでの出勤感や連帯感を高めるための位置アウェアネスの研究[3]、会話開始のためのきっかけを作るための同じ作業をしていることを相手に気づかせる存在のアウェアネス研究[15]、インフォーマルなコミュニケーションを促進させるインタレストアウェアネスの研究[16]などが積極的に行われた。このような研究の延長線上にナレッジアウェアネス[28]、情報取得アウェアネス[11]、WWW アウェアネス[18]、コミュニティアウェアネス[8][20]などのさまざまなアウェアネス研究が行われている。最近のグループウェア、CSCW 関連の国際会議では、何らかの意味でアウェアネス関連の研究と言えるものが激増している。

そこで報告者らは臨場感、雰囲気、気配、熱気、凄みなどのアウェアネスを伝達できる遠隔コミュニケーション技術の研究開発を行った。これにより、分散環境でも対面環境と同様な突っ込んだ遠隔会議[15]や遠隔教育[21]が可能となることが期待できる。さらに、日本学術振興会・未来開拓学術研究推進事業宮原プロジェクト「未来映像・音響の創作と双方向臨場感通信を目的とした高品位 Audio-Visual System の研究」の深い感激をも伝達できる電子的 A-V システムの提供により、深い感性をも共感できる双方向通信基盤が構築できることが期待される。

まず 3.4.2 節で、同期・対面環境で自明の存在感、実在感、臨場感を、双方向通信環境でも伝達できる遠隔アウェアネス通信の研究のサーベイ[14]を行い、報告者ら独自のアウェアネス通信の体系的分類を行った。それによると、アウェアネス研究の分類には、協調

作業におけるアウェアネス、情報・知識共有におけるアウェアネス、臨場感支援としてのアウェアネス、視覚・聴覚以外のアウェアネス、気配や状況のアウェアネスといった分類が考えられ、それぞれ活発な研究開発がされている。

それに対して報告者らは浅い感性を伝達する個別アウェアネス研究を 3.4.3 節で、それらを統合するアウェアネス研究を 3.4.4 節で、さらに、深い感性を伝達するアウェアネス研究について 3.4.5 節で述べる。報告者らは従来から非同期の分散環境である WWW 環境で、存在のアウェアネスや動作のアウェアネスを提供する環境の構築を行ってきた。高品位なテキスト、手書き文字、静止画像や動画像を用意することによって、どれだけ対面環境と同等あるいはそれ以上の A-V 環境に近づきうるか、そのために何をを用意すべきかが課題であったが、「高品位」A-V 環境を用いることにより、著しく感性の伝達が効果が高まることを明らかにした。

浅い感性の個別アウェアネス研究として、報告者らは情報取得アウェアネス[11]、WWW アウェアネス[18]、WWW 探索アウェアネス[22]、およびカンバセーションアウェアネス[4]という概念を実現する環境の構築について 3.4.3 節で述べる。それらの統合アウェアネス環境の構築については、デモ展示会場におけるコミュニティの知的活動を支援するコンテキストアウェアネス環境の構築事例[6][24]を、3.4.4 節で紹介する。

深い感性のアウェアネス研究の第一歩として、宮原プロジェクトの成果の脳科学的インプリケーションを明らかにする実験を行った。この結果は、3.4.5 節に要約されるように、高品位 A-V システムの静止映像を用いて高品位映像を表示すると、アルファ波が従来画像よりも発生するという興味深い成果[2]を得た。

勿論、深い感性を伝達するには、宮原プロジェクトのように視覚・聴覚の見えざる因子を探するというアプローチ以外に、触覚や嗅覚のような他の感覚情報を伝達することでアウェアネスを深めるアプローチがある。これに関して岡田らは多地点同期型メディア空間におけるアウェアネス支援方法を研究した。試作した芳香発生装置を用い、臭覚情報の提供の効果を確かめる臭覚アウェアネスの研究[19]を行った。また宗森らはローケーションアウェアネスの研究として、遠隔地での移動グループ間の気配の伝達実験を行った。お互いに離れており、かつ移動中の環境における気配といったアウェアネスを、双方向に伝達する方法として、音や振動の効果的な利用法を検討し、実際の遠隔地間でシステム[17]に適用した。

### 3.4.2 アウェアネス研究の分類と動向

アウェアネスを分類する軸にはさまざまなものがあり、多くの人々によってアウェアネス関連研究が行われているが、ここでは便宜上、5つの分類についてのみ述べる。

#### 3.4.2.1 協調作業におけるアウェアネス

協調作業のプロセスを、同じ場所に集合するコプレゼンス、他者の存在や動作に気づくアウェアネス、円滑な会話を進行できるコミュニケーション、そして実際に協調作業を行うコラボレーションという四つの軸で分類するのが松下・岡田の分類[16]である。中川のWWWアウェアネス[18]はこの存在と動作のアウェアネスを、WWWの利用者がリアルタイムに利用できるようにしたものである。なおチャットベースのツール CHOCOA (<http://www.fujitsu.co.jp/hypertext/free/chocoa>) も存在のアウェアネスを利用したツールである。

#### 3.4.2.2 情報共有、知識共有の場としてのアウェアネス

山上は「協調行動過程支援において必要となる情報共有過程に関して、グループメンバーが相互認知し、気づくという概念」[28]として、ナレッジアウェアネスの概念を提唱した。すなわち、オフィスにおける知的生産性を向上させるための、情報の共有を促進するコミュニケーション手段という概念の提唱である。このナレッジアウェアネスの特殊なものとして、組織におけるノウハウ知の共有促進、組織知の再利用促進をはかる仕掛けを実装したのが門脇の情報取得アウェアネス[11][12]である。この研究の方向性は、ナレッジマネジメントにおける情報共有あるいは知識共有の重要性として再認知されつつある。なお学習環境の討論の誘発を目的としたナレッジアウェアネス研究[21]もある。

#### 3.4.2.3 臨場感支援としてのアウェアネス

形式知を暗黙知に変換する、あるいは暗黙知を暗黙知のまま伝達するには対面環境と同様な臨場感を提供する必要がある。そのようなツールをそれぞれ内面化支援ツール、社会化支援ツール[13]と呼ぶことにすれば、それらはどのような機能を提供すべきであろうか。既存のマルチメディアグループウェアが視覚・聴覚情報の伝達装置であることに注目し、アイコンタクトのみならず画面上のどこを見ているかの動作をアウェアさせる装置が石井の Clear Board[9]である。

#### 3.4.2.4 視覚、聴覚以外のアウェアネス

臨場感は視覚、聴覚以外の情報の伝達も伴って、初めて成立するという立場で、臭覚、触覚、味覚情報までも分散環境で伝達しようという研究開発もある。岡田らの研究では、森林やお化粧の匂いを伝達する特殊な装置[19]を研究開発中である。化粧品や花の仮想空間上でのバーチャルショッピングには必要な機能かもしれない。なお味覚のアウェアネスの研究のみ、不思議なことに未だ行われていないようである。

#### 3.4.2.5 気配や状況のアウェアネス

ビットの世界と物理世界を融合した気配や状況を伝達できる環境を構築しようというの



が石井のタンジブルビッツプロジェクト[10]である。各種センサによる情報獲得、レーザを含む多様な出力装置を活用し、気配や状況に応じたウェアネスを提供することにより、デジタル世界に相互交流できる物理世界を実現できる。デジタル空間と物理空間が融合する、あるいは仮想世界と実世界が融合する新しい知の空間を設計する場合、新しい建物のアーキテクチャをデザインするときなどに有効であろう。

### 3.4.3 浅い感性を支援する個別ウェアネス研究

本節では、浅い感性を支援するウェアネス研究のうち、報告者らの個別研究成果のトピックについて述べる。

#### 3.4.3.1 情報取得ウェアネス

山上のナレッジウェアネスの概念を組織知の共有・再利用プロセスに活用しようという研究から生まれた。門脇は「システム内に蓄積されている情報の再利用を促すために、コンピュータを用いて、情報の存在、周囲の情報共有活動などをユーザに気づかせる技術」として、情報取得ウェアネス概念[12]を導入する。具体的には時節の視点からのウェアネス、組織の視点からのウェアネス、他者の視点からのウェアネスを提供した情報共有促進支援システム GGG (information Grasp with fishGlobe for GoldFISHes) [11]を実現した。上記 3 種のウェアネスを付随機能として付与した GGG を実際に 13 人のユーザが 3 ヶ月に渡って試用、評価実験した。その結果、共有情報の検索数、参照数でそれぞれ 1.3 倍、1.5 倍の増加、複数参照数で平均 5.4 倍の増加が見られ、本ツールの有効性が確認された。

#### 3.4.3.2 WWW アウェアネス

ネットワーク上の分散環境の会議に変えた途端、対面環境の会議では当たり前の視覚情報の一部がウェアされなくなる。特に最も多用される WWW 環境で、複数のユーザが同一のホームページを覗いているとき、同一のホームページを見ているのはだれか、彼らが今そのホームページのどこを見ているか、それを見て感じた印象は何か、これらを専用の WWW アウェアネスサーバを通して、「透かす」ように見せることにより、WWW 上のコミュニケーションやコラボレーションをサポートするのが、富士通北陸システムズと報告者らで共同開発した Web アウェアネス表示環境 WebCoordinate[18]である。それぞれの人固有の「赤ペン」を表示できるようにしており、ソフトウェア開発の分散レビューや実際の大教室での講義運用などが試みられている。

### 3.4.3.3 WWW 探索アウェアネス

WWW 上で複数の人間で協力しあってブラウジングしようとしたとする。その際、報告者らは既存の WWW 環境では、複数の人間が探索しているという行動の把握、しかも他者がどの辺りの WWW 空間を探索しているかの実態を認識できない。そこで、共同して WWW 空間を探索するのに有用な URL 表示の可視化とそこでの視点のリアルタイム表示を行い、ブラウジングを協調して行う行為を WWW 探索アウェアネス（あるいはコラボレーションブラウジング）と呼ぶ。WWW 探索アウェアネス情報を提供するシステムを坂本が試作[22][23]した結果、他人の存在、動作をある程度予測でき、他ユーザとの円滑なコミュニケーションを促進し、一人では見つけられないホームページを多数見つけ出せることを実証した。ホームページ間の位置関係の空間配置は、リンクのジャングルを探索する際の手掛かり情報を提供する。納豆ビュー[27]の探索モードでの改良版とも言える。WWW 探索アウェアネス環境を図 3.4-1 に示す。

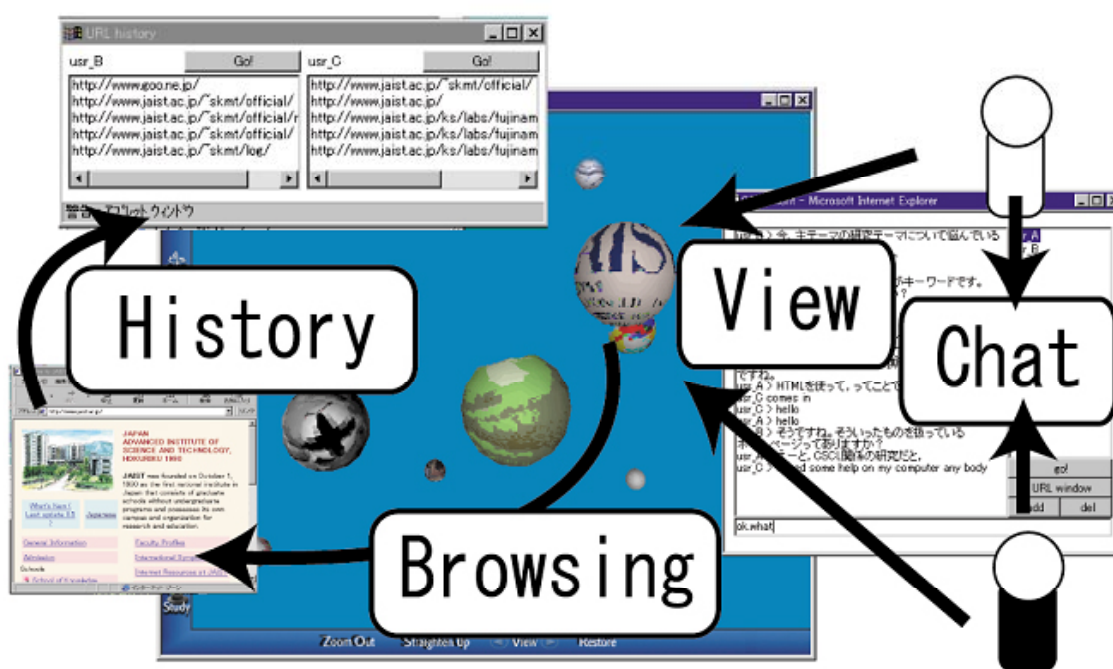


図 3.4-1 WWW 探索アウェアネス環境

### 3.4.3.4 カンバセーションアウェアネス

インターネット上でテキストベースコミュニケーションを行う割合は圧倒的に多いが、その際仲間同志のコミュニケーション過程のリアルタイムでの実態を分析表示する社会性

指向ツールはほとんど存在しない。実世界における社会的コミュニケーションと同様なだれが話者で、だれが聴者なのか、だれがだれだれに対して発話しているのか、聴者はだれに対して発話を返したか、だれが特定のキーワードに関して特に発話頻度が高いか、あるいは発話したのかといった発話状況をリアルタイムに表示することで、カンバセーションウェアネスの提供を試みた。これらのメタ情報をリング上の空間配置状況としてリアルタイムに表示するシステムを、伊藤が構築[4][5]した。その結果、カンバセーションウェアネス表示機能のない通常のグループメールに比べて、メッセージ総数と平均文字数で、それぞれ 50%と 71%の増加が見られた。また対話関係の成立や継続対話回数の増加も見られた。特に、継続的対話関係の支援において、本システムがきわめて有効であることが実証された。カンバセーションウェアネス環境を図 3.4-2 に示す。

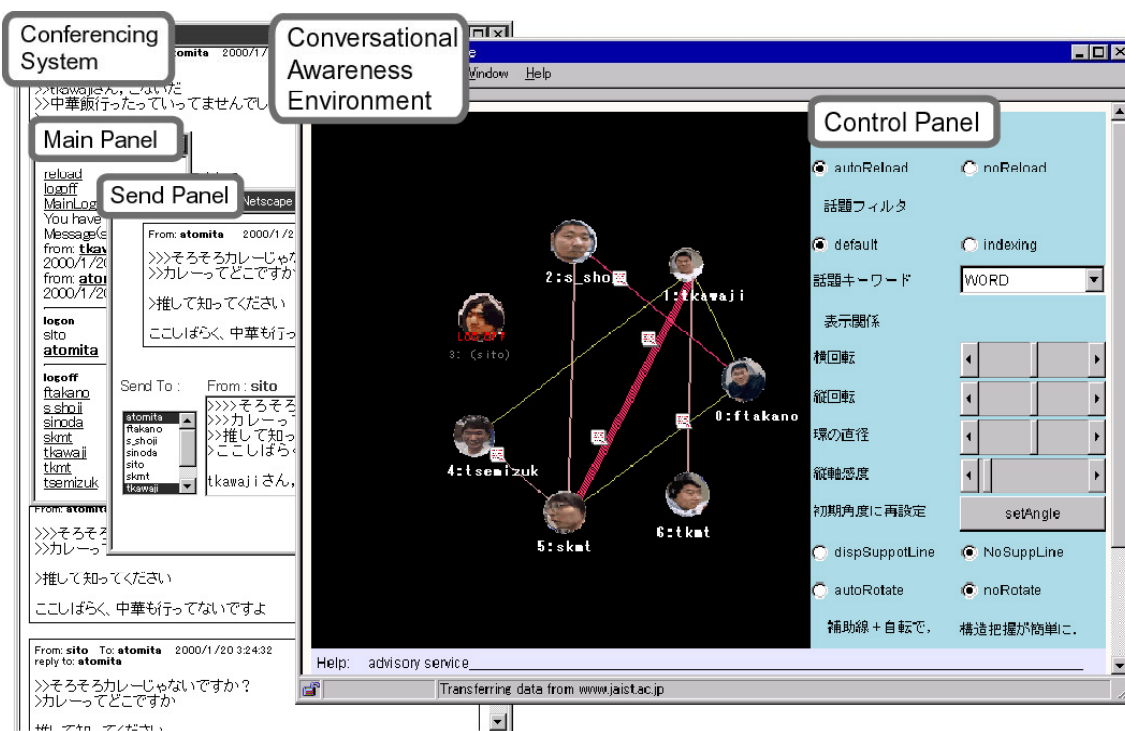


図 3.4-2 カンバセーションウェアネス環境

3.4.4 浅い感性を支援する統合ウェアネス研究

浅い感性を支援する個別ウェアネス研究の統合・集大成として、図 3.4-3 に示されるデモ展示会場内のコンテキストウェアネス環境の構築がある。デモ展示会場で赤外線バッチを着けると、特定の展示物の情報が個々人のパームガイドに表示される。その表示を読み、気になることや気づいたことをどんどんメモとして書き込んでいくと、メモ書きのアノテーションからコミュニティ内の情報交換やインフォーマルコミュニケーションの偶発的開始がおり、これにより新しい知識が創発される。また展示物を見た感想をテキスト入力すると、その感想テキストが漫画風に要約されるというエンタテイメント性の高いツールも試作した。

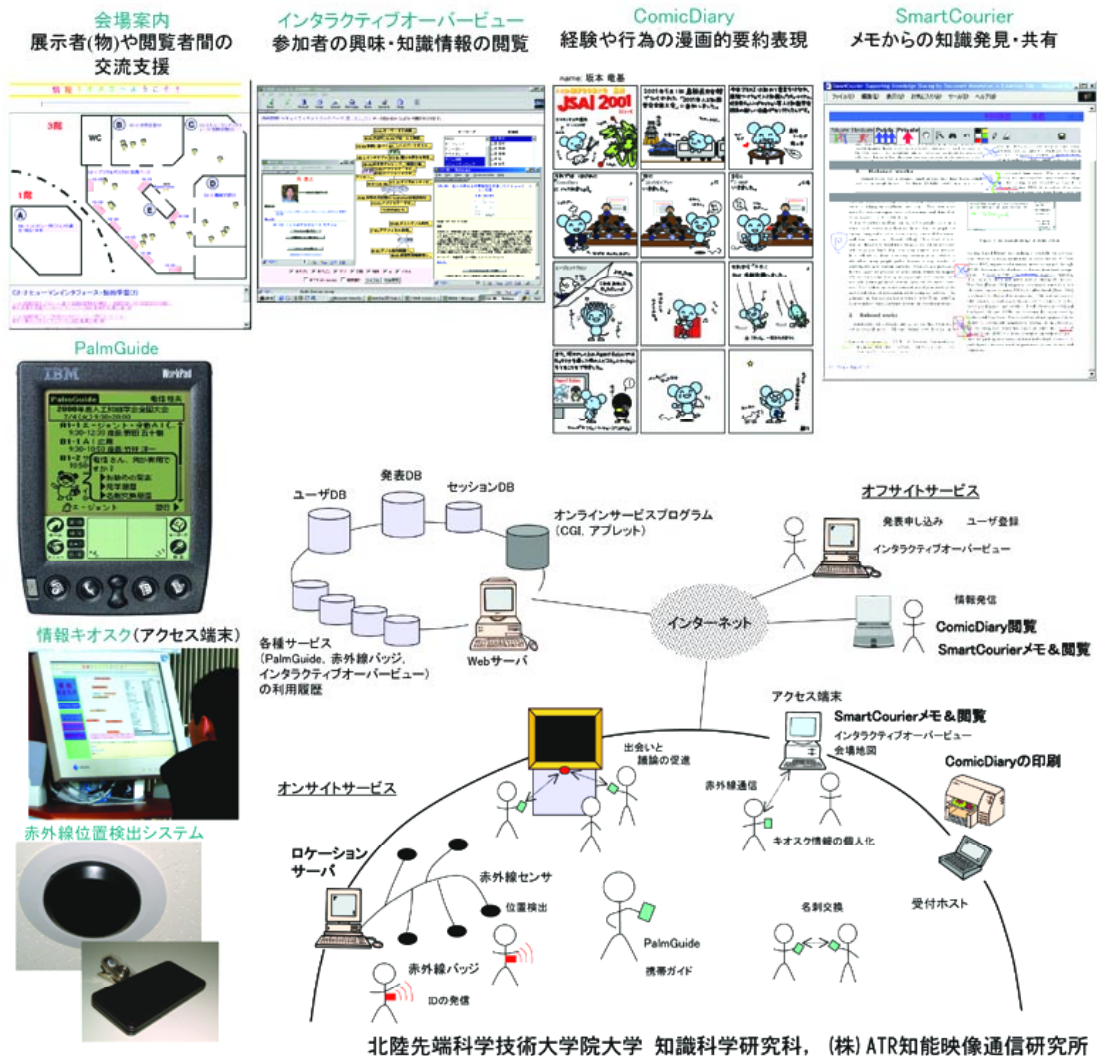


図 3.4-3 コンテキストウェアネス環境の構築

#### 3.4.4.1 ComicDiary: 経験や行動履歴の漫画的要約表現[25][26]

多くの要素から成り立っている情報に対して、その構成要素の関係把握を視覚的に容易化させることを目的として、情報の視覚化に関する研究は盛んに行われてきた。しかし、これらの研究は単に要素同士を、有向グラフなどを用いてそのまま視覚化しているにすぎず、周囲の状態や経験や思い出といった、対象物とそれを取り巻く状況や前後関係などを表現するにはいたっていない。本研究では、漫画を用いた経験や思い出の視覚化手法を提案した。これは、C-MAP という展示会場における個人化されたガイドシステムのログを用いて、その会場で何を見たのか、何に興味を持ったかなどを、漫画を用いて表現する、親しみやすい視覚化手法である。

この個人化された漫画生成サービスは、展示会場においてユーザが見て感じたことをストーリー化し、その流れに沿った、親しみやすい表現手法である漫画を自動生成することによって、展示会での思い出をより留めてもらう効果をもった視覚的記憶媒体を提供することを目的としている。具体的に、今回は C-MAP システムという個人化された展示ガイドシステムによって得られる個人的なサービス授受の時系列データや興味データと、ユーザにかかわらず静的に判明している会場データやタイムテーブルなどを組み合わせることによって自動的にストーリーを生成、漫画化し、印刷物として個々のユーザに配布することにより実現した。

状況表現手段として漫画を使用する利点は以下の通りである。

- (a) 親しみやすい。
- (b) 一覧性が高い。
- (c) 表現力が高い。
- (d) 時系列に沿った表現が可能。

経験や行動履歴を表現するにあたり、報告者らが対象としているユーザの種類は広範囲におよぶ。例えば、今回のプロトタイプ之母体である C-MAP は、さまざまな種類の展示会場におけるサービスの提供を対象としているので、ユーザの年齢層は広域におよび、訪れる多くの閲覧者に受け入れられやすい媒体であることは、この漫画要約サービスにおいて重要な意味をもつ。よって、(a)、(b) のユーザに対する親和性は大変重要である。また当然のことながら、展示会場での経験を静的なメディアに写像するためには、(c)、(d) に代表される表現力も重要な要素である。報告者らが普段目にする表現方法の中で考えると、これらのことを同時に兼ね備えた表現方法という意味で漫画は稀有な存在である。例えば、文章は (c)、(d) は確実に兼ね備えているが、(a)、(b) に関しては弱い。箇条書きにすれば (b) の部分はクリアされるが、代わりに (c) を損なうことになる。また、画像は (d) の要素が欠如しており、動画は (a)、(c)、(d) には長けているが、(b) の能力が致命的に低い。これら利点と欠点は文字や画像に固有の特徴に起因するものであり、そ

の両方を含んでいる漫画は両者の良い点と悪い点を多重に継承しているといえる。さらに、これらの利点と、何より漫画にはストーリーを表現するための手法としての歴史的実績があるという点で、報告者らは漫画を経験や行動履歴を表現することに対して信頼に足りる、適切な表現手法であると結論づけた。

一方で、漫画は

- (e) 正確さに欠ける。
- (f) コマ数を制限すると、多くのことを表現できない。

などの今まで述べた特徴とトレードオフの関係にある欠点をもっている。これらは漫画のもつ仕様なので、制作者はこれら制約の中でいかに表現するか、また制約を逆に利用するかを考えなければならない。さらに、生成という観点から見た場合、漫画は上記の欠点とは別に、より根本的な以下の問題をかかえている。

- (g) 作成に専門的スキルを要する。

この問題は主に素材の作画とベースとなる物語の生成に分けることができるが、そのどちらも自動生成という観点からは解決が難しい。しかし、本研究で行ったサービス提供のプロトタイプの実験にあたっては、根本的な絵やストーリーが自動生成されているかどうかは問題ではないと考え、今回はこの部分には焦点を当てず、漫画制作の経験者のスキルによって解決した。

このプロトタイプに対するユーザの反応と有意義性を確かめるため、第 15 回人工知能学会全国大会において、運用実験を行い、使用感に関するアンケートを採取した。ここでは、本システムが出力した漫画と、見学した対象のアブストラクトをテキストのリストとして表現する C-MAP の別サービスを比較した。これによると、両者の有意義性に有意差は見受けられないが、漫画による経験の表現自体は利用者にまったく抵抗なく受け入れられていることが判明した。また、アンケートのフリースペースに書かれた意見には、「〇〇の見学の記録を入れて欲しかった」など、漫画の内容に関するものが多く、漫画の表現を用いていることに対する意見は皆無であった。今後は、ユーザの希望する内容と漫画家が書く内容のマッチングに関する機能付加を、ユーザモデリングの技術を応用して行っていく予定である。なお本研究は第 15 回人工知能学会全国大会ベストプレゼンテーション賞を獲得した。

#### 3.4.4.2 SmartCourier: アノテーションからの知識発見・共有[6][7]

本研究は、展示会場における参加者個人の状況や興味といったコンテキストに応じて、展示見学に関連する個人化された情報提供による効率的な情報獲得の支援や、参加者間に

共通する知識や興味から参加者コミュニティにおける交流や知識共有を支援することで、展示見学への参加を通してなされる知識流通、知識創造の促進を目標としている。ここで構築した **SmartCourier** は、ユーザによる展示会場に関連する資料群へのアノテーション行為からその興味関心を抽出し、これを材料として情報の個人化や知識共有を支援するシステムである。

本研究では、ユーザの興味関心を抽出する対象としてアノテーションに注目し、その利用形態を個人的利用と組織的利用の2つに分けて考える。最も一般的なアノテーションの利用は個人的に行われる。すなわち、ユーザが閲覧するさまざまな資料や論文などの中から、特に注目すべき箇所を視覚的に附加し明示することで、ユーザ自身による情報の継続的活用、再利用性を高める一般的行為であると言える。このような行為は報告者らが研究対象とした展示会場でも数多く観察された。すなわち、情報の効率的収集と記録のため、視聴すべき演目や発表プログラムにマーキングしたり、興味深い発表へのコメントなどをその資料へ記しておいたりするような作業は、そのユーザの興味関心に応じて、自然に行われる作業であると言える。

一方、このようなアノテーションは、継続的組織において共有される場合もある。例えば小規模な研究グループのように、ユーザがある程度相互に興味関心の趣向を把握しているようなコミュニティにおいて、資料に関連するアイデアなどの知識の共有や効率的な情報取得のため、相互に役にたつような情報を評価し、これを共通のマーキング記号を使って示すといった作業が行われる。例えば、報告者らの研究室では、参考となる論文について、それを読んだ者が、内容の難度と質を示すシールを貼りつけ、コメントを残していくことで、資料検索をより効率的に行っている。しかしながら、このようなアノテーションの共有による効率化は、展示会場のような不特定多数の参加者が存在し、集団としては一過性のものである場合、物理的に書類の共有が困難であるというほかに、その展示会場において参加者が持つ興味関心に差異があり、他者のアノテーションが役に立たないといった問題がある。

そこで、本研究では、これらの問題に対処するため、ユーザによる資料へのアノテーション行為から、ユーザの興味モデルを構築し、これによって興味・関心に適応的な情報共有と情報検索のためのガイダンスを行うシステムを構築する。本システムのクライアントは、展示会場のブースに設置される液晶タブレットを備えた PC キオスクマシン上で動作することを前提に構築された。クライアントはウェブブラウザ上の **Java Applet** として動作し、ユーザはクライアントを通して、展示会場に関連する種々の資料を検索、閲覧し、アノテーションを記入することができ、記入したアノテーションは、インターネットを通して展示会終了後に確認することができる。これらアノテーションの対象となる資料については、そのイメージとアノテーション認識のための文書書式情報、キーワード情報を **Content Database** としてサーバ側に保持する。システム中の **Stroke Manager** は、ユーザのペンストロークをその形状とストローク間の時間的距離的近さから分類し、文書中の

どのキーワードに関連づけられたアノテーションであるかを決定する。この情報をもとに **User Profile Generator** がユーザの興味モデルを生成する。興味モデルは、アノテーション数によって重みづけられたキーワードセットによって表される。

ユーザは個人的に資料を閲覧しアノテーションを記入することができるほか、興味関心が近い参加者同士で互いのアノテーションを共有することができる。また、一般に、未読の資料であっても、興味関心が近い他者がその資料を評価していれば、似た評価を下す可能性が高いという前提を利用した協調フィルタリングを用いて、未読資料のリコメンデーションによる情報の効率的取得を支援している。これらの機能により、ユーザは小規模な継続的組織におけるアノテーションの共有が果たす、情報取得の効率化と知識の共有といった役割を、一過性の集団でしかない展示会場においても享受することができる。本システムにおいて、ユーザ間の類似度はユーザ興味モデル間の相関係数によって求めている。また、協調フィルタリングの評価予測は、類似ユーザグループの興味モデルを使い、平均2乗誤差法を用いて行っている。また、ユーザのアノテーション行為と資料検索による情報の再利用を効率化するツールとして、手書きペン入力されたアノテーション記号の検索やこれらの一覧表示といった機能も備えている。

本システムは、上記人工知能学会全国大会において公式サービスの1つとして提供し、複数の参加者に利用された。アンケートなどを集計し、より詳細な使用実験を行い、研究内容は学会誌[7]に掲載された。

#### 3.4.5 深い感性を支援するアウェアネス研究

高品位映像の評価技術についての研究は、まだ始まったばかりであり、その確立は複合現実感などを含む新たな感性評価技術の基盤作りとして重要である。本研究では、高品位映像の客観評価法の1つとして自発脳波が利用できないかを検討し、高品位画像システムにより提示される R、G、B 各 4096 階調の再現が可能である高品位映像について脳波を指標とする客観評価法の開発[2]を試みた。

脳波測定実験では、高品位映像と基準映像をそれぞれ鑑賞しているときのアルファ帯域波にどのような違いが生じるのかを明らかにし、アルファ帯域波を客観評価指標として導入することを試みた。映像の提示方法は映像1枚につき100秒間、インターバルを25秒間においてランダムに4回提示して、これを1ブロックとした。100秒間のうち前半50秒間、後半50秒間のどちらかに基準映像を提示し、残りの50秒間は高品位映像を提示した。高品位映像と基準映像のどちらが先であるかはランダムである。実験回数は1日3ブロック（休憩含む）を目安として計12回を連日で行い、毎日最初の1ブロックは実験環境の慣れを考慮して予備的な計測とし、時間効果を考慮して2ブロックのうちランダムに選ばれた1つを解析対象としたため、1名当たり4日間で計4ブロックが解析対象とされた。



上記の提示方法により、解析対象とした 100 秒間のうちの 50 秒ずつについての順序効果、時間効果は均等化、ランダム化されている。100 秒間連続して映像を提示したのは、高品位映像と基準映像の違いである階調数の変化が脳波に与える影響以外の要因を極力抑えるためである。解析対象としたデータは上記提示方法により得られた 100 秒間の提示を 1 件として、計 192 件である。一般にアルファ波は後頭部において振幅が大きく、出現頻度も高いことや、前頭部は瞬きによるアーチファクトが混入しやすいことから後頭部周辺の O1、O2、P3、P4 について解析を行った。なお図 3.4-4 に示すように国際 10-20 法に基づき脳波電極を装着し、EOG、ECG も同時に計測した。

電極: 銀皿電極使用

国際10/20法 10ch

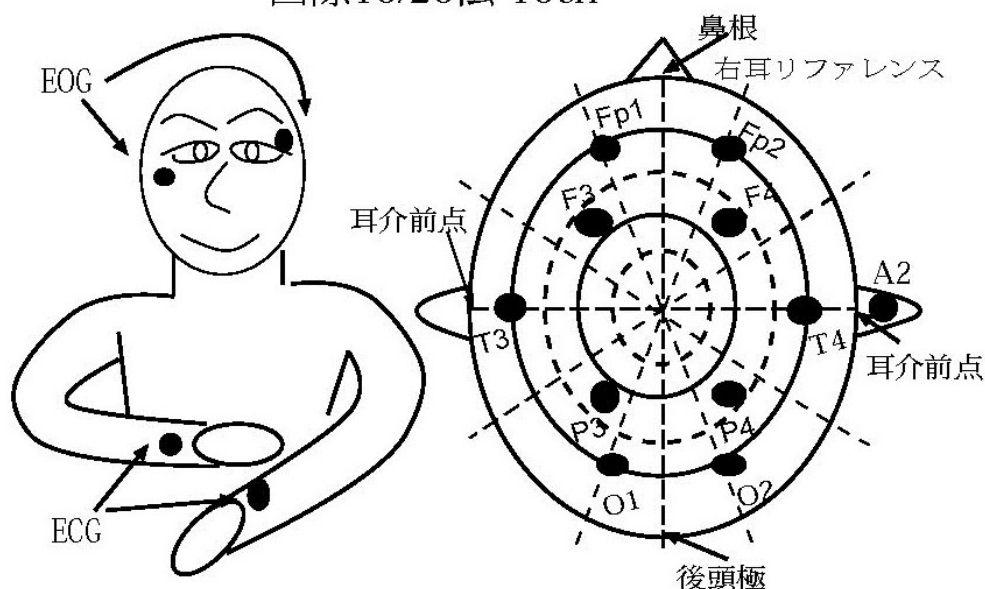
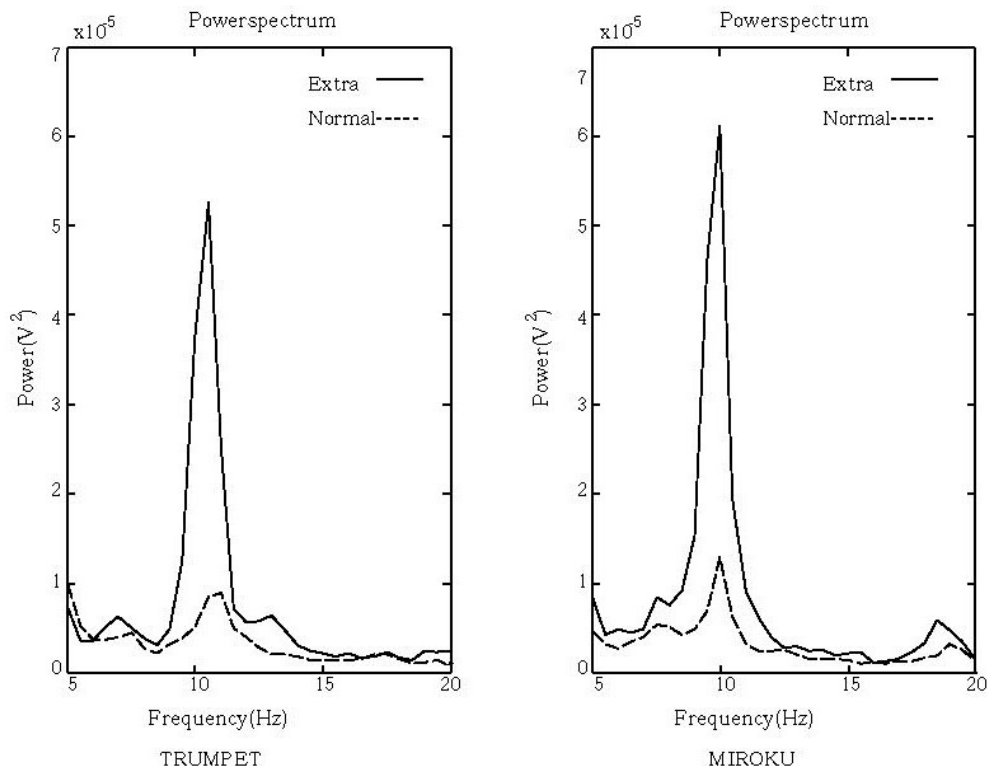


図 3.4-4 電極配置図



映像“TRUMPET”，映像“MIROKU”  
による被験者Cの場合。

図 3.4-5 脳波解析結果

図 3.4-5 の実験により得られたパワースペクトルの例を示す。アルファ帯域である 8～13Hz において、基準映像観察時よりも高品位映像観察時にパワーの増大が認められる。このような高品位映像観察時のアルファ帯域波パワーが増大する傾向を有意水準 5%で検定した。その結果、解析対象とした 192 件すべてにおける両側 t 検定では有意水準 5%で有意であった。各映像 (48 件) ごとの検定については映像「MIROKU」、映像「FACE」において有意水準 5%で有意であった。他の 2 枚については、有意水準 5%は満たしていないが、高品位映像観察時のほうが、アルファ帯域波パワーが高い傾向にあることが認められた。また、主観評価実験により得られた主観評価値と脳波測定実験により得られた客観評価値との相関係数 0.85 であり、比較的高い相関が得られた。したがって、有意水準に満たない映像についても、主観評価結果との相関は高い。

本研究では、高品位画像システムにより提示される R、G、B 各 4096 階調の再現が可能

である高品位映像について、脳波を指標とする客観評価法の開発を試みた。高品位映像と基準映像の間に定量的な差があることを差分電力により確かめ、同時に、変化位置を観察するために高品位映像と基準映像それぞれの微分映像を求め、差分処理により分析した。その結果、階調数の違いによる微妙な階調変化のある領域が観測された。そして、脳波を指標とした客観評価法の有効性を検証するために、高品位映像の物理要因・特性が主観評価実験結果、客観評価実験結果に与える影響をそれぞれ比較し、主観評価実験結果と客観評価実験結果の関係を求めた。その結果、分析された評価映像の性質と各評価実験結果との間に対応がある可能性が示唆された。基準映像に対して高品位映像は主観評価法、客観評価法においてともに高い評点を与える傾向にあることが示され、主観評価法と客観評価法は比較的高い相関関係にあることが示唆された。以上により、本研究で提案した脳波計測による客観評価法が高品位映像の評価として有効である可能性が示唆された。ただし、本実験は評価映像枚数が4枚という制約における評価実験であるため、高品位映像の選定方法については今後さらなる深い検討が必要である。

今後の課題として評価者および映像の選定方法を体系的に検討する予定である。本研究で提案した客観評価法が、高品位映像を必要とする深いアウェアネスや複合現実感などを含む新たな感性評価技術の基礎的基盤として貢献することを期待したい。

### 3.4.6 おわりに

本報告では、最近グループウェアや CSCW 研究における主要研究課題の1つであるアウェアネス研究について、報告者らの研究成果を中心に、その研究開発動向を述べた。アウェアネス研究は分散環境において、同期・対面環境以上の臨場感を提供するには如何にするかという問題提起から出発したアイコンタクト提供という研究から出発し、存在のアウェアネス、動作のアウェアネス、嗅覚・触覚のアウェアネス、形式知のみならず暗黙知のアウェアネス、メタ情報やメタ知識のアウェアネスの研究と進展しており、その研究開発の最前線は止まることを知らず、前進している。新世代グループウェアの設計者が対面環境では失われた臨場感を相手に提供したいとき、自分の使用するマルチメディア、ヒューマンメディアの特性を十二分に理解し、新しい世代のグループウェアのためのアウェアネス環境を構築しなければならない。

今後の課題として、美的、精神的に高度化した人々が、21世紀指向の Visual-Audio システムに要求する深い感性をノンコンタクトで客観的に評価する方法を、脳科学の知見を考慮して、模索していく必要がある。脳波、fMRI、PET、振動などの生体計測装置を駆使した分析法で得られた知見を、宮原プロジェクトの高品位 Visual-Audio システムの研究でこれまでに得られた知見と統合し、どのような深い感性を伝達したいかに応じた臨場感アウェアネスにあふれた新世代のグループウェアを構築したい。そのため、匂いの伝達や触

覚の伝達の研究をスタートしておくことが、深い感性情報の伝達には必要である。報告者らは感性から理性を除いた「深い感性」の空間にひそんでいる見えざる因子を発見し、それをデジタル技術に翻訳するという名人芸的プロセスを通して、臨場感、雰囲気、気配、熱気、凄みなどのアウェアネスを伝達できる 21 世紀型 Visual-Audio システムを構築できると確信している。そのような研究は 21 世紀なかばにおいて、情報ソムリエあるいは知識ソムリエの活躍するマルチメディア脳科学という学問分野に昇華するであろう。

#### 参考文献

- [1] P. Dourish and S. Bly: Supporting Awareness in a Distributed Work Group, in Proceedings of CHI'92, pp. 541-547, ACM, 1992.
- [2] 林秀彦, 國藤進, 宮原誠: 高品位映像の評価—脳波を指標とする客観評価法—, 映像情報メディア学会誌, 56, 6 (2002).
- [3] 本田新九郎, 富岡展也, 木村尚亮, 岡田謙一, 松下温: 在宅勤務者の疎外感の解消を実現した位置アウェアネス・アウェアネススペースに基づく仮想オフィス環境, 情報処理学会論文誌, Vol. 38, No. 7, pp. 1454-1464 (1997).
- [4] 伊藤禎宣, 國藤進: カンパセーションアウェアネス支援: カンパセーション状況の視覚化による新たなコミュニケーションツールの提案, 人工知能学会第 39 回人工知能基礎論研究会, 東京電機大学, AI 学会資料 SIG-FAI-9903, pp. 87-92 (Nov. 1999).
- [5] S. Ito and S. Kunifuji: Supporting Conversational Awareness in Text-based Conference System, Proceedings of KES'2000 Vol. 1, pp. 221-224, University of Brighton (Aug. 2000).
- [6] S. Ito, Y. Sumi, and K. Mase: Supporting Knowledge Sharing by Document Annotation at an Exhibition Site, Proceedings of the 15th Annual Conference of Japanese Society for Artificial Intelligence, Shimane Kenmin-Kaikan (May, 2001).
- [7] 伊藤禎宣, 角康之, 間瀬健二, 國藤進: SmartCourier: アノテーションを介した適応的情報共有環境, 人工知能論文誌, 17 (2002).
- [8] T. Ishida (ed.): Community Computing and Support Systems, Springer, Lecture Notes in Computer Science 1519 (1998).
- [9] 石井裕: グループウェアのデザイン, 共立出版 (1994).
- [10] 石井裕: Tangible Bits: 情報の感触/気配の伝達, 情報処理, Vol. 39, No. 8, pp. 745-751 (1998).
- [11] 門脇千恵, 爰川知宏, 山上俊彦, 杉田恵三, 國藤進: 情報取得アウェアネスによる組織情報の共有促進, 人工知能学会誌, Vol. 14, No. 1, pp. 111-121 (Jan. 1999).
- [12] 門脇千恵: 組織情報の共有過程分析に基づく情報共有促進支援法の研究, 北陸先端科学技術大学院大学情報科学研究科博士論文 (Mar. 1998).

- [13] 國藤進: オフィスにおける知的生産性向上のための知識創造方法論と知識創造支援ツール, 人工知能学会誌, Vol. 14, No. 1, pp. 50-57 (Jan. 1999).
- [14] 國藤進, 加藤直孝, 門脇千恵, 敷田幹文: ナレッジマネジメント時代のグループウェア, 日科技連出版社 (July 2001).
- [15] 松下温, 岡田謙一: コラボレーションとコミュニケーション, 共立出版 (1995).
- [16] 松浦宣彦, 日高哲雄, 岡田謙一, 松下温: VENUS: Interest Awareness を利用したインフォーマルコミュニケーション環境, 情報処理学会論文誌, Vol. 36, No. 6, pp. 1332-1342 (1995).
- [17] 宗森純, 宮内絵美, 牟田智宏, 吉野孝, 湯ノ口万友: 電子鬼ごっこ支援グループウェアの試作と適用, 情報処理学会グループウェア研究会, 2001-GW-39, pp. 25-30 (Mar. 2001).
- [18] 中川健一, 國藤進: アウェアネス支援に基づくリアルタイムな WWW コラボレーション環境の構築, 情報処理学会論文誌, Vol. 39, No. 10, pp. 2820-2827 (Oct. 1998).
- [19] 永野豊, 本田新九郎, 大澤隆治, 太田賢治, 重野寛, 岡田謙一, 松下温: 仮想空間内の風と香りの表現手法, 情報処理学会第 58 回全国大会 (Mar. 1999).
- [20] 西田豊明, 畦地真太郎, 藤原信彦, 角薫, 福原知宏, 矢野博之, 平田高志, 久保田秀和: パブリック・オピニオン・チャンネル, NTT コミュニケーション科学基礎研究所主催 CMCC 研究会第 2 回シンポジウム論文集, pp. 59-66, 奈良県新公会堂 (Sep. 1999).
- [21] 緒方広明, 矢野米雄: アウェアネスを指向した開放型グループ学習支援システム Sharlok の構築, 電子情報通信学会論文誌 D-II, Vol. J80-D-II, No. 4, pp. 874-883 (1997).
- [22] 坂本竜基, 國藤進: コラボレーションブラウジング: WWW アウェアネスを利用した新しいブラウジング方式の提案, 人工知能学会第 10 回"AI シンポジウム'99", 早稲田大学国際会議場, AI 学会資料 SIG-J-9901, pp. 97-102 (Dec. 1999).
- [23] R. Sakamoto and S. Kunifujii: Collaborative World Wide Web Browsing System through Supplement of Awarenesses, Proceedings of KES '2000, Vol. 1, pp. 233-236, University of Brighton (Aug. 2000).
- [24] R. Sakamoto and S. Kunifujii: A Visualization of Users' Contexts on Proxy Server, Proceedings of the 15th Annual Conference of Japanese Society for Artificial Intelligence, Shimane Kenmin-Kaikan (May 2001).
- [25] 坂本竜基, 中尾恵子, 角康之, 間瀬健二: 経験や行動履歴の漫画的要約表現, 2001 年度第 15 回人工知能学界全国大会, 島根県民会館 (May 2001).
- [26] R. Sakamoto, K. Nakao, Y. Sumi, and K. Mase: ComicDiary: Representing Individual Experiences through a Comic, SIGGRAPH 2001 (Aug. 2001).
- [27] 塩澤秀和, 西山晴彦, 松下温: 「納豆ビュー」の対話的な情報視覚化における位置づけ, 情報処理学会論文誌, Vol. 38, No. 11, pp. 2331-2342 (1997).

- [28] 山上俊彦, 関良明: Knowledge-awareness 指向のノウハウ伝播支援環境: CATFISH, 情報処理学会, 93-DPS-59-8, pp. 57-64 (1993).

## 3.5 音声情報処理の本格的アプリケーション構築に向けての展望と課題 —VoiceXML を中心として

報告者: 児島宏明委員

### 3.5.1 はじめに

近年、音声情報処理技術を応用した大規模で実用的なアプリケーションの 1 つとして、音声ポータル (Voice Portal) が、また、そのようなシステムを構築するための共通の基盤として VoiceXML が注目されている。本報告では、VoiceXML を中心として、音声対話を利用した大規模アプリケーションの可能性に関する展望と課題について述べる。

### 3.5.2 音声情報処理技術とアプリケーションの動向

音声認識技術に関しては、この 10 年間に、隠れマルコフモデル (HMM) と統計的言語モデル計モデルに基づいた手法が確立され、この枠組みに適合するアプリケーションについては、認識率が長足の進歩をとげた。しかし、適合しないアプリケーションについては、近年の重点課題として多くの研究者が取り組んできたにもかかわらず、本質的な解決となるような手法は見出されていない[1]。適合するアプリケーションとは、連鎖統計の学習に用いられるコーパス (多くの場合は新聞記事) から大きく外れないような文章が認識対象となるディクテーション (口述筆記) や、小規模な文法で記述できるような定型的な文章が認識対象となる限定的ドメインのタスクで、マイクや回線などの特性や雑音といった環境が変化しないようなアプリケーションである。適合しないアプリケーションとは、大きな多様性を含む話し言葉や、話題の移り変わる会話、高度な意味処理、多様で変化する雑音が存在する実環境、音響的な特徴が学習サンプルの統計的な中心から外れる老人や幼児や非母語話者、などを対象とするようなアプリケーションである。

このため、ここ 2~3 年間は、前者のように現在の技術に適合する範囲内において、その対象を拡大していくフェーズであったと言える。このような戦略は、音声認識の研究分野の内側から見れば、順調に成果をあげ、評価されていると言える。しかし、一般ユーザの評価は必ずしも高くない。例えば、Dataquest の報告によれば、50%のユーザが購入から 90 日以内に、パフォーマンスが期待通りでないことを理由に音声認識ソフトウェアを放棄しているという。このような状況を打開し、「使える」音声認識システムを実現するには、どうすればよいだろうか。

そのような方向性として、これまでの試みを集約した大規模で実用的なアプリケーションを構築することが、当面の目標の1つと考えられている。すなわち従来は、限定的な対象を扱った限定的なアプリケーションが、それぞれ独立に開発されてきたため、需要の拡大に限界があり、また、ユーザにとっても、ドメインが限定されることで、融通のきかない、使いにくいシステムという印象につながっていた。それらを、共通の基盤の上に構築し、統合して大規模化できるような枠組みがあれば、アプリケーションが大きく拡大する可能性があると考えられる。これは、インターネットの応用がHTMLおよびWWWの枠組みによって急激に拡大したことの、アナロジーとしてとらえることもできる。

このような共通基盤として注目されているのがVoiceXMLである。VoiceXML自体は、さまざまな音声システムに対応できる汎用性を備えているので、長期的には、上記のような展開が期待できるが、当面のアプリケーションとしては、音声ポータルを主なターゲットとしている。音声ポータルは、小規模な文法で済む場合が多い点や、回線の特性がある程度限定できる点(携帯電話など無線系を含むと多様性は増えるが)など、現在の技術水準とも適合性の高いアプリケーションといえる。

#### 3.5.3 音声ポータルの現状

前節で述べたように、長年にわたる改良にもかかわらず、音声認識システムはユーザにとって使いやすいものと認知されるに至らず、いくつかあった音声認識専門企業も、老舗の米Dragon SystemsがL&H(Lernout & Hauspie)に買収され、そのL&Hも従来音声を手がけていなかったScansoftに買収されるなど、厳しい状況が続いている。その一方で、業界紙Speech Recognition Updateによれば、音声認識ソフトウェア・装備・サービスの市場について、1997年には3億5600万ドル規模にすぎなかったものが2003年には226億ドル規模にまで急成長すると予測している。このような市場拡大の原動力と期待されているのが音声ポータルで、Cahners In-Stat Groupは、音声ポータルの売上規模が2005年に16億ドルに達するとの見通しを発表している。

音声ポータルは、主として電話を用いてサーバにアクセスし、音声認識や音声合成による対話を通じて、サーバやインターネット上の情報にアクセスできるようなシステムである。言い換えれば、「CTI(Computer Telephony Integration)」と呼ばれる電話による自動化システムに、音声対話による自動処理、すなわち「IVR(Interactive Voice Response)」を導入したものである。ここでは、従来のコールセンターのような企業ごとのサーバも含めたIVRシステムを、広い意味で「音声ポータル」と呼ぶことにする。

アプリケーションの例としては、以下のようなものがある。



- ・ ビジネス情報へのアクセス  
(例えば、企業のフロントデスク、企業間の自動発注システム、サポートデスク、旅行やホテルなどの予約システム、ホームバンキングなど)
- ・ 公共情報へのアクセス  
(例えば、天気、交通、学校、経済、ニュースなどの情報)
- ・ 個人情報へのアクセス  
(例えば、カレンダー、住所録、電話帳、備忘録など)
- ・ コミュニケーション  
(例えば、電子メールやボイスメールの送受信など)

音声ポータルに関するこのような強気の予測の根拠の1つとして、インターネットに接続されたパソコンの総数に比較して、通常の電話器の総数が圧倒的に大きいことが挙げられる。しかし実際には、単純に数量的な比較だけではなく、具体的なニーズを把握する必要がある。具体的な応用場面で期待される要因として、自動車の情報化と携帯電話の普及がある。運転中の自動車からインターネットにアクセスする手段としては、画面ベースのWWWブラウザよりも、電話による音声対話の方が適している場合が多い。また、携帯電話に関しては、日本ではi-modeのような文字ベースのインターネットアクセス手段が普及し、大画面の端末に人気があるのに対し、米国では通常の音声通話機能以外は、あまり受け入れられていない。これが、日本よりも米国において音声ポータルへの期待が高まっている大きな要因の1つと見ることもできる。これ以外にも、視覚に障害のあるユーザのインターネットへのアクセス手段としての意義もある。

このように、特に米国において、音声ポータルへの産業的期待は大きく高まり、これに関連した多くの企業が設立されている。音声ポータル専門の音声認識ソフトウェア開発会社としてNuance、SpeechWorksなどが、またポータル構築会社として、Tellme Networks、BeVocal、HeyAnita、VoiceGenieなどがある。1998年10月に、これらの企業が参加して電子商取引アプリケーションに採用される音声認識技術の標準化を目指して、「V-Commerce Alliance」が設立された。これがVoiceXMLの標準化につながっていくことになる。Nuanceは、SRI (Stanford Research Institute) からスピンアウトした技術者を中心に設立された企業で、1999年10月に音声対応ブラウザ「Voyager」を発表し、音声ポータル開発ツール「V-Builder」などを販売している。Tellmeは、AT&TやIntelなどと提携して、音声ポータルによる企業のコールセンターの自動化を進めると同時に、ポータル上の広告も収益源として想定している。これに対してBeVocalは、広告ではなく、ポータルやサーバ上のサービスへのアクセスに対応した手数料を、主な収益源として想定している。また、Qwest、Genuity、iBasis、ITXCなどの通信キャリアは、音声ポータルを導入して、自社のネットワークにトラフィックを引き込むことにより、通信収入の向上を期待している。このように、企業によって様々な異なるビジネスモデルで、音声ポータル

ルの事業化を進めている。

ほかにも、製品としては、SpeechWorks が、電話で証券の売買を行う「Tele\*Master」というシステムを構築し、米国のオンライン証券会社の E\*Trade で使われている。また、Lucent Technology の「PhoneBrowser」や、IBM の「WebSphere Voice Server」といった音声ポータルシステムが販売されている。また Intel は「Voice Portal Platform」、という音声ポータルサーバシステムを販売している。さらに、これまで音声認識に関して目立った動きのなかった Sun Microsystems も、2002 年 3 月になって「VoiceTone」と呼ぶサーバシステムを発表した。Micorsoft も、次節で述べるように SALT という規格を提唱して、Microsoft .NET (dot net) Speech SDK (Technical Preview) という開発ツールを発表している。

日本でも、NTT コミュニケーションズが「V ポータル」[2]と呼ぶ音声ポータルを運用し、2001 年 8 月から試験サービスを、2002 年 1 月から本格サービスを開始している。また、日本テレコムが「Voizi」[3]と呼ぶ音声ポータルの試験サービスを 2001 年 7 月から開始している。

#### 3.5.4 VoiceXML

##### 3.5.4.1 VoiceXML の背景と経緯

VoiceXML[4][5]は、前節で述べた音声ポータルに代表されるような、音声（音）による対話システムを作成するための、XML (eXtensible Markup Language) 仕様の記述言語である。

1995 年に、AT&T ベル研で PML (Phone Markup Language) が開発され、これに基づいて、AT&T の PhoneWeb や Lucent の TelePortal などのサービスや、Motorola の VoxML や、IBM の SpeechML などの記述言語が開発された。これらの成果を VoiceXML として統合し標準化することを目的として、1999 年 4 月にこれら 4 社を中心に、「VoiceXML Forum」[6]が結成された。また、W3C (World Wide Web Consortium) では、1998 年 10 月に、音声ブラウザ (Voice Browser) に関するワークショップを開き、1999 年 3 月に「音声ブラウザワーキンググループ (Voice Browser WG)」を結成した。VoiceXML Forum と、Voice Browser WG との関係は、今のところ Forum で仕様を検討して WG に提案し、WG が規格として承認する、という形態をとっている。日本では、2001 年 6 月に、沖電気工業、日立製作所、日本アイ・ビー・エムなどを中心に、「XML コンソーシアム VoiceXML 部会」[7]が結成された。Microsoft は、VoiceXML Forum に加入せず、VoiceXML とは距離を置いていたが、2001 年 10 月に SALT (Speech Application Language Tags) という新たな規格を提案し、「SALT Forum」を設立した[8]。Intel や Cisco などは、両方のフォーラムのメンバーとなっている。

VoiceXML の仕様は、VoiceXML Forum から 1999 年 8 月に Version 0.9 が、2000 年 3 月に Version 1.0 が発表され、2000 年 5 月に W3C の規格として承認された。現在の最新版は、2001 年 10 月に Version 2.0 のワーキングドラフトとして公開されている[9]。

#### 3.5.4.2 VoiceXML の概要

VoiceXML は、WWW ページを構築するのと同様の容易さで、音声応答アプリケーションを構築することを目的としている。VoiceXML で想定されているシステムの構成は、次のようになる。

- (a) PBX や VoIP ゲートウェイ。
- (b) VoiceXML インタプリタ。
- (c) ドキュメントサーバ。
- (d) インプリメンテーションプラットフォーム。

(a) は、ユーザからの電話を受けるための PBX や、これを計算機ネットワークにつなぐ VoIP (Voice over IP) ゲートウェイである。(b) は、VoiceXML を解釈し、実際に対話処理を実行する処理系である(「VoiceXML ブラウザ」とも呼ばれる)。(c) は記述された VoiceXML ドキュメントを保存したり、通常の HTML ベースの WWW リソースにリンクしたりするための WWW サーバである。(d) は、音声認識や音声合成を行う処理系や、時間切れなどのイベントを生成する機能などで、「VoiceXML インタプリタ」に付随する形にすることもできるし、別サーバとして接続する形で構成することもできる。「VoiceXML インタプリタ」と「ドキュメントサーバ」との間は、通常の HTTP プロトコルでやり取りされ、「ドキュメントサーバ」は通常の WWW サーバと同様の構成で済む点が、導入が容易という VoiceXML の利点につながっている。

VoiceXML の設計上の特徴としては、以下のような点が挙げられる。

- ・ 1 つのドキュメントに複数の対話処理を記述できるため、クライアント・サーバ間の通信を最小化できる。
- ・ 作成者が、機器の構成やハードウェア環境に依存した低レベルの処理を考慮しなくて済む。
- ・ ユーザの対話処理 (VoiceXML で記述) を、サービスの処理 (CGI で記述) から分離できる。
- ・ アプリケーションの移植が容易になる。
- ・ 簡単なメニューシステムから複雑な対話処理にまで対応できる。

VoiceXML が扱える入出力や操作には以下のようなものがある。

- 合成音声の出力。
- オーディオファイルの出力。
- 音声入力の認識。
- DTMF（ダイヤルトーン信号）入力の認識。
- 音声入力の録音。
- 対話の流れの制御。
- 電話の転送や切断など回線の制御。

### 3.5.4.3 VoiceXML の記法

VoiceXML の記述例を下に示す。

```
<?xml version="1.0"?>
<vxml version="2.0">
<menu>
  <prompt>
    Say one of: <enumerate/>
  </prompt>
  <choice next="http://www.sports.example/start.vxml">
    Sports
  </choice>
  <choice next="http://www.weather.example/intro.vxml">
    Weather
  </choice>
  <choice next="http://www.news.example/news.vxml">
    News
  </choice>
  <noinput>Please say one of <enumerate/></noinput>
</menu>
</vxml>
```

これは、電話サービスのジャンルを音声で選択する例で、<prompt>と</prompt>で囲まれた部分は、合成音声で出力される。会話の例は次のようになる。

サーバ（PC）："Say one of: Sports; Weather; News."  
 ユーザ（人）："Astrology."

サーバ (PC) : "I did not understand what you said."  
(システムの標準エラーメッセージ)  
サーバ (PC) : "Please say one of: Sports; Weather; News."  
ユーザ (人) : "Sports."  
サーバ (PC) : (<http://www.sports.example/start.vxml> へ飛ぶ)

また、次のように、音声の代わりに DTMF (ダイヤルトーン信号) で選択するように記述することもできる。

```
<menu>
  <property name="inputmodes" value="dtmf"/>
  <prompt>
    For sports press 1, For weather press 2, For News press 3.
  </prompt>
  <choice dtmf="1" next="http://www.sports.example/start.vxml"/>
  <choice dtmf="2" next="http://www.weather.example/intro.vxml"/>
  <choice dtmf="3" next="http://www.news.example/news.vxml"/>
</menu>
```

また、次の例のように、ユーザの回答を音声認識した結果を、変数に保存して参照することもできる。

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<vxml version="2.0" lang="en">
<form>
  <field name="city">
    <prompt>Where do you want to travel to?</prompt>
    <option>Edinburgh</option>
    <option>New York</option>
    <option>London</option>
    <option>Paris</option>
    <option>Stockholm</option>
  </field>
  <field name="travellers" type="number">
    <prompt>How many are travelling to <value expr="city"/>?</prompt>
  </field>
```

```

<block>
  <submit next="http://localhost/handler" namelist="city travellers"/>
</block>
</form>
</vxml>

```

この場合、ユーザの入力（例えば"London"）が参照されて、サーバは"How many are travelling to London?"のような合成音声音声を出力する。

#### 3.5.4.4 VoiceXML 関連の W3C 規格

このほかにも W3C では、VoiceXML に関連する規格をいくつか策定している。入力として複雑な文章を扱う場合には、「SRGF (Speech Recognition Grammar Format)」[10]と呼ばれる形式に従って、文脈自由文法に基づく文法を記述できる。このほかにも、単語連鎖統計の形で文法を規定する Stochastic (N-Gram) モデルに関する規格[11]も検討されているが、優先順位が低く策定は進んでいないようである。これらの文法に従って、入力文から構文解析木が生成される。この構文解析木から意味構造を抽出する際の記述言語として、「SI Tag (Semantic Interpretation Tag)」[12]と呼ばれるタグが策定されている。これは、ECMA Script[13] (JavaScript など) のサブセットとなっている。これとは別に、意味構造を XML 的に記述する NLSML (Natural Language Semantics Markup Language) [14]も策定しようとしているが、こちらはあまり進展していない。音声合成に関しては「SSML (Speech Synthesis Markup Language)」[15]が、発音辞書については、「Pronunciation Lexicon Markup Language」[16]が策定中である。また、回線の制御に関しては「CCXML (Call Control XML)」[17]が規定されている。このほかにも、「Multimodal」[18]や「Reusable dialog」[19]などが検討されてきたが、規格としてはまとまらなかったようである。

#### 3.5.5 音声アプリケーションの課題

上で述べたように、音声技術のキラーアプリケーションとして、音声ポータルに対する業界の期待は大きい。しかし、米国で音声ポータルの構築が本格的に始まってから、約 2 年、日本で約 1 年たったが、現在のところ、一般ユーザに広く利用されるまでに普及する兆しは、まだ見えていない。ただし、Allied Business Intelligence (ABI) 社の予測では、2006 年末までに音声ポータルのユーザ数は約 2 億 9,100 万人で、そのうち、それほど頻繁には利用しないユーザが 25%、比較的頻繁に利用するユーザが 40%、常習的なユーザが 35%としている。このような予測の背景として、狭い意味での他目的ポータルとしての利

用は、まだ目立つほどにはなっていないが、従来のコールセンターや、電話取引のようなシステムに、目立たない形で徐々に浸透している状況がある。例えば、前述の米 E\*Trade 証券の音声取引システムは、顧客の約 30%に利用されているという。

しかし、音声ポータルを含む IVR システムには、多くの課題が残されている。まず、画面ベースのブラウザのように、情報を一覧することができないことや、道順などのように、図や絵による表示のほうが適している情報への対応は、本質的な問題である。また、サーバから引き出した利用結果を保存する方法として、効率的な方法が見出されていない。また、ビジネスモデルとして、情報の間に広告を埋込む場合、利用者にとっては、画像の広告以上に煩わしいものとなる可能性が高い。また、一覧可能性とも関連して、分かりやすいメニューを設計することも課題となる。

一覧可能性の問題は、従来からある DTMF ベースの電話サービスにおいて、最もユーザにとってのフラストレーションとなっている問題であり、IVR 技術によってこれが解決されるならば、大きな意義がある。もし、音声ポータルにおいて人間のオペレータを相手にするような自然な対話の実現できれば、使いやすさは文字ベースのものを上回る可能性がある。しかし、これは技術的には本質的に困難な課題であり、10 年以内に実現できる見込みはないと考えられる。そのため、音声ポータルのメニュー設計によっては、DTMF ベースのメニューがそのまま単語音声認識に置換わっただけのものになりかねない。しかし、逆にこれは VoiceXML ベースのシステムの利点でもある。VoiceXML が DTMF 入力もサポートしていることにより、従来の DTMF での応答システムをそのまま、VoiceXML ベースのシステムで構築することができ、そこから段階的に、音声認識に置き換えていき、さらに入力として受けつけられる文法を複雑化して、徐々に自然な対話に近づけていくことができる。これは、そのときどきの技術水準に応じた使いやすい音声ポータルを構築できる可能性につながる。

そのほかに、大画面携帯電話や携帯情報端末 (PDA) などを想定し、音声だけでなく、文字や画像など、他のモダリティを併用する方向も重要である。上で述べた SALT は、このようなマルチモーダルをサポートを「売り」にしている。W3C でも、上で述べたように規格としてはいちど中止になっているが、マルチモーダルワーキンググループ[20]として研究が続けられており、XHTML や XForms や SMIL (Synchronized Multimedia Integration Language) などとの統合が検討されている。XHTML や XForms は、それぞれ HTML や入力フォームを拡張した規格であり、SMIL は、マルチメディアでのプレゼンテーションを主なターゲットとした規格である。日本でも VoiceXML へマルチモーダルを導入する研究が活発に行われている[21][22]。

以上のように、音声ポータルなど VoiceXML をベースとした IVR システムは、1990 年代のディクテーションシステムに続いて、この 10 年間の代表的な音声アプリケーションとなると考えられる。これを手がかりとして、音声アプリケーションの構築が促進され、大規模化や高度化が進むことが期待される。また、それにより社会慣習的な課題、すなわ

ち、音声入力システムに対する人々の慣れの問題も徐々に解消され、それに適合した新たな社会規範が形成されていくことが期待される。

#### 参考文献

- [1] 人間主体の知的情報処理技術に関する調査研究 IV, 日本情報処理開発協会先端情報技術研究所 (2001).
- [2] <http://www.ntt.com/v-portal/>
- [3] <http://www.voizi.net/>
- [4] Bob Edgar: The VoiceXML Handbook -- Understanding and Building the Phone-Enabled Web --, CMP Books (2001).
- [5] <http://www.w3.org/Voice/>
- [6] <http://www.voicexml.org/>
- [7] [http://www.xmlconsortium.org/wg/voice/voice\\_index.html](http://www.xmlconsortium.org/wg/voice/voice_index.html)
- [8] <http://www.saltforum.org/>
- [9] <http://www.w3.org/TR/2001/WD-voicexml20-20011023/>
- [10] <http://www.w3.org/TR/speech-grammar/>
- [11] <http://www.w3.org/TR/ngram-spec/>
- [12] <http://www.w3.org/TR/2001/WD-semantic-interpretation-20011116>
- [13] <http://www.ecma.ch/ecma1/STAND/ECMA-262.HTM>
- [14] <http://www.w3.org/TR/nl-spec/>
- [15] <http://www.w3.org/TR/speech-synthesis/>
- [16] <http://www.w3.org/TR/lexicon-reqs/>
- [17] <http://www.w3.org/TR/2002/WD-ccxml-20020221/>
- [18] <http://www.w3.org/TR/multimodal-reqs/>
- [19] <http://www.w3.org/TR/reusable-dialog-reqs/>
- [20] <http://www.w3.org/2002/mmi/>
- [21] 植田ほか: VoiceXML のマルチモーダル化の検討, 情報処理学会研究報告, Vol. 2001, No. 100, pp. 43-48 (2001).
- [22] 桂田ほか: 音声対話記述言語 VoiceXML と MMI 記述言語 XISL の比較, 情報処理学会研究報告, Vol. 2001, No. 100, pp. 49-54 (2001).



## 3.6 Digital Libraries, Metadata, Interoperability

報告者: 杉本重雄委員

### 3.6.1 はじめに

1990年代のインターネットと World Wide Web (WWW) の急速な発展により我々の情報環境は大きく変化した。我々は日常的に Google や Yahoo! を用いて必要な情報を探している。これは机の上から世界中で発信される情報にアクセスし、必要な情報を探すことを意味する。このようなことは 10 年前には考えられなかったことである。

Digital Library (電子図書館、デジタル図書館、ここではデジタルライブラリあるいは DL と記す<sup>1)</sup>) はインターネット上での重要な応用分野と位置づけられ、研究開発が積極的に進められてきた。こうした研究開発は大きく分けると新しい情報技術の開発を目的とする研究中心のものと、図書館などをベースにして実際的な DL 環境を提供するものに分けることができる。前者の代表格はアメリカの NSF 他の共同助成で進められた Digital Library Initiative (DLI) であろう[1]。DL という分野から容易に理解できることであるが、そこには多様な技術要素が含まれており、またそれらの統合技術も重要な要素となっている。そこで進められた研究は、図書館で直接そのまま利用できるシステムの開発を目指したものとは言えないが、応用分野と計算機科学・情報学分野を結びつけた新しい取り組みであり、かつ図書館に注目を集めたという点からは大きな役割を果たしたものであると言える。また、NSF による National SMETE Digital Library (NSDL)<sup>2)</sup> プログラムは、数学・科学技術・工学分野での教育・学習資源の開発とそれらの利用のための基盤の研究開発を目指している[2]。これは教育学習コンテンツの開発をもサポートするものであり、インターネット上に教育学習コンテンツとその利用環境(すなわち図書館)の実現を目的とするものである。一方、後者には大学図書館や国立図書館などで進められた新しい図書館環境を提供するものや、大学などの組織による新しい情報資源の蓄積と流通の取り組みがある。図書館による取り組みとして典型的なものにはアメリカ議会図書館の American Memory に代表される歴史資料・貴重資料のデジタル化のプロジェクトがある。大学図書館では、電子出版物の提供、資料の電子化と提供、インターネット情報資源へのアクセス支援、ネットワーク時代に適した新しい図書館環境の提供といった幅広い取り組みがな

---

<sup>1</sup> 「図書館」ということばから、従来の建物ないしは施設を基盤としたサービスを連想させることが多い。ここでは機能中心に述べるため「館」という連想を避けるため「デジタルライブラリ」と記している。

<sup>2</sup> SMETE は Science, Mathematics, Engineering and Technology Education に対応する。

されてきている。

ネットワーク上での情報資源の流通は、情報資源に関する情報の記述、すなわちメタデータの重要性を広く認識させることになった。例えば、教育・学習資源、ビデオデータ、地理的・空間的情報、電子商取引、資料保存というように、メタデータは多様な分野で開発が進められている。こうしたいろいろなメタデータ規則の中で、インターネット上での情報資源の発見を目的として Dublin Core Metadata Element Set (DCMES) と呼ばれるメタデータ規則の開発が進められてきた。DCMES は、多様な分野の多様な情報資源に関する記述を前提として開発されてきたメタデータ規則である。DCMES の取り組みにおける基本的かつ重要な視点は、多様なコミュニティで作り出される多様な資料（情報資源）に対して相互利用可能なメタデータのための規則を作り上げること、すなわち interoperability（相互運用性、相互利用性）であろう。

interoperability がインターネット上の情報流通、DL にとって重要であることは言うまでもない。例えば、1997-98 年に DL の研究戦略を議論するワーキンググループでも重要な研究トピックとして議論されている[3]。また、2001 年の Joint Conference on Digital Libraries (JCDL 2001)<sup>1</sup>での基調講演において、Clifford Lynch は DL における interoperability について述べた<sup>2</sup>。さまざまな草の根的活動から多様な情報資源が作り出されるインターネット環境においては、標準規格に基づくトップダウンで厳密な相互利用性だけでなく、ボトムアップでより緩やかな相互利用性が求められると考えられる。

以下、本報告では、3.6.2 節で 1990 年代に進められた DL に関するいろいろな研究開発活動について概観し、簡単にまとめる。3.6.3 節では、Dublin Core の解説とメタデータに関して今後求められると考えられる視点について述べる。3.6.4 節では interoperability に関して簡単に考察する。なお、報告者は本ワーキンググループのこれまでの報告書で DL やメタデータに関して多く述べてきた。そのため、本報告では詳しい内容については以前の報告を参照するのみにとどめる<sup>3</sup>。

## 3.6.2 Digital Library

### 3.6.2.1 DL のための新しい技術の研究開発の活動

1990 年代にはいくつもの DL あるいはそれに関連する領域の研究助成プログラムが欧米で進められた。前述の DLI はそうしたものの代表格で、1994 年からの 4 年間に進めら

---

<sup>1</sup> 以前は ACM と IEEE-CS がそれぞれ Digital Library に関する国際会議を行っていたが、2001 年からそれらが統合し JCDL となった。http://www.jcdl.org/参照。

<sup>2</sup> 米国 Coalition for Networked Information の Executive Director。講演題目は "Interoperability: the Still-Unfulfilled Promise of Networked Information"。

<sup>3</sup> 報告書の URL は、http://www.icot.or.jp/FTS/REPORTS/Report-index-J.html。

れた Phase 1 (DLI-1) と 1998 年から始められた Phase 2 (DLI-2) に分かれている。DLI-1 では規模の大きな 6 プロジェクトが進められたのに対し、DLI-2 では大きなものから小さなものまで数多くのプロジェクトが進められてきている。ヨーロッパでは、イギリスの eLib プログラムやドイツの GlobalInfo があるほか、EU での第 5 フレームワークの中の Information Societies Technology (IST) プログラムの中で研究助成が進められた。

DLI が持っていた特徴は、複数の研究助成母体による共同助成であったことと、計算科学、情報学の研究者、図書館や出版社といったコンテンツ所有者が参加するという異分野をつなぐプロジェクトであったことである。DLI-1 では、NSF のほか、DARPA と NASA が研究助成母体として参加した。一方、DLI-2 の場合、助成母体が NSF、DARPA、NASA に加えて、議会図書館 (LoC)、人文基金 (NEH)、医学図書館 (NLM) に広がり、自然科学・技術分野から医学、人文社会科学分野までカバーしている。DLI-1 は、計算機科学と情報学を基礎とする DL のための要素技術の研究を中心に進められた。一方、DLI-2 はよりコンテンツ指向を強めている。また、デジタル資料の長期保存といった図書館の基本機能に関するものもある。DLI-1 のアナウンスがなされたのは 1993 年の後半である。これは NCSA Mosaic が現れた時期と重なる。DLI-1 が開始されたのは、ちょうど WWW が爆発的に広がり始めた時期に重なる。これは、インターネットを用いて簡単に情報発信ができ、かつ大量の情報にアクセスができるということを、だれもが肌で感じることができるようになった時期であった。このことは DLI-1 にとって幸運であったと思われるが、その一方、激しいネットワーク環境の中で研究を進めなければならない、大変な時期であったとも言えよう。これまでの報告の中で DLI-1 および DLI-2 のプロジェクトについては触れているので、ここでは全体的な特徴についてのみ触れることにする<sup>1</sup>。

DL 研究は、情報検索や自然言語処理、文書処理、データベースなどさまざまな要素情報技術を総合的に利用するための新しい技術の開発であると言える。そこでは、要素技術そのものの重要性はもちろんのこと、それらを対象コンテンツと組み合わせること、また、利用者に適した総合的な利用環境を提供することが強く求められるものである。

DLI-1 の研究プロジェクト募集の資料[4][5]によると、研究領域が、(1) 任意の形式のデータを取り込むための技術、(2) 多様な情報に関する情報検索、フィルタリング、抄録作成などの技術、(3) 多様でかつ大量の情報を扱う DL のためのネットワークプロトコルなどの技術、となっており、基本的に技術中心であることがわかる。一方、DLI-2 の資料によると[6]、大枠として研究とテストベッド開発に分けられている。さらにこの「研究」が (1) 人間中心分野として、ユーザインタフェースや情報発見のためのソフトウェアやモ

---

<sup>1</sup> DLI-1 に関しては平成 8 年度の報告

<http://www.icot.or.jp/FTS/REPORTS/H8-reports/H08NAI-WG/H0903R1ch34.html#anchor-ch3-7>、DLI-2 および NSDL に関しては平成 12 年度の報告

<http://www.icot.or.jp/FTS/REPORTS/H12-reports/H1303-AITEC-Report4/AITEC0103R4.html/AITEC0103R4-ch3-6.htm> を参照。

デルなど、情報を作り出し、探し、利用するという人間の活動を助けるための研究、(2) コンテンツおよびコレクション中心分野として、多様なデータの取り込み、メタデータ、相互運用性、分野依存の対象とそれを扱う技術といった情報資源を扱うための研究、(3) システム中心の分野として、多様でかつ大量の情報資源を多様な利用者対象と利用者環境において提供するための研究、の3分野に分けられている。また、DLI-2では国際間での共同プロジェクトを推進した。DLI-2では、違う視点を持つ研究者を集めることをより推進したと言える。これは1997年に開かれたSanta FeでのワークショップのDLに関する研究戦略の議論の結果を反映したものである[7]。

DLI-2と並行して、NSFではDLに関連する分野の研究助成として2つのプログラムが始められた。ひとつはInformation Technology Research (ITR) [8]であり、情報技術一般の研究を対象とする非常に大きなプログラムである。もうひとつがNational SMETE Digital Library (NSDL)である。DLI-2の一部はNSDLのプロジェクトとして進められている。インターネットの教育利用に関しては、K-12<sup>1</sup>として広く進められてきており、大学学部レベルでの教育とあわせてK-16といわれることもある。現在、あちこちの大学で遠隔学習プログラムの提供が進んでおり、インターネットを利用した学習、教育は重要なトピックとなっている。そうした状況においてNSDLの果たす役割は大きいと感じる。

#### 3.6.2.2 図書館を基礎としたデジタルライブラリの開発

従来、研究は大学や研究所に所属しないと困難であるとされてきた。研究に必要な資料を持つ図書館が近くにないと研究が進められない、というのがその理由の1つである。図書館（あるいは図書館のようなサービス）が研究者にとって重要であることは、従来とまったく変化していない。しかしながら、インターネット上での情報流通の発達によって、図書館に行く前に（あるいは行かずに）、インターネットで情報を探すということが一般化してきていると思える。もちろん、行動パターンの違いは分野や個人に依存するところが大きいことは言うまでもないが、こうした傾向は先端的な自然科学・技術分野でより顕著であるように感じられる。こうした環境下、この10年の間に図書館での電子的な情報環境は、大学図書館を中心として大きく変化してきた。現在では、多くの図書館がホームページを持ち、図書検索サービスなどを提供している。図書館におけるDLサービスの典型的な機能を下に並べる。

- (1) 電子的に出版される資料の提供（典型的なものは電子雑誌やデータベース）
- (2) 貴重資料や歴史資料の電子化と提供
- (3) 大学など図書館が所属する組織・機関から出版されるさまざまな情報資源の電子

---

<sup>1</sup> K-12はKindergarten to 12th Gradeを意味し、幼稚園から高校までの教育におけるインターネット利用が進められている。

## 化とアクセスの支援

- (4) インターネット上の情報資源へのアクセス支援
- (5) ネットワークを介したレファレンスサービス
- (6) 電子化資料の長期保存
- (7) 図書館内外における DL 機能を利用した総合的な利用環境の整備

以上の項目に関して簡単に考察したい。

### (1) 電子的に出版される資料の提供

電子的に出版される資料の提供は、出版物を利用者に提供するという意味では図書館としてごく基本的な機能である。学術雑誌の電子出版の発展とともに最近では数千タイトルの雑誌を提供する大学図書館も普通になってきている。一方、小規模の学会などでは、独自に電子ジャーナルに対応することが必ずしも容易ではない。遡及的な電子化や電子ジャーナルの長期保存などの問題がある。

### (2) 貴重資料や歴史資料の電子化と提供

貴重資料や歴史資料の電子化は、図書館がもつ貴重資料そのものへのアクセスを制限する一方、電子的な複製物によりアクセス性を飛躍的に高めるという、「保存とアクセス」の両方の観点から進められてきている。また、電子化はこれからも進められていくと思われる。一方、電子化された資料が実際に役に立つかどうかについても十分に考慮されないといけない。単に、電子化しインターネット上で提供するだけではなく、教育利用での利用性の促進や評価、DLI-2に見られるような新しい情報技術との組み合わせによる新しい分野の開拓といった活動が求められる。

### (3) 大学などから出版されるさまざまな情報資源の電子化とアクセスの支援

大学などの組織が出版する情報資源の電子化とアクセス支援は発信とアクセスの両面での技術が求められる。例えば、資料の電子化のために適切な技術を選ぶこと、資料に適切なメタデータを与え組織化することが求められる。また、発信される情報を再利用するための技術も求められる。大学や研究所といった組織に限らず、電子政府のような、巨大な情報発信元を想定してみるとこうした機能の必要性が理解しやすい。

### (4) インターネット上の情報資源へのアクセス支援

インターネット上での情報資源アクセスを効率的に行うためのポータルサイトあるいはゲートウェイサイトの重要性は広く認められている。実際にこうしたサービスを実現するには、「有用な情報資源」あるいは「高品質な情報資源」を見つけ、それらに対して利用者にとって有用な情報を持つ高品質なメタデータを記述し、それらを適切にメンテナンスすることが求められる。しかしながら、現状ではこれらの多くが人手に頼らざるを得ない。

### (5) ネットワークを介したレファレンスサービス

レファレンスサービスは、図書館員によるサービスであり、これをネットワーク上で提

供しようという取り組みが進められている[9][10]。レファレンスサービスには、質問者が何を聞きたいのかを明確化することなども必要とされる人間主体の作業である。そのため、サービスそのものを機械化することは困難であると思えるが、レファレンスサービスを支えるための環境を作り上げることが、DLの研究開発にとって重要であることは疑えない。

#### (6) 電子化資料の長期保存

資料の長期保存は図書館の機能としては基本的なものである。ところが、情報媒体技術の進化の激しさは、電子媒体に蓄積された資料、特にもともと電子的に蓄積された資料（Born Digital）の長期にわたる保存を難しくしている。アメリカ議会図書館の場合、デジタル資料の長期保存が重要課題として報告されており[11]、また議会からそうした課題を中心としてDLをさらに開発するための大きな予算を与えられている。

#### (7) 図書館内外におけるDL機能を利用した総合的な利用環境の整備

図書館内の計算機とネットワークの環境は、この10年で大きく変わってしまったように思われる。公共図書館ではまだこれからであるかもしれないが、大学図書館では、端末用のパソコンがずらっと並んでいることはもう珍しくない。図書館内外（あるいはキャンパス内外）での利用環境の違いを少なくしていくための技術や制度が求められると思われる。

全体を通して、この10年間でずいぶん図書館および利用者の情報環境は変化したことは間違いない。一方、諸外国と比較してわが国を見た場合、図書館とそれ以外、特に研究開発部門との連携が弱いと感じられる。また図書館間での連携も弱いように感じる。研究開発能力を高めるためには必要な情報が容易に手に入る必要があることは言うまでもない。例えば、アメリカの場合、Digital Library Federationという組織が早くから作られ、情報の共有やプロジェクトを進める母体にもなっている。わが国を見ると、コンテンツをより使いやすくするための環境整備に十分な手当がなされていないように感じる。

#### 3.6.2.3 デジタルライブラリに関するその他の話題、取り組み

インターネット環境における新しい取り組みとしていくつか興味深いものがあるのでそれを紹介したい。最初のもは、プレプリント、テクニカルレポート、学位論文などを電子的に作られた資源を蓄積提供するサービス（リポジトリ）による協調プロジェクトであるOpen Archives Initiative（OAI）である[12]。これは別個に作り上げられてきたサービス間でコンテンツを協調的に提供する環境を作り出すことを目指しており、例えば資源に関するメタデータの収集（metadata harvesting）のためのプロトコルを決めている。このプロジェクトは、複数のリポジトリによる横断的な利用が可能になるのみならず、収集したメタデータを利用した付加価値サービスの可能性も含んでいる。

上のようなサービスの広がりや学術情報の流通経路を変える可能性を持っている。学術雑誌に掲載される論文が学術情報の中心であることは疑えない。ところが、学術雑誌の価

格が高く、かつ学術領域の広がりとともに雑誌の数が増えていくことは大学図書館と研究者にとっては頭の痛い問題である。研究者にとっては学術雑誌に論文を掲載することが自身の業績評価につながる一方、研究者の大学の図書館がその雑誌を購入していない、あるいは購入の継続を打ち切ってしまい自分の論文が掲載された雑誌を図書館で読めない状況も出てくる可能性がある<sup>1</sup>。その一方、WWW上に無料で公開されるテクニカルレポートや学位論文などの学術情報資源はどんどん増えている。査読システムによって quality は保証されるが手に入れにくい（高価な、あるいは限られた場所でしか利用することのできない）雑誌に掲載された論文と、quality は自分で判断しなければならないが WWW 上で容易に手に入れることのできる論文のどちらが本当に役に立つのか、といった問題に関する議論が必要なのであろう。

もう1つの話題は、電子資料の長期保存に関する取り組みである。これまで図書館や文書館では紙の資料を中心とした資料保存を行ってきた。先に述べた貴重資料の電子化のように「保存とアクセス」の両面から資料の電子化が進められる一方、電子媒体による出版の発展は電子出版物の保存という観点から、これまで以上に困難な問題をもたらしている。納本図書館である国立図書館、わが国の場合は国立国会図書館にとって、これは非常に大きな問題である。電子出版物は大きく分けて CD-ROM などのパッケージに入れて出版されるものと、ネットワーク上で出版されるものがある。前者をパッケージ系出版物、後者をネットワーク系出版物と呼ぶことにする。パッケージ系出版物の場合、媒体技術の進歩が早いために蓄積媒体を利用するためのハードウェアがなくなってしまう可能性に対処しなければならない。パッケージ系、ネットワーク系いずれの場合も、コンテンツの利用に特別のソフトウェアあるいは利用環境が必要とされる場合、そうしたソフトウェア環境を稼動可能な状態で保存しつづけることが難しい。また、WWW上で提供される資料の場合、どのような資料を保存の対象とすべきかを定めること（保存対象範囲の定義）が必ずしも容易ではない。また、電子ジャーナルの場合、印刷物と異なり出版社はライセンスのみを提供することになるので実体の保存に對しだれが責任を持つのかといった問題が生じる。電子資料の保存に関してはコンテンツの内容的問題、知的財産権にかかわる問題など、社会制度的な問題もあるが、技術的にも克服しなければならない点が多くある。次に、電子資料の保存に関するいくつかの活動を紹介してみたい。

(1) DLI-2 の中で進められているミシガン大学とリーズ大学（イギリス）の共同プロジェクト CAMiLEON は Emulation をベースにした電子資料の保存方法の開発を目的と

---

<sup>1</sup> 図書館における電子出版物（特に電子ジャーナル）の導入、提供などに関する種々の問題に関して知識と情報を共有し、意見交換をする取り組みが進められている。例えば、アメリカを中心とする国際的な取り組みである SPARC (the Scholarly Publishing and Academic Resource Coalition, <http://www.arl.org/sparc/>)、わが国での国立大学図書館協議会による取り組みなどがある。図書館での現実的な対応に関してはここでは触れない。

している[13]。また、コーネル大学のプロジェクト PRISM でも電子化資料の保存を扱っている[14]。

(2) スタンフォード大学を中心として進められているプロジェクト LOCKSS (Lots of Copies Keep Stuff Safe) は、電子ジャーナルの保存のためにジャーナルへのアクセスのために用いられる PC 上にコピーを残す仕組みを作り、それをネットワーク上の多数の PC で動作させることによって、実質的に資料を保存するという取り組みを進めている[15]。

(3) 北欧、オーストラリア他の国立図書館において、WWW 上の出版物を収集し、保存する取り組みを進めている[16]。図書館によって、内容による選択的収集を行っている館と、内容によらない網羅的収集を行っている館がある。

(4) Alexa Internet 社では、インターネット上の資料のアーカイブ作りを目的として WWW 上の資料をロボットで収集し、蓄積・提供するサービス Internet Archive を進めている[17]。ここでは、例えば「1999 年の WWW 上の資料」を探すことのできるサービス Wayback Machine を提供している。しかしながら、動的に作成されるページやデータベースの中でしかアクセスできないページなどロボットによる自動収集の制約を含んでいる。

#### 3.6.3 メタデータについて

##### 3.6.3.1 メタデータとは

メタデータは簡単には「データに関するデータ (Data about Data)」と定義される。この定義に基づくと、目録、索引、抄録、シソーラスといったものから、資料の識別子、書評や内容の Rating まですべてメタデータに含めて考えることができる。利用者はインターネットや DL 上で情報資源を見つけ出し、その内容を評価し、利用する。メタデータなしにはこうした操作を行うことは困難である。メタデータは従来の目録と同じものと(乱暴には)言うこともできるが、メタデータは内容とメディアの多様化、利用者と利用環境の変化などに対応するべく提案されてきているものである。例えば、図書館での目録は、基本的には、1 冊の本、1 タイトルの雑誌に対して付与されている。これは、図書館での資料の取り扱いの単位に対応する。一方、インターネット上では、単純にとらえると URI が与えられる資料単位が操作の基本単位ということになり、図書館での資料とインターネット情報資源の違いが良く理解できる。また、メタデータに関して、記述の視点と利用の視点(例えば、情報資源の発見)の違いに留意する必要がある。記述の視点からは、できるだけ正確かつ詳細に対象資源に関する記述をすることが望まれる。利用の視点からは、一般の利用者による情報資源検索には詳細な情報が利用されることは少ない。また、異なる規則・基準に基づいて作られたメタデータを横断的に検索する場合やそれらを収集して利用する (metadata harvesting) 場合には詳細情報は必ずしも用いられない。

現在、応用に応じていろいろなメタデータ規則が提案されている。その一例を下に示す。



- Dublin Core Metadata Element Set: インターネット上の情報資源の発見を目的として定義されたもの[18] (昨年度までの報告書などに詳しい解説がある)。
- IEEE Learning Object Metadata (IEEE LOM) : 学習、教育情報資源に関するメタデータ。IMS、Ariadneなどで遠隔教育・学習分野で利用されている[19][20][21][22]。
- MPEG-7: オーディオ・ビデオデータのためのメタデータ[23]

ほかに、政府行政情報、地理情報、環境情報、電子商取引、倫理的内容などいろいろな分野でのメタデータ開発が進められている。

### 3.6.3.2 Dublin Core Metadata Element Set について

Dublin Core Metadata Element Set (DCMES) はインターネット上での情報資源の発見を目的として開発が進められてきたもので、現在いろいろな分野で応用が進められている。図書館や美術館・博物館での利用、教育情報資源への適用などが進められてきているほか、ヨーロッパ各国での政府情報のためのメタデータとして利用されている<sup>1</sup>。欧米を中心として標準規格化が進んでいる<sup>2</sup>。わが国でも大学図書館を中心として利用が進んできており、国立国会図書館、国立情報学研究所でもネットワーク情報資源のためのメタデータ規則として適用している[24]。また、昨年10月東京で第9回のDublin Core Workshopを、規模を拡大した国際会議として開催した<sup>3</sup>。

報告者は、これまでDublin Coreは昨年度までの報告書などで紹介してきたのでここでは詳しく述べず、Dublin Coreの基本的な考え方と利用の現状について述べる。なお、基本となる15エレメントについては表3.6-1に示す。

#### \* 基本概念

##### (1) 「情報資源の発見」のためのメタデータ

DCMESは情報資源を発見するためのメタデータ、すなわち利用者が情報資源を探すための情報資源の記述データを提供することが本来の目的である。「情報資源の記述」という観点から見ると、できるだけ詳細に定義されたメタデータエレメントや厳密な意味定義のなされた統制語彙に基づく記述が望ましいように思われる。しかしながら、利用者の観点からは、情報資源を探すために詳細な記述を用いることは多くない。そのため、情報資源の発見の観点から、できるだけ基本的な内容のエレメントからなるメタデータ記述規則を

<sup>1</sup> DCMIのホームページからプロジェクト (<http://dublincore.org/projects/>) や政府情報のためのメタデータとしての採用例 (<http://dublincore.org/news/adoption/>) を見ることができる。

<sup>2</sup> ヨーロッパ: CEN/ISSS CWA13874。アメリカ: ANSI Z39.85。

<sup>3</sup> <http://www.nii.ac.jp/dc2001/>に会議の論文集がある。

### 3 研究開発の新しい展開と内外の動向

決める努力がなされてきた。

#### (2) コア・メタデータ・エレメント・セット (Core Metadata Element Set)

報告者はコア (Core) ということばが大きな意味を持っていると理解している。ネットワーク上にはさまざまな分野の、さまざまな種類の、さまざまな目的で作られた情報資源がある。これらに共通に適用できるメタデータ記述規則を作るには、それぞれの分野や目的で必要とされるエレメントをすべて含むように定義する方法と、すべての分野や目的で共通に利用可能であると考えられるエレメントのみを含むように定義する方法がある。前者の方法をとると巨大な規則を作ることになり、実現性に疑問がある。Dublin Core は後者の方法をとったものである。

表 3.6-1 Dublin Core の基本 15 エレメント

| 要素名               | identifier  | 定義および説明                              |
|-------------------|-------------|--------------------------------------|
| タイトル              | Title       | 情報資源に与えられた名前。                        |
| 作成者               | Creator     | 情報資源の内容の作成に主たる責任を持つ実体。               |
| 主題およびキーワード        | Subject     | 情報資源の内容のトピック。                        |
| 内容記述              | Description | 情報資源の内容の記述。                          |
| 公開者 (出版者)         | Publisher   | 情報資源を利用可能にすることに対して責任を持つ実体。           |
| 寄与者               | Contributor | 情報資源の内容への寄与に対して責任を持つ実体。              |
| 日付                | Date        | 情報資源のライフサイクルにおける何らかの事象に対して関連づけられた日付。 |
| 資源タイプ             | Type        | 情報資源の内容の性質もしくはジャンル。                  |
| 形式                | Format      | 物理的表現形式ないしデジタル形式での表現形式。              |
| 資源識別子             | Identifier  | 与えられた環境において一意に定まる情報資源に対する参照。         |
| 情報源 (出处)          | Source      | 現在の情報資源が作り出される源になった情報資源への参照。         |
| 言語                | Language    | 当該情報資源の内容の言語。                        |
| 関係                | Relation    | 関連情報資源への参照。                          |
| 対象範囲<br>(空間的・時間的) | Coverage    | 情報資源の内容が表す範囲あるいは領域。                  |
| 権利管理              | Rights      | 情報資源に含まれる、ないしはかかわる権利に関する情報。          |

### (3) Simple Dublin Core (Simple DC) と Qualified Dublin Core (Qualified DC)

Dublin Core の開発において、はじめ 13 の基本エレメントからなるエレメントセットが合意され、その後 15 エレメントに定義しなおされた。これが Simple DC と呼ばれるもので、現在までに ANSI などの標準規格となっているものである。一方、この開発過程で、詳細な記述のための限定子 (qualifier) に関する議論が行われてきた。この過程では、「エレメントをより詳細に表す」といっても基本的な性質が異なるものがあることが明確化されてきた。例えば、作成者を詳細 (あるいは正確に) 表す方法として、正確な記述のための典拠ファイルのレコード名など作成者に与えられたユニークな識別子で記述する方法と、作成者の名前、所属、連絡先などの組で記述する方法がある。前者の場合にはどのような典拠ファイルを使っているのか、あるいはどのような統制語彙を用いて表しているのかを示す必要がある。後者の場合、作成者である人 (あるいは組織) を表す値の構造、すなわち下位のエレメント (サブエレメント) の定義が必要である。このほか、「詳細に表す」ということには、「作成者」の意味が小説や論文の「著者」であるのか、「作曲家」あるいは「作詞家」であるのかといった、エレメントの意味の詳細化の場合がある。Dublin Core では値の表現形式に関する指定 (すなわち統制語彙やデータ形式) を指示するもの (Encoding Scheme qualifier、コード化スキーム限定子) と、エレメントの意味の詳細化を表すもの (Element Refinement qualifier、エレメント詳細化限定子) の 2 種類の限定子を持っている。一方、上のサブエレメントに対応するものは限定子としては含まれていない。これは、作成者に関するサブエレメントを、人を表すためのエレメントセットと理解できる。すなわち、サブエレメントの場合は、別個のエレメントセットとして定義することができる。このようにして定義された限定子を含む DCMES が Qualified DC と呼ばれる。限定子の定義については平成 12 年度の報告書に示しているのでここでは省略する。

### (4) Warwick Framework

1996 年春にイギリスのウォーリックで開催された第 2 回のワークショップで提案された Warwick Framework は Dublin Core の位置づけとネットワーク上でのメタデータのとらえかたの上で大きな役割を果たしている。Warwick Framework は図 3.6-1 に示す概念的なモデルを与える。Warwick Framework は複数のメタデータ規則による記述 (パッケージと呼ぶ) をひとつの入れ物 (コンテナ) に入れ、まとめるというものである。MARC などの従来目録規則では、それ自身で閉じたものとして定義されてきたが、インターネット上での情報資源や利用方法の多様性に対応するには、単一の規則のみで記述することをせず、いくつかの記述規則を利用してひとつのメタデータを記述する枠組みを与えている。この考え方は、WWW コンソーシアムの Resource Description Framework (RDF) の考え方に影響を与えている。

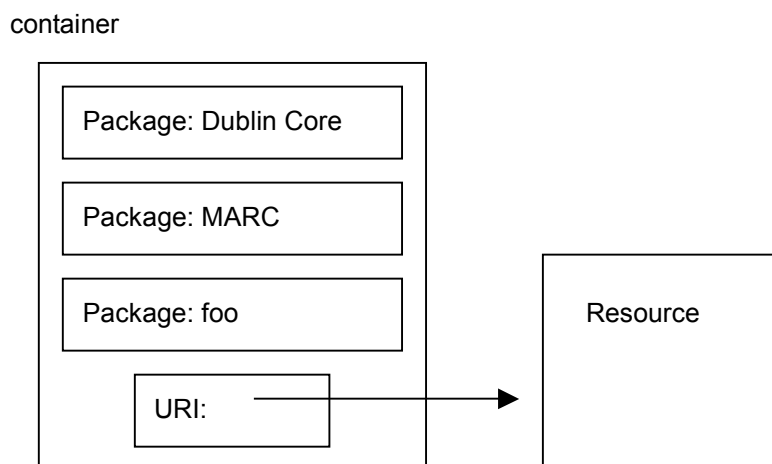


図 3.6-1 Warwick Framework

#### \* Dublin Core の利用

##### (5) Dublin Core の適用方法

現在、DCMI 推奨となっているメタデータの記述要素は Simple DC の 15 エレメントと DCMI 推奨限定子、および DCMI Type Vocabulary である。一方、これまで Dublin Core はいろいろな応用に利用されてきている。こうした応用では、推奨エレメント・限定子に加え、それぞれの目的に則して独自の拡張がなされていることが多い。例えば、国立国会図書館のメタデータ規則の場合、主題の記述のために NDC が限定子として利用されているが、NDC は DCMI の推奨限定子には含まれてはいない。また、Open Archives Initiative に見られるように、異なるメタデータ規則の下に作られたメタデータの共通部分として Dublin Core を利用することもある。一般に、メタデータの記述の際にはできるだけ「高品質な記述」が求められる。「高品質な記述」には、できるだけ詳細であること、できるだけ内容の記述が統制されていることが含意されている。一方、前にも述べたように情報資源の発見の側からは、あまりに厳密な規則はエンドユーザーには意味のないものである。

##### (6) Dublin Core の具体的な記述形式

DCMI では具体的な記述形式の定義を HTML 以外には積極的には進めてこなかった。その理由は、記述形式の定義を進めることによってエレメントの意味定義に影響が出る可能性を避けたためである。一方、Dublin Core を利用して作られた応用サービスでは、それぞれのシステム構成に応じた表現方法が用いられている。現在、DCMI では、XML および RDF での記述形式の定義を進めている。

##### (7) 新しいエレメントや限定子の導入

上に述べたように実際の応用においては、必要に応じたエレメントや限定子の導入がな

されている。一方、DCMI の中でも新しいエレメントや限定子の導入が進められている。例えば、教育分野での利用に関して、どのような利用者に適した資源であるのかを記述するための **audience** エレメントやそれに関連する限定子が提案されている。政府情報や図書館への適用に関しても同様に議論が進められている。

#### (8) エレメントセットとアプリケーションプロファイル

Simple DC では 15 のエレメントの定義に加えて、「すべてのエレメントは省略可能 (Optional) であり、かつ繰り返し可能 (Repeatable) である」と定義している。これに対して図書館での Dublin Core の利用に関するワーキンググループ (DC-Libraries) では、図書館での利用では、例えば「タイトルエレメントは必須、主題エレメントは適用可能な場合必須、内容エレメントは記述を推奨」というように提案している。こうした応用ごとの適用規則をアプリケーションプロファイル (application profile) と呼んでいる。そこで、メタデータ規則をエレメントセットとアプリケーションプロファイルに分けて定義しなおすことにする。エレメントセットはエレメントや限定子など、メタデータとして対象情報資源の性質を記述するための語彙とその意味を定めるものであり、アプリケーションプロファイルはエレメントセットを対象分野に適用する際に定める適用規則である。例えば、Simple DC の定義は、15 エレメントの名前とその意味の定義に加え、「すべてのエレメントは省略可能 (Optional) であり、かつ繰り返し可能 (Repeatable) である」としている。この場合、15 エレメントの定義がエレメントセット、15 エレメントに関して省略可能かつ繰り返し可能と定義している部分がアプリケーションプロファイルということができる。また、アプリケーションプロファイルは複数のエレメントセットをもとにして定義することも可能である (図 3.6-2 参照)。

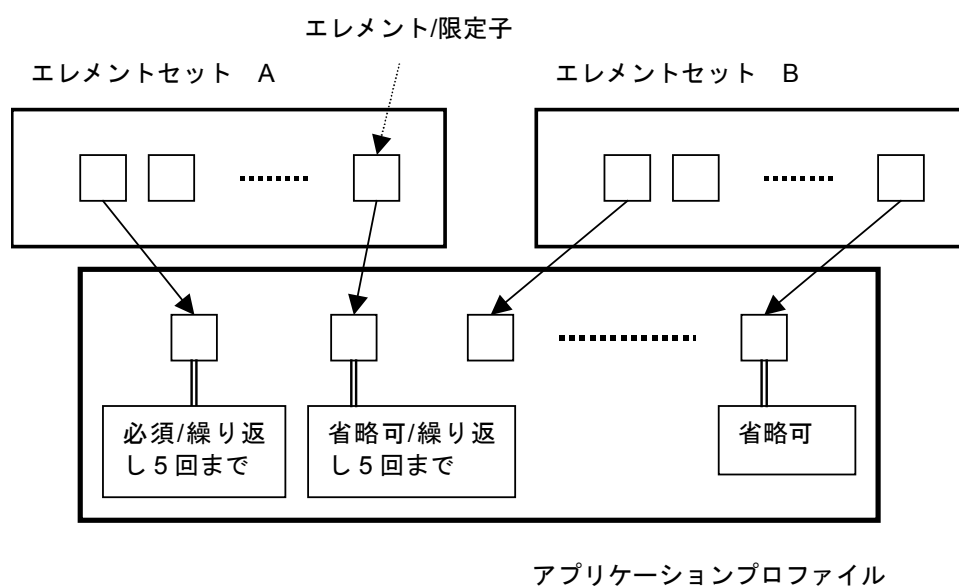


図 3.6-2 エレメントセットとアプリケーションプロファイル

#### \* Dublin Core の開発と維持

##### (9) Dublin Core Metadata Initiative (DCMI)

Dublin Core の開発は、OCLC による支援の下に草の根参加者によって進められてきた。Dublin Core の開発の初期の段階では OCLC や NSF 他の助成の下で進められてきたが、Simple DC の合意が得られ標準化を進める必要が出てきたころから DCMES の開発と維持の活動を行う組織を明確にする必要が出てきた。そのため、Dublin Core 開発を支えてきた OCLC がホストする形で DCMI と呼ばれる組織が 1999 年ごろから構成されてきた。その後 Dublin Core の利用がさまざまな分野で広がるにつれ、新しいエレメントや限定子の導入をも含めた DCMES の維持、地域ごとのコミュニティによる Dublin Core の Localization などいろいろな活動の必要性がでてきた。こうした将来の活動に向け、現在 DCMI の再編成が進められている。

##### (10) 限定子の拡張と Dumb-Down (ダム・ダウン) 原則

上に示したように、分野や応用に応じた限定子が導入されている。限定子が無秩序に定義されるのを防ぎ、Dublin Core としての相互利用性を保つため、DCMI では限定子の導入に関する基本原則を決めている。これは Dumb-Down 原則と呼ばれるもので、「限定子付きの表現から限定子を取り除いても、当該エレメントとその値の間で矛盾が生じてはならない」というものである。例えば、「有効になった日付 (Date Valid) は 3 月 20 日である」(エレメント: 日付 (Date)、限定子: 有効になった (Valid)、値: 3 月 20 日) から「有効になった (Valid)」という限定子を取り除いて「日付 (Date) は 3 月 20 日である」としても、矛盾は生じない。一方、サブエレメントの場合を考えてみよう。例えば、「作成者の名前は杉本重雄、所属は図書館情報大学」(エレメント: 作成者 (Creator)、限定子: 名前と所属、値: 杉本重雄と図書館情報大学) の場合、限定子を取り除くと「作成者は杉本重雄、図書館情報大学」となり、図書館情報大学が作成者になってしまうので矛盾する。この例から、先に述べた値の構造を表す限定子 (サブエレメント) が一般的には Dumb-down 可能ではないことが理解でき、Qualified DC の中に含まれていない理由がわかる。

##### (11) 分野や地域によるコミュニティ

これまでの開発においては DCMI の場での標準作りと応用ごとの DCMES の適用が行われてきたが、今後は政府情報や教育情報といった応用分野の要求に基づく開発、地域ごとの要求、特に英語圏以外の要求に基づく開発が進むと考えられる。また、DCMES の各国語への翻訳、各国での標準の維持なども求められる。こうした活動を進めるには分野や地域ごとのコミュニティを作り上げる必要がある。

##### (12) わが国での状況

わが国では、大学図書館を中心に Dublin Core の利用が広まってきており、日本語で書かれた情報資源に対するメタデータの記述方法に関する知見も蓄積されてきている。また、

Dublin Core の国際会議 (DC-2001) として東京の国立情報学研究所で開催した。しかしながら、現時点では JIS による標準化はまだ行われていない。日本語あるいは日本の情報資源のためのエレメントや限定子などに関する議論や意見交換を行うための組織作りが求められている。

### 3.6.3.3 メタデータ規則の利用

ネットワーク上には多様な情報資源が提供されると同時に、それらの利用目的、利用方法、利用環境も多様である。このことは、ひとつの情報資源であってもいろいろな視点からメタデータを付与することの必要性を意味する。例えば、教材として利用するビデオ資料の場合、教育資料としてのメタデータとビデオデータとしてのメタデータの両方が求められる。ネットワーク上での情報資料検索の場合、政府関係資料や地理情報資料というように分野を特化して利用する場合はそれぞれの資料向けのメタデータが望ましいが、広い範囲から資料を探す場合は Dublin Core のようなコアメタデータが望ましい。また、メタデータの記述対象のとらえ方もさまざまである。Dublin Core の場合は、情報資源を対象とし、情報資源を扱う主体は視野の外である。IFLA の FRBR モデルでは、対象を *work* (知的生産物のもともとの内容)、*expression* (知的生産物を何らかの表現手法により表したものの)、*manifestation* (出版物として知的生産物を実体化したものの)、*item* (出版物の 1 点) の 4 段階でとらえている[25]。電子商取引用の *indecis* の場合、人 (*people*)、物 (*stuff*)、取引 (*deal*) が記述対象となる[26]。Lagoze らによる ABC metadata model ではイベントを重要な要素としてとらえている[27]。

ネットワーク上ではいくつものデータベースに対する横断的な検索が有用である。例えば、Z39.50 は分散環境上での横断的な検索に広く利用されているプロトコルである。横断検索を実現するには、共通のメタデータ規則の下に検索質問を受けつけることが必要であり、そのため、メタデータ規則間あるいはメタデータエレメント間の対応関係の定義 (Crosswalk) が必要とされる。Crosswalk をできるだけ機械的に作成、維持することが求められる。

WWW コンソーシアムによる Resource Description Framework は WWW 上でのメタデータ記述のための共通基盤を提供するための取り組みである。RDF は Model and Syntax と Schema の 2 つに分かれている。前者はメタデータ記述のための構文とその基になるモデルを定義し、後者はメタデータのエレメントや限定子などメタデータ規則に含まれる語彙の定義を与える仕組みを定義している。その後の Semantic Web の取り組みは、WWW 上での情報流通のための意味的基盤を与えようというものであるが、現時点での中心は、RDF Schema によるメタデータを表す属性の語彙の定義を含め、WWW 上での情報交換のための語彙の定義 (ontology あるいは vocabulary) が中心的な話題となっているようである。Dublin Core においても、メタデータを記述するための統制された語彙に関する議論は多くあり、こうした問題は古くて新しい問題の 1 つのようである。

先に述べたようにネットワーク上でいろいろなメタデータ規則が用いられる。Dublin Core の範囲内でもいろいろな拡張がなされている。メタデータの流通性、相互利用性を高めるには、メタデータ規則に関する情報の流通性を高めることが求められる。そのため、メタデータ規則を登録し、提供するサービスであるメタデータスキーマレジストリ（あるいは、メタデータレジストリ）の重要性が認められている。例えば、EU の Schemas プロジェクトでは多数のメタデータ規則を収集、登録している[28]。図書館情報大学では DCMES を中心として多言語によるメタデータレジストリを実現している[29]。DCMI でもレジストリに関するワーキンググループを作り、DCMI レジストリの構築を進めている。メタデータレジストリは、単にエレメントや限定子の定義の参照記述を人間が見るためだけに限らず、エレメントや限定子に対し、ネットワーク上で一意に定まる識別子を与えることでエレメントなどの意味定義の基底を与えること、エレメントなどを長期にわたって維持管理するための基盤となることが期待される。さらに、新しいエレメントや限定子を定義する上で参考となる情報を与えるといった付加価値サービスを行うことができると考えられる。

#### 3.6.4 Interoperability

Interoperability（相互運用性、相互利用性）は DL における重要なトピックのひとつである。1997 年から 98 年にかけて進められた EU と NSF の共同助成による DL の研究戦略を議論するワーキンググループにおいても interoperability は主要なトピックとして議論がなされた[30]。Interoperability に関するワーキンググループの報告書[31]の中で階層的なモデルを提案し、ネットワーク上に分散した DL 間での interoperability を実現するための課題について議論している。

Dublin Core では Semantic Interoperability ということばキーワードとして用いられてきた。直訳すると意味的相互運用性である。これは「特定の表現形式に依存することなくメタデータを記述するエレメントの意味（だけ）を決め、意味的な相互利用性を保とう」というものである。Dublin Core では、HTML による基本的な記述形式を除いては具体的な表現形式に関する議論はほとんど行われてこなかった。表現形式に関する議論を避けてきた背景には、意味と表現形式の議論は必ずしもうまく切り分けられないためエレメント定義を進める過程において表現形式に関する議論に引っ張られることを避ける必要があったこと、具体的な表現形式はシステムの実現方法に負うところが大きいことがある。また、エレメントの意味定義に関する合意が形式定義のために損なわれることを避けたためであるとも言える。

2001 年 6 月に開催された Joint Conference on Digital Libraries 2001 (JCDDL '01) の Clifford Lynch による基調講演は interoperability をテーマとするものであった。同氏は



講演の中で、Dumb-Down (ダムダウン) 原則と Open Archives Initiative (OAI) に触れていた。Dumb-Down 原則は、詳細レベルでの interoperability が保証できないときに記述の詳細度を落とすことで interoperability を守るための原則を与えている。また、OAI の metadata harvesting は、個別に定義されたメタデータを横断的に収集するためのプロトコルを決めている。そこでは共通のエレメントを決める必要があり、Dublin Core を基礎とするエレメントを決めている。

Lynch の基調講演の際に、ミシガン大学教授の Margaret Hedstrom から「Digital Library における重要な課題であるデジタル資料の保存 (Preservation) は時間的相互運用性 (Temporal Interoperability) と言ってもよいのではないか。」という指摘があり、Lynch もそれに同意していた。デジタル資料の保存には、何らかの媒体に蓄積された資料の保存だけでなく、資料を利用するための技術も保存する必要がある。ところが、技術の進歩の速さは技術の保存を困難にしている。資料を利用するために必要なハードウェアやソフトウェア技術が消えてしまうことによる資料の実質的消失は大きな問題として広く認識されている。

Hedstrom は「マイクロフィルムによる保存は原資料の機能をすべて保存しているものではなく、それと同じことがデジタル資料にも当てはまってもよいのではないか。作られたときとまったく同じ状態で、技術とともに保存しなければならないものばかりとは限らない」と述べている<sup>1</sup>。デジタル資料の保存のために、時を越えてどのような精度の interoperability を求めるかが現実的な課題であることが理解できる。資料保存というと、1 次コンテンツをいかに保存するかのみ注目が向きがちであるが、利用された技術、コンテンツの保存履歴などメタデータとして記録し、保存する必要があることは言うまでもない。

メタデータのコミュニティからは ontology ということが良く聞こえてくる。大規模知識ベースなどの AI 指向のコミュニティで続けてこられたものが、ネットワーク上での意味的な interoperability の必要性とともに (再び?) 注目されたものと理解している。上の EU-NSF ワーキンググループでも Information Model として ontology に関する議論をしている。一方、もともと図書館分野では統制語彙が広く用いられてきたので、それ自身は新しいひびきはあまりない。しかしながら、これまでの統制語彙 (あるいはそのために作られたシソーラスや辞書) が、いわば閉じた世界での利用を前提としていたのに対し、ネットワーク上では異なる語彙の基に作られたデータ (あるいはデータベース) 間での interoperability、いわば開かれた世界での interoperability がより重要になる。

話題を Dublin Core に戻したい。Dublin Core の議論の中で、interoperability の観点から重要であると考えられるトピックに、上に示した Dumb-down 原則のほかにエレメントセットとアプリケーションプロファイルの分離とメタデータスキーマレジストリがある。

---

<sup>1</sup> 図書館情報大学で開催した第20回デジタル図書館ワークショップでの講演による[30]。

エレメントセットとアプリケーションプロファイルを別々に考えることはエレメントおよび限定子の意味定義と、応用に適用する際の記述上の規約を分離したととらえることができる。両者を分離することで Crosswalk の作成や Dumb-down のための基礎となるメタデータ規則の定義がより明確にできる。先に示したように、メタデータスキーマレジストリはメタデータの interoperability のために重要な役割を持っている。Dublin Core のレジストリの場合で考えると、namespace の管理、各国語訳と英語で記されたオリジナルの記述との間の管理、継続的な開発に伴う版の更新といった点でレジストリの果たす役割は大きい。

現在のインターネットでは、さまざまなコミュニティによって多様な情報資源が発信され、蓄積されている。言語、文化、分野の違いを越えて相互に情報資源を利用できるようにすること、また時間を越えて情報資源の利用性を保つことの重要性は明らかである。コミュニティごとの高い自由度を保ちつつ、コミュニティ間の interoperability が求められると言える。しかしながら、これを実現することは必ずしも容易なことではない。従来 interoperability を得るためには詳細に決められたプロトコルに従うことが前提であったように思うが、Dumb-down 原則や temporal interoperability の議論にもあるように、緩やかな相互利用性・相互運用性についても理解を深める必要があるように感じる。

#### 3.6.5 おわりに

国家情報基盤 (National Information Infrastructure)、世界情報基盤 (Global Information Infrastructure) ということばとともに 1990 年代のはじめから半ばにかけて、DL に関する研究開発が大きな期待とともに始められた。WWW の爆発的な広がりとともに、我々の情報環境は急激に変化した。例えば、現在、報告者自身の研究活動はインターネットおよびインターネット上のいろいろなツールなしにはまったく成り立たない。DL という分野がインターネットに密接に関連したものであることが理由の 1 つではあろうが、研究者一般についてかなり共通していることであるとも思う。本報告の執筆にあたって、雑誌や論文集の論文、種々の報告書など、ほとんどインターネット経由で得ることができた。その中にはかなり 1992-93 年に作られたものも有り、この分野では「古い」資料も含まれている。

図書館現場での DL サービスもこの 5、6 年でかなり進んだと思われる。例えば、大学図書館で提供される電子ジャーナルのタイトル数も最近では非常に増加した。それでも「電子図書館 = 歴史資料の電子化」というイメージがいまだに強いようにも感じる。このイメージだけでとらえられたままでは、DL としては失敗なのであろう。図書館 (あるいはそれに似た機能を持つ組織) はネットワーク上での出版、情報流通の増加に対応した中継ぎとしてのサービスを提供することが基本であると思う。その一方で図書館自身による情

報資源の開発と提供が求められていることも確かである。求められる機能として重要なものに、サブジェクトゲートウェイやネットワーク上でのリファレンスサービスなど情報資源へのアクセス支援サービスがある。こうしたサービスは人手中心に進められている一方、こうしたサービスの実現を支援するための情報技術は十分であるとはいえない。先に、述べた資料の長期保存技術とともにこうした情報技術の開発が望まれる。

DL 分野はいろいろな要素情報技術を総合的に応用することが求められる分野である。EU-NSF 共同の DL の研究戦略に関するワーキンググループで選ばれた 5 分野 (知的財産権、Interoperability、情報発見、メタデータ、多言語情報アクセス) は、グループの活動から 4 年ほどたつが、これらが情報技術の観点から見た DL の重要な基盤分野であることは今も変わっていないと思う。DLI-1 と DLI-2 を比較すると、DLI-1 では各プロジェクトの中でいろいろな要素情報技術の研究が見られたのに対し、DLI-2 の方がよりコンテンツ指向であると考えられる。これは特定の応用分野のコンテンツとそれに適した情報技術開発を組み合わせた中小規模のプロジェクトが多いためでもあろう。また、DLI-2 にはデジタル資料の長期保存やサブジェクトゲートウェイ関連の技術開発など、DLI-1 が行われていた間の WWW の発展と電子的な情報資源の急激な増加によって必要性が明確になったものも含まれている。こうした点が DL 分野の特徴を表しているともいえよう。また、教育・学習 (e-Learning、遠隔学習、遠隔教育)、電子政府、電子商取引など隣接分野での DL のための技術や知識の適用も重要な話題であると考えられる。

最後に、メタデータ分野の観点から、DL において重要と思われる点をいくつか挙げてみたい。

#### (1) 多様な interoperability とメタデータ規則の流通

ネットワーク上では、それぞれの応用にあわせて規則を決めることのできる自由度が求められる一方、地理的距離、時間的距離、言語や文化の違い、分野の違いを超えてネットワーク上で相互にデータやサービスを利用し合うための共通性が同時に求められる。そこではいろいろなレベルでの相互利用性が求められることになる。同時に、そうした interoperability を支えるにはメタデータ規則 (メタデータスキーマ) をネットワーク上で流通させる仕組みが重要である。

#### (2) メタデータを作り、維持するための技術

現時点では高品質であることを求められるメタデータの作成には人手がかかる。また、そうしたメタデータを作成する対象を選定すること、すなわちある程度のコストをかけてもメタデータを作る価値がある情報資源であると判断することにもコストがかかる。加えて、ネットワーク情報資源の特性として内容の更新が容易であるという点がある。この点はメタデータの維持管理を難しくしている。こうした点に対する情報技術開発が望まれる。

#### (3) 多言語メタデータ

言うまでもなくインターネット上では多様な言語が利用されており、多言語情報アクセ

スのための情報技術はメタデータの作成や利用においても重要である。

#### 参考文献

- [1] Digital Library Initiatives, <http://www.dli2.nsf.gov/> (DLI-2 のホームページ。DLI-1 へのリンクもある)。
- [2] NSDL, <http://www.ehr.nsf.gov/EHR/DUE/programs/nsdl/> (<http://www.smete.org/> も参考)。
- [3] NSF-EU Collaboration, <http://www.iei.pi.cnr.it/DELOS/NSF/nsf.htm>
- [4] NSF 93-141 - Research on Digital Libraries, <http://www.nsf.gov/pubs/stis1993/nsf93141/nsf93141.txt> (Sep. 1993)。
- [5] 杉本重雄: デジタル図書館へのアプローチ - DL 関連研究分野に関して, デジタル図書館, No. 3, [http://www.dl.ulis.ac.jp/DLjournal/No\\_3/05-sugimoto.html](http://www.dl.ulis.ac.jp/DLjournal/No_3/05-sugimoto.html) (Mar. 1995)。
- [6] DIGITAL LIBRARIES INITIATIVE - PHASE II, <http://www.nsf.gov/pubs/1998/nsf9863/nsf9863.htm> (1998)。
- [7] Report of the Santa Fe Planning Workshop on Distributed Knowledge Work Environment: Digital Libraries, <http://www.si.umich.edu/SantaFe/> (Sep. 1997)。
- [8] Information Technology Research, <http://www.itr.nsf.gov/>
- [9] Internet Public Library, <http://www.ipl.org/>
- [10] Collaborative Digital Reference Service, [http://www.dl.ulis.ac.jp/DLjournal/No\\_3/05-sugimoto.html](http://www.dl.ulis.ac.jp/DLjournal/No_3/05-sugimoto.html)
- [11] Committee on an Information Technology Strategy for the Library of Congress, Computer Science and Telecommunications Board, National Research Council, LC21: A Digital Strategy for the Library of Congress, 238p., <http://books.nap.edu/html/lc21/> (2000)。
- [12] Open Archives Initiative, <http://www.openarchives.org/>
- [13] CAMiLEON - Creating Creative Archiving at Michigan and Leeds: Emulating the Old on the New, <http://www.si.umich.edu/CAMILEON/>
- [14] Information Integrity in Distributed Digital Libraries, <http://www.prism.cornell.edu/>
- [15] LOCKSS Web Site, <http://lockss.stanford.edu/index.html>
- [16] PADI - Preserving Access to Digital Information (National Approaches), <http://www.nla.gov.au/padi/topics/65.html>
- [17] Internet Archive, <http://www.archive.org/index.html>
- [18] Dublin Core Metadata initiative, <http://dublincore.org/>

- [19] ARIADNE - Alliance of Remote Instructional Authoring and Distribution Networks for Europe, <http://www.ariadne-eu.org/>
- [20] IMS Global Learning Consortium, Inc., <http://www.imsproject.org/>
- [21] 先端学習基盤協議会, <http://www.alic.gr.jp/>
- [22] Greenberg, J. (ed.): Metadata and Organizing Educational Resources on the Internet, Haworth Press, 302p. (2000).
- [23] The MPEG Home Page, <http://mpeg.telecomitalia.com/>
- [24] デジタル図書館, No. 22 (Mar. 2002) (<http://www.dl.ulis.ac.jp/DLjournal/>に掲載予定).
- [25] IFLA, Functional Requirements for Bibliographic Records, <http://www.ifla.org/VII/s13/frbr/frbr.pdf> (1998).
- [26] Rust, G.: The <indec> metadata framework, <http://www.indec.org/pdf/framework.pdf> (June 2000).
- [27] Lagoze, C. and Hunter, J.: The ABC Ontology and Model, Journal of Digital Information, Vol. 2, issue 2, <http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Lagoze/> (Nov. 2001).
- [28] Forum for Metadata Schema Implementers, <http://www.schemas-forum.org/>
- [29] Nagamori M., et al.: A Multilingual Metadata Schema Registry Based on RDF Schema, Proceedings of International Conference on Dublin Core and Metadata Applications 2001, <http://www.nii.ac.jp/dc2001/proceedings/product/paper-31.pdf>
- [30] An International Research Agenda for Digital Libraries - Summary Report of the Series of Joint NSF-EU Working Groups on Future Directions for Digital Libraries Research, <http://www.iei.pi.cnr.it/DELOS/NSF/Brussrep.htm> (Oct. 1998).
- [31] Birmingham, W., et al.: EU-NSF Digital Library Working Group on Interoperability between Digital Libraries, <http://www.iei.pi.cnr.it/DELOS/NSF/interop.htm> (Oct. 1998).
- [32] Margaret Hedstrom: Digital Preservation: Problems and Prospects, デジタル図書館, No. 20, pp. 3-15, [http://www.dl.ulis.ac.jp/DLjournal/No\\_20/1-hedstrom/1-hedstrom.html](http://www.dl.ulis.ac.jp/DLjournal/No_20/1-hedstrom/1-hedstrom.html) (Mar. 2001).

## 3.7 デジタルコンテンツのトランスコーディングとインタフェースロボット

報告者: 長尾確委員

### 3.7.1 デジタルコンテンツのトランスコーディング

デジタルコンテンツがあたりまえのものとして世の中に溢れ出したのは 20 世紀の情報技術の進歩からすると必然的であっただろう。そして、それら膨大なコンテンツを活用するための技術もさまざまなものが発明され、進歩をとげていくことは間違いがない。これまでは、ともかくコンテンツを作成して流通させることが主目的であったのに対し、これからは、それらのコンテンツをいかに賢く利用するか、あるいは、いかに多様に、多目的に利用するか、ということが最も重要な課題になると思われる。

コンテンツ技術には、デジタルコンテンツを作成・保存・伝達する技術のほかに変換・加工・再利用する技術も含まれる。本報告では、特に後者に力点を置いている。

デジタルコンテンツの変換・加工・再利用の形態には、例えば、パーソナライゼーションとアダプテーションがある。デジタル放送の映像や Web ページなどのデジタルコンテンツをユーザの好みに応じて変換することをパーソナライゼーションと呼び、それらのコンテンツを PC や PDA (Personal Digital Assistant) や携帯電話などのデバイスの特性に合わせて変換することをアダプテーションと呼ぶ。

本報告では、デジタルコンテンツのパーソナライゼーションとアダプテーションを合わせたものをトランスコーディングと呼ぶ。現状では、インターネットへのアクセスは PC 経由で行われることが多い。しかし、この様相は近年、急激に変わりつつある。PC に加えて、携帯電話や PDA、テレビ、カーナビなどを使ってインターネットにアクセスする機会がますます増加するだろう。このとき重要となるものがトランスコーディングである。例えば、PC で表示することを前提にして作成した Web ページを携帯電話などで表示する場合、画像の縮小やテキスト部分の圧縮といった操作を自動的に行う必要がある。トランスコーディングには、少ない伝送容量を使ってサーバからクライアントにコンテンツを配信できるという利点のほかに、ユーザの嗜好に応じた理解しやすいコンテンツを生成できるといった利点がある。

コンテンツプロバイダやサービスプロバイダは、それぞれの機器に対応したコンテンツを個別に用意しなくても済む。具体的な応用例としては、PC 向け Web コンテンツのトランスコーディングによって、携帯電話向けのコンテンツを生成するといった利用法がある。

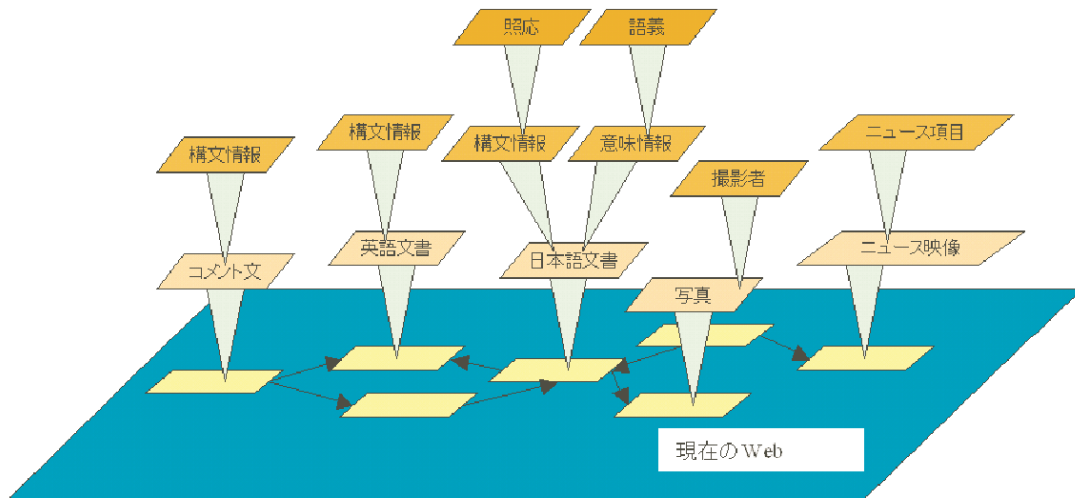


図 3.7-1 アノテーションによる Web コンテンツの拡張

### 3.7.2 セマンティックトランスコーディング

このトランスコーディングをさらに進めて、テキストの要約などの内容に基づく処理の精度を高める工夫を盛り込んだのが、報告者の提案するセマンティックトランスコーディングである[8]。具体的には、コンテンツに含まれるテキスト文要素に言語的なアノテーション（補足情報）を加えることによって、要約や翻訳などの自然言語処理の精度を大きく向上させることができる。例えば、アノテーションによってコンテンツに含まれるテキスト文の意味を明確にすると、正確な要約や翻訳が期待できる。コンテンツにアノテーションをつける手間が増すが、重要な情報にはアノテーションをつけて正しく伝達し、共有すべきという考えに基づいている。このアノテーションはコンテンツの内容理解を促進するものと位置づけられる。現在、報告者らは原著者を含む多くの人々が文書の内容に関する補足的情報を付加できるような枠組み作りや、その情報を加味して文書を読者に適した形に加工する仕組み作りに取り組んでいる。

図 3.7-1 で示されるように、アノテーションは、現在の Web に上位構造を作る基盤になる。現在の Web コンテンツが最下層で、アノテーションはコンテンツに情報をつけ加えるメタ（上位）コンテンツ、さらにメタコンテンツに対するメタコンテンツのように階層をなしている。

セマンティックトランスコーディングは、ユーザが指定した Web 上の新聞記事などのコンテンツを任意の圧縮率で要約して表示したり、テレビ番組などの映像データからユーザ

の好みに応じた話題だけを抜き出して、ダイジェスト映像を作成したりといったことを可能にする。さらに、要約したコンテンツを翻訳したり、テキストを音声化して聴いたりすることもできる。コンテンツサーバにおかれたテキスト、画像、音声、映像などのコンテンツはトランスコーディングプロキシによって、ユーザの使用するデバイス（PC、携帯電話、カーナビなど）や、ユーザの要求（概要をつかみたい、母国語で読みたい、声で聞きたい、など）に合わせて加工される。このとき、アノテーションと呼ばれる付加情報を用いて、より精度の高い要約・翻訳を行う。アノテーションはアノテーションサーバに蓄えられている。

セマンティックトランスコーディングの手法を使って、具体的には HTML 文書などの Web コンテンツがかかえる、以下の 3 つの課題を解消できるだろう。

- HTML (HyperText Markup Language) ではレイアウトなどの文書の表現については規定している。しかし、文書の意味などといった内容に関してはほとんど何も規定していない。この点を改善するために SemanticWeb という構想が、World Wide Web Consortium (W3C) において議論されている[12]。
- HTML などで記述したハイパーテキストは、各文書間のネットワーク構造を記述できる。ただしリンク情報が常に正しいとは限らず、その修正ができるのはもとの文書の著者だけである。
- Web 文書の著者は一般にその読者のことを考慮して著作してはいない。なおかつ著者と読者の間に立って吟味・調整する役割の人間も通常はいない。

Web は、新しいスタイルの文書のあり方を示したという点において革新的だったと言えるだろう。Web コンテンツの自由度の高さは疑いようがない。しかし、現状では Web コンテンツを読者が読みやすいような体裁に機械的に変換することは非常に困難である。

図 3.7-1 にあるように、従来の Web コンテンツは 1 枚の平面上に存在する要素群としてとらえることができる。セマンティックトランスコーディングでは、Web コンテンツを平面から立体に拡張する手法を提案する。コンテンツの各要素に意味や文書構造を示すアノテーションを付加する。このことによって Web コンテンツに、コンテンツの各要素の意味や文書構造を記述した上位構造を築くことができる。代表的なアノテーションの例としては、リンク元の文書に埋め込まれていないハイパーリンクである外部リンクや、コンテンツに対するコメントなどが挙げられる。アノテーションを作成して公開することが容易になれば、Web コンテンツの表現力は大幅に高まり、その利用価値が飛躍的に向上するだろう。

セマンティックトランスコーディングは、基本的にテキストコンテンツの処理を中心としたものであるが、その手法は映像や画像などの非テキストコンテンツの加工にも応用され、マルチメディアデータを含むコンテンツに適用できる。



### 3.7.3 アノテーションとトランスコーディング

さて、アノテーションあるいはメタデータに対するこれまでのさまざまな取り組みでは、コンテンツを高度化するという目的は共有しているが、コンテンツの検索や分類以外の具体的な目的への利用可能性についてあまり考慮されていない。これに対して報告者らのアプローチでは、特にトランスコーディングへの応用を目指し、いくつかのシステムをすでに実装し、実験を進めている。

セマンティックトランスコーディングで用いられるアノテーションには、大きく分けて3つの種類がある。テキスト文の言語構造などを付与する言語的アノテーション、画像やハイパーリンクなどのコンテンツを構成する各要素に対するコメントアノテーション、ビデオ映像などマルチメディアデータの意味的構造を記述するためのマルチメディアアノテーションがある。

具体的な応用例としては、Web ページの要約や翻訳、レイアウト変換、「音声→テキスト」や「テキスト→音声」といった変換、映像から要約した映像への変換などが挙げられる。さらに、複数のコンテンツからユーザの好みに合った新規のコンテンツを生成するといった応用も視野に入っている。実際にアノテーションを付与する手法は、コンテンツの種類によって異なる。

#### 3.7.3.1 アノテーションエディタ

コンテンツに効率良くアノテーション情報を付加するために、報告者らはアノテーションエディタと呼ばれるオーサリングツールを開発している。これをアノテーションデータの自動生成や編集に用いる。アノテーションエディタは Java アプリケーションとしてユーザ側のクライアント上で利用できる。生成したアノテーションデータは、アノテーションサーバへ送信され、分類/格納される。

アノテーションデータの記述には、XML を用いる。アノテーションを作成した Web ページが更新された場合、更新前のアノテーションデータの再利用をはかる<sup>1</sup>。更新前のアノテーションデータを参照しながら、更新後のアノテーションデータの再構成を試みるような仕組みを取り入れる予定である。

このエディタを使って、ユーザは言語構造（構文や意味に関する構造）に関するアノテーションをテキスト文に付加したり、コンテンツ内の画像や音声といった要素にコメントをつけたりすることができる。言語構造に関するアノテーションは自動生成できる。ただし、その構造にあいまいさが含まれる場合は、ユーザがアノテーションエディタを操作し

---

<sup>1</sup> ユーザがトランスコーディングを要求した Web コンテンツが更新されたかどうかを知るために、DOM (Document Object Model) ハッシュと呼ばれる手法を利用する[6]。DOM ハッシュでは、Web コンテンツの内部構造に関してハッシュ値の演算を行う。どの HTML エレメントが更新されたかを知ることができる。

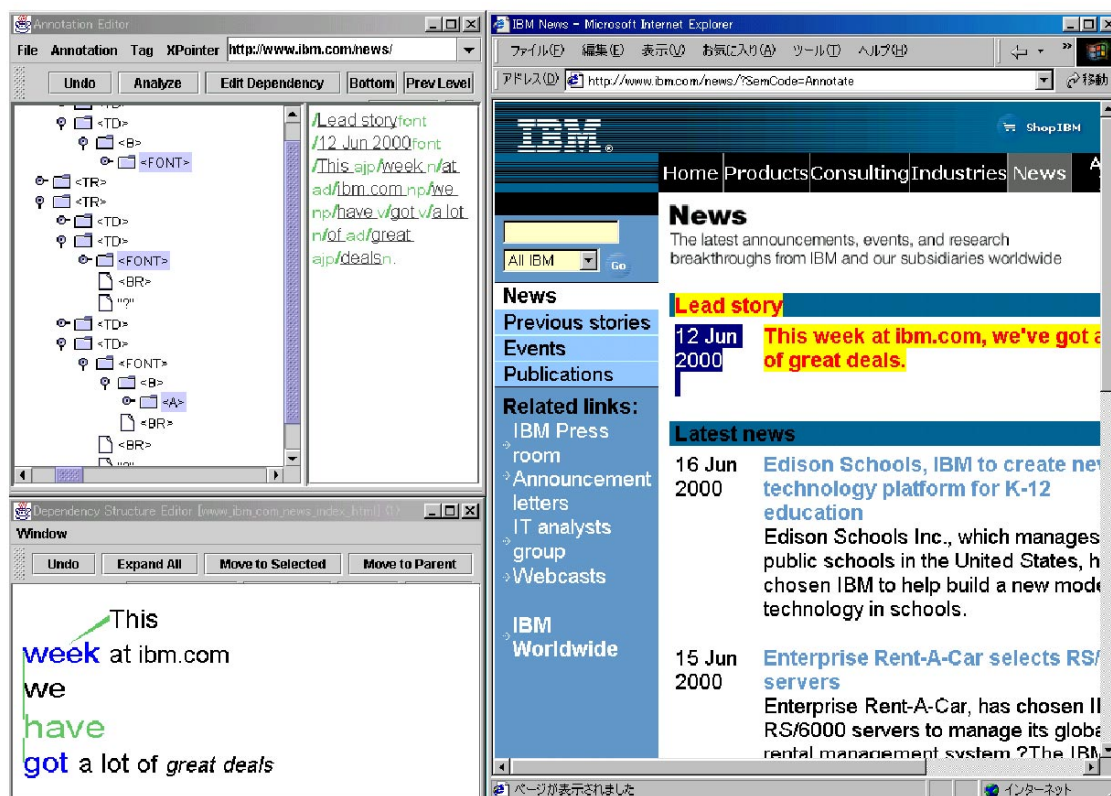


図 3.7-2 アノテーションエディタの画面例

て修正する。言語構造の表示は、グラフィカルに表示してその構造が把握しやすいように工夫した。

図 3.7-2 は、アノテーションエディタの画面例である。これは、テキストに含まれる言語構造を自動解析し、その結果を修正するツールである。言語構造を視覚的にわかりやすく表示し、簡単なマウス操作で修正できる。その他の機能として、HTML 文書の任意のタグ要素にコメント文をつけ加える機能がある。

さらに、報告者らはコンテンツの原著者がアノテーションを制御する方法を開発している。原著者が自分のコンテンツに対する一切のアノテーションを許可しない場合、コンテンツ内にアノテーションを許可しないという記述を入れることによって、アノテーションサーバがそれを認識し、アノテーションエディタからの要求がきても拒否するような仕組みになっている。

### 3.7.3.2 言語的アノテーションとテキストトランスコーディング

テキスト文書のトランスコーディングとは、言語的アノテーションを用いたテキスト文書の加工を指す。言語的アノテーションは、修飾-非修飾関係といった言語構造や固有名詞、動詞などの品詞や語義などをコンテンツに関連づける。具体的には、XML[14]形式のタグ

と属性を使って言語構造や語義をもとのコンテンツに付与する。

言語的アノテーションは、コンテンツに含まれるテキスト文の言語構造に関するアノテーションである。語間の係り受けや代名詞の指示対象、多義語の意味といった詳細な情報を含む。この言語的アノテーションは、ドキュメントの内容理解に大きく貢献し、テキスト文のトランスコーディング以外にも、例えば、内容検索や知識発見などにも利用できる。

言語的アノテーションの記述形式として、GDA (Global Document Annotation) と呼ばれるものがある[1]。これは、XML 形式のタグを利用して、文書の統語・意味・談話構造を明示化するものである。

GDA タグつきドキュメントは、例えば以下のようなものである。

```
<su><adp opr="loc"><adp opr="psr">人間の</adp><np sense="0f2e4c">細胞</np>には、
</adp><np syn="p"><np><vp><adp><adp><np sense="0f74e9">自動車</np>でいえば
</adp><adp opr="iob">アクセルに</adp> 当たり、</adp><adp opr="obj"><np
id="a1"sense="3be2c7">がん</np>を</adp><adp opr="gra">どンドン</adp>増殖する
</vp><n> 「<namep id="a2"><np eq="a1" sense="3be2c7"> がん </np><n
id="a3"sense="3bf4d0"> 遺 伝 子 </n></namep>」 </n></np> と、<np><adp><np
sense="107ab3"> ブレーキ</np> 役の</adp><n> 「<namep id="a4"><np eq="a1"
opr="obj"sense="3be2c7">がん</np><n sense="10d244 3cf57c">抑制</n><n eq="a3"
sense="3bf4d0"> 遺 伝 子 </n></namep>」 </n></np></np> がある。</su><su><adp
opr="cnd"><adp opr="sbj"><np><adp opr="psr"><np eq="a2 a4" sense="0face2">双方
</np>の</adp>バランス</np>が</adp>取れていれば</adp>問題はない。</su>
```

これらは統語構造を表しており、各エレメント（タグで囲まれた部分）は統語的構成素である。ここで、<su>は1文の範囲を表し、<n>、<np>、<vp>、<adp>、<namep>は、それぞれ名詞、名詞句、動詞句、形容詞句/形容動詞句（前置詞句、後置詞句を含む）、固有名詞句を表す。syn="p"は等位構造（例えば、例文中の「～がん遺伝子と～がん抑制遺伝子」）を表す。等位構造の定義は、係り受け関係を共有するということである。特に何も指定がない場合は、例えば、<np><adp opr="x">A</adp><n>B</n></np>はAがBに依存関係があることを表す。また、opr="x"は<adp>エレメントの関係属性を表している。また、sense=""は語義属性を表している（属性値としては、例えばEDR 単語辞書[5]の概念識別子が利用できる。また、1語が複数の語義を持つ場合は、属性値が複数になる）。

この形式を用いて、テキスト文書の要約や翻訳を実現する。例えば、このタグを使った要約のアルゴリズムは以下のようなになる[9]。

(1) 要素とその参照要素の間で重要度が等しくなり、それ以外では重要度が減衰するように重要度の計算（活性拡散）を実行する。

- (2) 重要度の拡散演算が終了した時点で、平均重要度の大きい順に文を選択する。
- (3) 選択された文で削除すると意味が通らない必須要素を抽出する。
- (4) 文の必須要素をつなげて文の骨格を生成し、要約に加える。代名詞などの参照表現の先行詞が要約に含まれない場合は参照表現を先行詞で置き換える。
- (5) 要約が指定された分量に達したときは終了する。まだ余裕がある場合は、次に重要度の高い文と省略した要素の重要度を比較して、高い方を要約に加える。

必須要素になり得る要素は、大きく分けて3つある。1つは要素の主辞である。主辞とは、ほかの要素に依存しない中心となる要素である。もう1つは内容、原因、条件、主題などの関係属性を持つ要素で、例えば、主語や目的語などが当てはまる。ほかに、等位構造が必須要素の場合は、それに直接含まれる要素も必須要素になる。等位構造とは、例えば2つの要素Aと要素Bが、ANDあるいはORの関係で結ばれている構造を指す。等位構造の要素のうちいずれを削除しても文の意味が変わってしまう。

報告者らの手法では、特定の個人の趣味や嗜好に、より柔軟に適合するように一連の処理を実行することが可能である。例えば、要約を開始する時点でユーザが任意のキーワードを入力して嗜好や興味を反映できるようになっている。ユーザが入力したキーワードと関連する単語を重要語として処理する。重要語を含む要素は重要度の初期値をかなり大きくとり、活性拡散を行う。さらに、ユーザの趣味や嗜好の学習機能も含まれている。キーワード設定の履歴に応じて、コンテンツに含まれる要素の重要度を定めることができる。これによって、要約システムは特にユーザからの入力がなくとも、そのユーザに特化した要約を生成することができる。このほか、テキスト文書のトランスコーディングの例としては翻訳が挙げられる。現在、英語から日本語、および英語からヨーロッパ言語（ドイツ語、フランス語、スペイン語、イタリア語）の自動翻訳をトランスコーディングとして実現している。英日翻訳に関しては、日本アイ・ビー・エムが開発した翻訳エンジンを使用している。PC向け翻訳ソフトウェア「インターネット翻訳の王様」の翻訳エンジンを、アノテーションを考慮して翻訳するように拡張したものを用いている[13]。

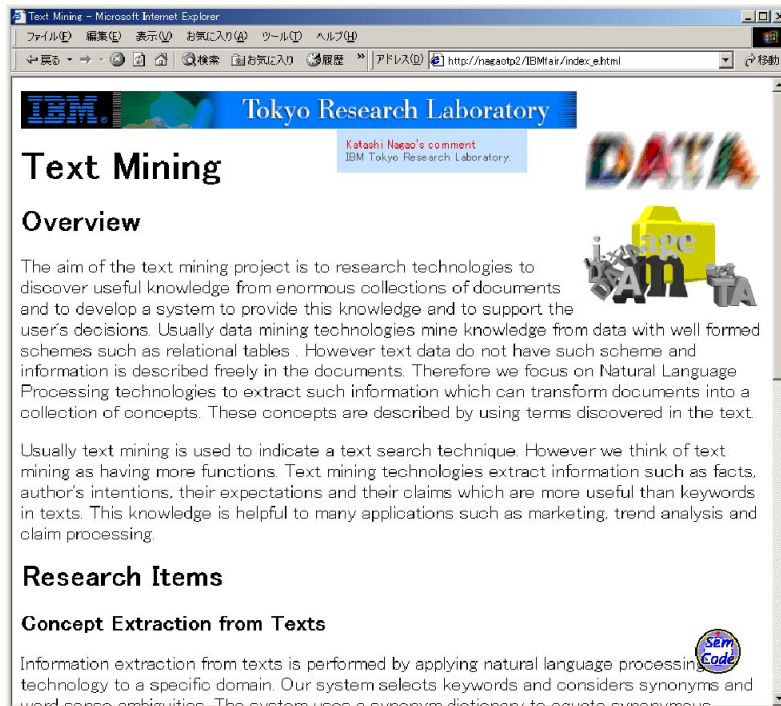


図 3.7-3 オリジナルドキュメント



図 3.7-4 要約・翻訳されたドキュメント

図 3.7-3 はオリジナルのコンテンツであり、要約・翻訳した結果が図 3.7-4 である。

翻訳エンジンによる誤訳の多くの部分は、語間の係り受け解析の失敗や、多義語の訳語選択の失敗による。係り受けや語義を明示化したアノテーションをコンテンツに付加することによって、翻訳精度の改善が見込める。将来的にはコンテンツの作者は、トランスコーディングによって誤解が生じるのを防ぐために、積極的にアノテーションをつけるようになると思われる。

近い将来、Web コンテンツを母国語でしか表現していない人々が、トランスコーディングによってさまざまな言語圏の人々に情報発信できるようになるだろう。

また、読者の理解度に合わせて専門的な表現をより一般的な表現に書き換えて、わかりやすい文章にする、パラフレーズトランスコーディングについても研究を進めている[2]。

#### 3.7.3.3 コメントアノテーションと音声トランスコーディング

セマンティックトランスコーディングシステムは、テキスト文などのコンテンツを音声合成によって、音声に変換することもできる。テキスト文だけでなく、画像とテキスト文が混在するコンテンツの場合、画像の説明にあたるコメントアノテーションを用いることによって、非テキスト要素も含めて音声化することができる。さらに固有名詞など、正しい読み方が音声合成用の辞書にない場合も言語的アノテーションによって、読み方を指定することができる。

コメントアノテーションは、主に非テキスト要素に対する任意のコメントを含むアノテーションである。コメントはテキスト文だけでなく、画像やハイパーリンクなども含むことができる。例えば、コメントアノテーションを含んだ画像をマウスポインタで指すと、その画像の説明ウィンドウがポップアップして表示される。あるいは画像とテキストを含むコンテンツ全体を合成音声で読み上げた場合、画像部分に関してはコメントアノテーションを参考にして音声化する。コメントアノテーションは以前から研究が行われている。例えば、コメントを管理するサーバと、コンテンツにコメントを加えて加工するプロキシを別個に用意するというものである[10][11]。これはコンテンツを共有するグループが、コンテンツに関する補足情報を効果的に共有できるように配慮したものである。基本的には、セマンティックトランスコーディングの枠組みもこれと同様である。ただし、コメントを付与する単位がコンテンツ全体ではなく、任意の HTML の要素に対して行える。コメントはそれを読む人間のためであると同時に、そのコンテンツを機械が理解して適切にトランスコードするための手段としてとらえている。

実際にコンテンツを音声に変換してユーザに配信するには 2 種類の方法があるだろう。1 つはトランスコーディングを担当する外部サーバが音声データを作成してクライアントに配信する方法である。クライアントが音声合成機能を持たない場合に有効である。例えば、携帯電話から Web ページにアクセスするときを使う。もう 1 つは、音声合成システムを備えたクライアントに加工したテキストデータを配信する方法である。この場合、音

声合成に適した形式にコンテンツをトランスコードすることになるだろう。MP3 のデコーダが内蔵された携帯電話が発売されるようになった。今後、配信された音声データを携帯電話に一時的に保存して、ユーザの都合の良いときに聞くといった使い方が盛んになると思われる。

#### 3.7.3.4 イメージトランスコーディング

画像のトランスコーディングは、ユーザが使用する機器の表示能力などに合わせてコンテンツに含まれる画像のサイズや解像度を変換する処理である。ただし、変換された画像は必ずもとの画像へのリンクを含むようにしてあるため、オリジナルのサイズや解像度で見たいときは、単にその画像をクリックすればもとの画像が現れる。

画像のトランスコーディングとテキストのトランスコーディングを併用することにより、ユーザの好みに応じて、画像と文書の表示バランスを変えることができる。表示する要約の分量や、画像と文書の表示バランスは、ページに埋め込まれた設定ウィンドウで調整する。

#### 3.7.3.5 マルチメディアアノテーションとビデオトランスコーディング

マルチメディアコンテンツのメタデータあるいはアノテーションは、これまで主に著作権管理、検索、選択などの応用を目指して設計されてきているが、さらに多くの応用を可能にする枠組みとしてセマンティックトランスコーディングが利用できる。

映像コンテンツをトランスコーディングする場合、まず映像コンテンツに含まれる音声のトランスクリプト(書き起こしたテキスト文)を用意する。このトランスクリプトに、意味構造や、シーンの変わり目のタイムコード、シーンごとのキーフレームの位置、映像の各シーンに登場するオブジェクトの名前とその出現位置(時間と座標)などをアノテーションとして付加する。

セマンティックトランスコーディングにおけるマルチメディアアノテーションでは、トランスクリプトを自動的に生成して、半自動的にアノテーションを作成できる。映像のシーンの変わり目も自動認識し、シーンに関するタグづけを支援する。

このシステムは、現在のところ、映像コンテンツの要約の生成、映像コンテンツからテキストと画像からなるコンテンツへの再構成や、ビデオ音声の翻訳などが実現できる。

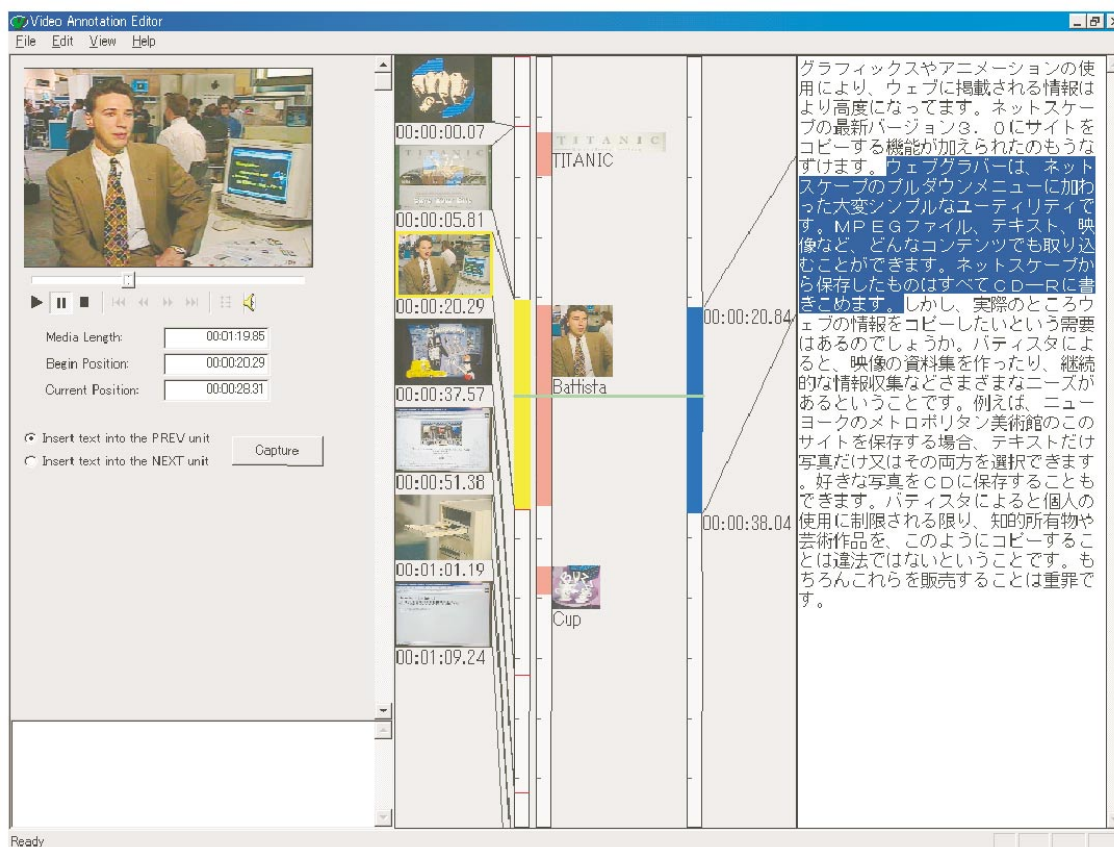


図 3.7-5 ビデオアノテーションエディタ

図 3.7-5 はビデオアノテーションエディタの画面例を示している。このエディタは、ビデオのシーンへの分割と音声部分のテキスト化を行う。自動処理の結果はインタラクティブに修正できる。

ビデオアノテーションエディタは、まず、ビデオデータを音声データと動画像データとに分離する。音声データについては、2 か国語放送のようにチャンネルごとに異なる言語音声収録される場合があるため、さらに左右のチャンネルごとに音声を分離する。両チャンネルの音声信号の差分の平均が閾値以上である場合は異なる音声とみなし、多言語音声識別・認識処理にかける。こうして得られるトランスクリプトは、タイムコードと認識された単語列および言語情報を含んでおり、次に示すような XML 形式のデータとして表現される。

```
<text lang="ja">
<w in="1.264000" out="1.663000">残す</w>
<w in="1.663000" out="2.072000">時間も</w>
<w in="2.072000" out="2.611000">少なくない</w>
```



```
<w in="2.611000" out="3.180000">ましたが</w>
<w in="3.180000" out="3.778000">最後に</w>
<w in="3.778000" out="4.856000">次近々発売</w>
<w in="4.856000" out="5.215000">される</w>
<w in="5.215000" out="5.934000">マイクロソフト</w>
<w in="5.934000" out="6.153000">の</w>
<w in="6.153000" out="6.462000">オート</w>
<w in="6.462000" out="6.802000">マップ</w>
<w in="6.802000" out="7.191000">トリップ</w>
<w in="7.191000" out="8.039000">パンダを紹介</w>
<w in="8.039000" out="8.538000">しましょう</w>
...
</text>
```

一方、動画像データはシーン検出処理およびオブジェクトトラッキングにより、タイムコードとシーンの代表イメージ（キーフレーム）を含むシーン情報データと、オブジェクトの名称と説明とイメージ、関連する URL へのリンク、出現タイムコード、画像上の位置の軌跡情報からなるオブジェクト情報データを、次に示すような XML 形式のデータで表現する。

```
<scene>
<v in="0.066733" out="11.945279" file="s0.jpg"/>
<v in="11.945279" out="14.447781" file="s1.jpg"/>
<v in="14.447781" out="18.685352" file="s2.jpg"/>
...
</scene>

<object>
<vobj begin="1.668335" end="4.671338" name="Davids"
desc="anchor" img="o0000.jpg" link="http://...">
<area time="1.668335" top="82" left="34"
width="156" height="145"/>
<area ... />
</vobj>
...
</object>
```

2 か国語放送のようなビデオデータの場合、シーン情報やオブジェクト情報のタイムコードとそれに対応するトランスクリプトのタイムコードとがチャンネル（言語）ごとに異なることが多いため、これらのアノテーションデータはチャンネルごとに作成し、各情報をまとめて記述する次のようなビデオアノテーションデータを作成する。

```
<vax file="D:¥demo.mpg">  
<text channel="0" src="D:¥demo¥0.vt"/>  
<text channel="1" src="D:¥demo¥1.vt"/>  
<scene channel="0" src="D:¥demo¥0.vs"/>  
<scene channel="1" src="D:¥demo¥1.vs"/>  
<object channel="0" src="D:¥demo¥0.vo"/>  
<object channel="1" src="D:¥demo¥1.vo"/>  
</vax>
```

ビデオトランスコーディングには、ビデオの文書化（ビデオからテキストとイメージの集合への変換）、ビデオの要約、ビデオの翻訳などが含まれる。それらについて簡単に説明する。

ビデオからテキストとイメージへの変換は、ビデオトランスコーディングの基本となる変換である。例えば、もし、クライアントのデバイスがビデオを再生することができない場合、ユーザはビデオのコンテンツにまったくアクセスできなくなってしまう。その場合、ビデオトランスコーダーはそれぞれのシーンを代表するイメージとそれぞれのシーンの内容を表すテキストを含めたドキュメントを作成してユーザに提示することができる。このドキュメントは、PC のブラウザに表示されるときは、ビデオプレイヤーを内部に埋め込んだマルチメディアプレゼンテーションとして機能する。つまり、イメージをクリックして、それに関連するシーンをプレイヤー上で再生したり、再生中のシーンのトランスクリプトを、背景色を変えて強調したりすることができる。また、シーン内の人物や対象物の領域を選んで、それに関連づけられたコメントなどの情報をポップアップウィンドウに表示することもできる。さらに、後述するように、生成されたドキュメントを、テキストトランスコーダを用いて要約あるいは翻訳することもできる。



図 3.7-6 ビデオドキュメント

図 3.7-6 は生成されたビデオドキュメントの画面例である。左上がビデオプレイヤーの画面でその下に字幕ウィンドウがある。中央の垂直のバーはタイムバーで、右のウィンドウには、シーンのキーフレームと、そのシーンのトランスクリプトが表示されている。

ビデオの要約は、まずビデオのトランスクリプトを要約して、その要約に対応するビデオシーンを抽出することによって行われる。これは、トランスクリプトが対応する音声の出現するタイムコードを含んでいるため、そのタイムコードを含むシーンを選択することで要約できる。



図 3.7-7 要約されたビデオドキュメント

図 3.7-7 は要約されたビデオドキュメントの画面例である。中央のタイムバーの青色の部分に要約が含まれる。また、要約に含まれないシーンのキーフレームは小さく表示される。もちろん、要約だけ再生させることもできるが、その他のシーンもキーフレームをクリックすることによって再生できる。また、テキストの要約と同様に、要約のサイズを変えたり、キーワードを入力したりするなど、インタラクティブに要約をカスタマイズすることができる。

ビデオの翻訳には、2 種類あり、トランスクリプトを翻訳してテキストとして表示する場合と、翻訳結果を音声合成によって音声にし、ビデオの再生と音声の再生を同期させることによって、他の言語のビデオを作成する場合がある。後者の部分は、まだ実現されていないが、近い将来にこの機能も統合される予定である。

報告者は映像コンテンツが今後重要な情報ソースになることを確信している。そのため、要約やフィルタリングに限定されない、コンテンツの再利用を可能にするさまざまな枠組みをできるだけ早めに用意しておきたいと考えている。ここでのアノテーションを利用した手法は、将来の枠組みに対しても容易に付加情報を変換して対応できるようにした。例えば、MPEG-7[7]のようなアノテーションの標準的なフレームワークが確立した場合にも

容易に移行できるだろう。

### 3.7.4 インタフェースロボット

ロボット技術の目覚ましい進歩に伴い今世紀には人間とコンピュータの対話のスタイルの劇的な変化が予想される。そこで、報告者らは人間とコンピュータの新しい対話スタイルを想定して、統合された普遍的なプラットフォームを適用し、それを通して、この分野で重要な役割を果たすことを目標とする。ここでは従来のようなコマンド入力中心の一方通行型の操作ではなく音声出入力システムを用いた双方向型の対話のメカニズムが中心になる。また、学習・適応・記憶をつかさどり、対話システムの主要なプログラムとなる感情ユニットを開発して、人間的なモダリティを備えた対話インタフェースの実現を目指している。こうしたシステムはさまざまな情報端末に対応する強力なツールとなり、それにより自然な対話の実現すれば多くのユーザが利用することになるだろう。

自然言語インタフェースのような従来のシステムでは細かい制約が多く自然な会話は成立しにくかった。ユーザにはコメントの正確さが求められるし、コンピュータは予期しない文章には応答できず文脈も強く制約される。こうした問題のため、このようなシステムにおける会話は目的を1つに絞った単純なものになってしまう。こうした会話は人間の日常的な会話とはかけ離れた不自然なものである。こうした理由から、その用途は電話案内などの特殊なものに限られており、一般的普及には至らなかった。

日常会話を観察すれば、その中でさまざまな内容が展開され複数の目的に向かった発話が同時になされていることがわかる。ある発話とそれに続く発話のつながりは明らかに希薄である。コンピュータはそのような日常会話の特性にあわせてシステムを実用に耐えるべく調整しなければならない。人間の会話というものは、本来、思いもよらない方向に発展する性質のものである。報告者らが提案する会話システムはこのような会話を扱うことによって人間の会話により近い自然な雰囲気を作ることが目標である。ここで提唱されるシステムは即応的発話ユニットとプランに基づく発話生成ユニットそしてそれらを統合する感情ユニットから成る。

即応的発話ユニットは入力に対して局所的な観点から出力を返す。一方プランに基づく発話生成ユニットは会話の全体的なフェーズを考慮しながら会話をコントロールする。さらに感情ユニットは前者が後者をコントロールし、後者が前者を活性化するような設計を備えている。感情ユニットは両者から情報を受け取り、複雑な感情状態を作り出す。その感情状態がまた両者に再供給される、という仕組みである。

このような柔軟な会話ナビゲーションユニットの導入により、文脈が刻々と変化し何の目的も持たないようでありながら大局的には1つの筋道に従って展開する人間の会話により近いインタラクションが現実的なものになるだろう。

テキストを主として処理を行う従来のテクノロジーに対して、報告者らは、文脈に基づく手法を適用する。そこではフレーズや断片的発話に関する口語的処理が中心となる。それによって文法からの逸脱、いい間違い、言い換え、言いよどみ、音声認識上の問題といった自発的対話への対処も可能になるだろう。

#### 3.7.5 会話ロボット Pong と QB

このようなアーキテクチャに基づいて、新たなヒューマンインタフェースとして会話ロボットに関する研究を進めている。報告者らの取り組むロボットは、手足の動きなど外観の機能や動きを追求するのではない。「人間にとって役に立つ、便利な助手のような存在」として利用できるようにするための仕組み、つまり知能の部分にフォーカスを当てている。人間とのコミュニケーションの手段は非常に身近な「会話」である。しかしながら、人間が機械に対して話しかける行為に抵抗を感じるという問題をまずはクリアしなければ役には立たない。人間は無意識のレベルで、存在感があり、動作する対象が自分に注意を向けていることを感じ取ると、それは「自分とコミュニケーションする対象なのだ」と直感的に認識していると思われる。つまり、直感的で臨場感のある「会話」によって、人間と機械がスムーズなコミュニケーションを行うために、「ロボット」という入れ物を利用するわけである。その1つが以下に紹介するロボット Pong である。

##### 3.7.5.1 Pong

Pong は、話しかけられると、ユーザの意図を理解し、必要な情報を集めて自動的に処理し、適切な応答の生成を試みる。さらにはユーザと行ってきた会話の履歴を記憶する能力を持たせることにより、ユーザの嗜好にあった行動をロボットが自発的に行ってくれるようになるだろう。また、ユーザに特化し同調するような仕組みとして感情を理解できるシステムもあらかじめ組み込んでおけば、ユーザがうれしいとき、ロボットもうれしいし、楽しいとロボットも喜んでくれるだろう。ロボットが知性と感情を持ちあわせるようになったとき、便利さだけでなく、人間を癒してくれる存在となるだろう。

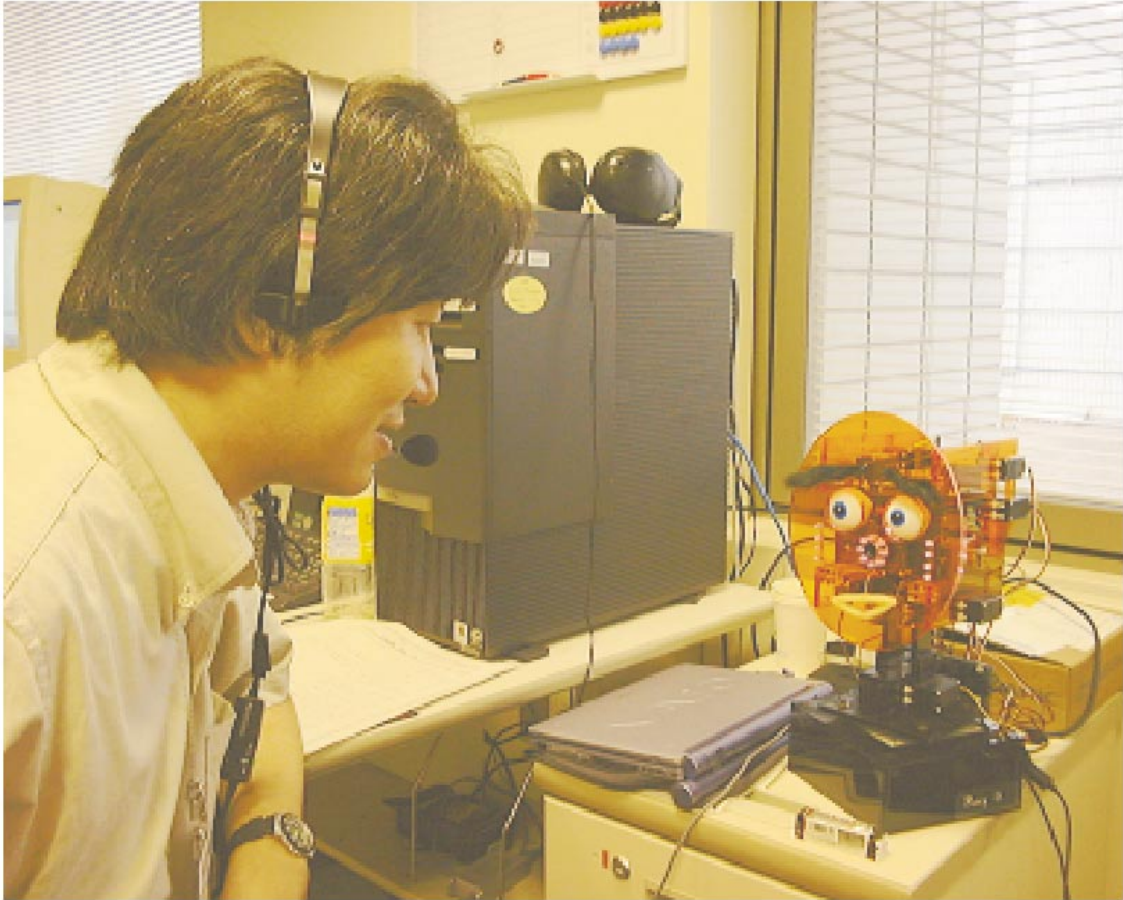


図 3.7-8 Pong との対話

図 3.7-8 は、Pong と報告者との対話風景である。

最近、ペット型ロボットなどにより人間とある種のコミュニケーションを行うロボットに関心が高まっている。しかし、エンターテインメント性だけを追求していると、いつか飽きられてしまうだろう。しかし、人間との会話や知識を処理するインテリジェンスを持たせたとき、人間にとって本当に信頼のおけるパートナーになり得るだろう。

以下は Pong との会話例である。

Pong 「こんにちは。ボクは Pong って言うんだ。アメリカの IBM 研究所で生まれたロボットなんだ。ボクは人間の言葉が少しだけわかるんだ。ボクに何か聞いてよ」

人間 「君は何ができるの？」

Pong 「いろいろできるよ。例えば計算とか」

人間 「じゃ、問題を出すよ」

Pong 「いいよ」

人間 「 $100 \times 200$  は？」

Pong 「100×200 の答えは 2 万だよ」

人間 「じゃ、100÷7 は？」

Pong 「100÷7 の答えは割り切れないよ。14.28571428……」

人間 「どうもありがとう」

Pong 「どういたしまして」

「ピンポン (卓球)」の「ポン」から名づけられたこのロボットは、もとは IBM アルマデン研究所で開発されたものである。人間の感情を理解できるコンピュータの開発を目指す「ブルーアイズ」プロジェクトの産物である[4]。赤外線 LED を頬に埋め込み、鼻の部分に仕込まれたカメラを使って人の顔などの画像を入力する。人の方向を向いて笑ったり、人が視界からいなくなると悲しそうな表情をしたりするなど、人間の行為に対して反射的な応答を示すことができる。ここで重要なのは、顔表情が人間とのコミュニケーションを促進する働きがあるということである。これは、コンピュータアニメーションでも同様なことが言え、インタフェースエージェントにキャラクタの外観を与え、表情やジェスチャーなどのノンバーバルモダリティを持たせる研究がいくつか行われた。現在では、インタフェースエージェントという名前を聞いたときに、コンピュータスクリーン上の擬人的なキャラクタを想像する人は少なくないであろう。しかし、スクリーン上のバーチャルなキャラクタに絶対的に不足しているのは、「彼 (彼女) が今そこにおいて、自分に関心を向けている」という存在感あるいはアウェアネスであろう。物理的な存在感ほど人間の直感に訴えるものはない。バーチャルリアリティが目指しているのも、まさにその部分をソフトウェアによって制御可能にすることだと思われる。

インタフェースが物理的な存在感を持ち、人間と会話ができ、働きかけができるものであるなら、親密度はぐっと増し、コンピュータ (の埋め込まれた環境) はとつきにくく、使いづらいというようなイメージを払拭できるかも知れない。

そのような発想から生まれたのが、Pong のような会話ロボットである。ロボットに話しかけるだけで、テレビを見なくても、新聞を読まなくても、その日のニュースや情報を得ることができるようになる。例えば、以下のような具合である。

人間 「Pong、インターネットってわかる？」

Pong 「もちろんわかるよ。インターネットを使うといろんな情報が手に入るよ」

人間 「じゃあ、日経新聞のページが見たい!」

Pong 「日本経済新聞のホームページだね。スクリーンに出すよ」

人間 「読んでくれる？」

Pong 「東京証券取引所の先物取引で……」

人間 「英語のページも見たいな。ブルーアイズのホームページをお願い」

Pong 「ブルーアイズのページを出すよ」



人間「英語わかる?」

Pong「もちろんわかるよ。でも、まだあまり上手じゃないけど」

人間「説明してくれるかな」

Pong「ページに書いてあることを説明するよ。ブルーアイズは……」

今後、すべての家電にコンピュータが搭載されネットワーク化される、いわゆる「情報家電」の時代がやってくると言われている。家庭内がネットワークでつながると、それらを使いこなすにはどうしたらいいか、また新たな問題が発生してしまうだろう。これまでならテレビのスイッチを入れればテレビがついたように、操作する対象と操作内容が1対1の関係であったためにわかりやすかったものが、さまざまな機械が依存関係を持ち、何を操作すると何に影響するのかがわかりにくくなってしまいうだろう。このような依存関係をすべて人間に把握させるのは困難なので、インタフェースとしてのロボットにその役割を担ってもらおうということである。ロボットに情報家電のすべての機能および依存関係を把握させておき、ロボットに自分の要求を言葉で伝達すると、ロボットは人間の意図と状況を認識し、家電を操作し、状況を報告する、という具合になるだろう。そのような初めて、ロボットは人間にとってなくてはならない存在になると思う。

Pong には、音声認識、音声合成、対話処理技術のほかに前述のセマンティックトランスコーディング技術が統合されている。このように、会話ロボットはさまざまな技術を統合するテストベッドにもなる。

最後にインタフェースエージェントとロボットの今後の方向性について述べる。物理的な存在であるロボットの弱点はそのままの形では情報世界を移動できないことである。しかし、ロボットの記憶や知識は人間にとって重要なものであり、例えば外出先であっても、それを利用したい状況は多いと思われる。人間がロボットを引き連れて移動することも考えられるが、それよりも、ロボットの記憶や知識を受け継いだ（ソフトウェア）インタフェースエージェントを携帯型システムに常駐させ、移動中であってもインタラクション可能にするやり方が有効であろう。

この場合、物理的な存在感は感じられないが、以前のインタラクションの記憶が継承されていれば人間にとってロボットとエージェントは一貫したものと感じられ、電話で遠隔地の人間と話すように、エージェントと会話して、要求を伝達することができるだろう。

このように、インタフェースエージェントと会話ロボットはその多くの部分を共有し、記憶の一貫性を保ち、状況に応じてインタラクションのスタイルを変えられるようになると思われる。

また、ロボットとエージェントは、コミュニティを媒介し、人同士のコミュニケーションを促進するシステムとしても機能する。例えば、会話ロボットは遠隔地にいるもの同士が非同期でコミュニケーションする場合に、物理的な存在感を備えた、擬似的な対面性を実現することができる。また、長い経験を通じて個人情報を獲得したロボットやエージェ



図 3.7-9 QB との対話

ントは、その人の外在化された記憶あるいはパーソナリティを持つシステムとして機能できるようになるだろう。これは、ロボットやエージェントが文字通り人間の代理人になるということであるが、もしロボットやエージェントの信用や責任という問題がうまく技術的あるいは社会的に解決できれば、人間は自分のコピーあるいは分身を作って同時に複数の作業ができるようになるだろう。

#### 3.7.5.2 QB

QB は、Pong の後継機となる、表情豊かな移動型会話ロボットである。このロボットは、Pong と同様に、人間の言葉を聞いて、さまざまな質問に答えてくれる。例えば、明日の天気や、今日のニュースや、スポーツの結果など、Web コンテンツから必要な情報を探し出して音声で伝達することができる。また、QB には人のような顔があり、その表情は会話に臨場感や親近感を与えてくれる。そして、頭や手を触ると、QB はときに喜び、ときに嫌がったりもする。これは、感情表現によって、ユーザの注意をひきつけるためである。

図 3.7-9 は、QB と報告者との対話風景である。QB は、Pong と違って自律的に移動できるので、人間は QB に話しかけるのにワイヤレスマイクの機能を持つ PDA (携帯情報端

末) を用いる。ユーザはこの PDA を使って、認証に必要なユーザ ID やユーザの現在位置を伝達することもできる。

また、QB は自らの判断で場所を移動し、離れている人とも会話しようと試みる。例えば、遠くにいる QB を声で呼ぶと、近くまで寄ってきて話を始める。つまり、QB には自らの位置と人間の位置を認識して、自律的に移動する能力がある。

QB には自分の家があり、バッテリーのチャージをするスペースが用意されている。そのため、QB は自分のエネルギーが少なくなると、家に帰っていく。QB の家は、インターネットから得られた情報を個人ごとの要求に合わせて加工する機能が組み込まれていて、QB とは無線で通信し情報交換することができる。

さらに、QB は、複数の人の声を聞いて、それぞれの人に合った答え方をする。生きた掲示板のように、多くの人の意見を収集して、個別に取り次いだりするメッセージャーとしての働きもすることができる。将来は、QB のようなロボットが多くの人々の間に浸透していき、コミュニケーションを活発化させる役割を果たすようになると思われる。そして、人間の質問に気軽に答えてくれるような知的なパートナーになってくれることだろう。

QB の主な目的は次の通りである。

- ・ 次世代のユーザインタフェースとして、音声対話機能、およびサーバとの通信機能を持ったインタラクティブなロボットを提案し、試作・実験を行う。
- ・ ロボットは、物理的な存在感と、ある種の感情を含めた多様な擬人的表現によって、ユーザの興味を刺激する。
- ・ 多様な文脈に適応したロボットの応答によって、ユーザの継続的な使用を促進する。
- ・ ロボットはサーバと連動して、インターネットを通じたさまざまな情報サービスを、ユーザによる複雑な手順を代行して提供する。
- ・ ロボットはユーザを記憶し、その特性に合わせて応答を変え、コンテンツをパーソナライズする。

### 3.7.6 おわりに

Pong や QB を通じて、会話ロボットのコンセプトはかなり明確になったものと思われる。また、これらのプロトタイプシステムを開発して実験した経験から、新しいユーザインタフェースとしての会話ロボットが、PC に代わって一般家庭に普及する可能性が大いにあることを実感している。

また、ユーザに提示されるコンテンツは、高度に個人化されたものになるだろう。本報告で紹介したセマンティックトランスコーディングが、そのための重要な技術の 1 つになるものと確信している。

#### 謝辞

セマンティックトランスコーディングとインタフェースロボットは報告者と学生研究員（細谷真吾、白井良成、大平茂輝、東中竜一郎、米岡充裕、伊藤大輔、Kevin Squire、梅澤猛、福岡俊樹、片桐由希子、斉藤美紀、小熊崇）との共同研究である。諸氏に感謝する。言語的アノテーションに関しては、GDA プロジェクトと連携して行われている。プロジェクトリーダーの産業技術総合研究所サイバーアシスト研究センターの橋田浩一氏には、常に議論の相手をしていただいている。ここに記して感謝する。

#### 参考文献

- [1] Koiti Hasida: Global Document Annotation, <http://www.i-content.org/GDA/>.
- [2] 東中竜一郎, 長尾確: アノテーションを用いて Web ドキュメントを分かりやすく提示する方法, 第3回インターネットテクノロジーワークショップ (WIT2000) 論文集 (2000).
- [3] Masahiro Hori, et al: Annotation-based Web Content Transcoding, Proceedings of the Ninth International WWW Conference, pp. 197-211 (2000).
- [4] IBM Almaden Research Center: BlueEyes Project Web Page, <http://www.almaden.ibm.com/cs/blueeyes/>.
- [5] Japan Electoronic Dictionary Research Institute: Electoronic Dictionary, [http://www.ijjnet.or.jp/edr/J\\_index.html](http://www.ijjnet.or.jp/edr/J_index.html).
- [6] Hiroshi Maruyama, Kent Tamura, and Naohiko Uramoto: XML and Java: Developing Web applications, Addison-Wesley (1999).
- [7] Moving Picture Experts Group (MPEG): MPEG-7 Context and Objectives, <http://drogo.cselt.stet.it/mpeg/standards/mpeg-7/mpeg-7.htm>.
- [8] Katashi Nagao, Yoshinari Shirai, and Kevin Squire: Semantic Annotation and Transcoding: Making Web Content More Accessible, IEEE MultiMedia, Vol. 8, No. 2, pp. 69-81 (2001).
- [9] Katashi Nagao and Koiti Hasida: Automatic Text Summarization Based on the Global Document Annotation, Proceedings of COLING-ACL '98, Vol. 2, pp. 917-921 (1998).
- [10] Martin Roscheisen, Christian Mogensen, and Terry Winograd: Shared Web Annotations as a Platform for Third-party Value-added Information Providers: Architecture, Protocols, and Usage Examples, Technical Report CSDTR/DLTR, Computer Science Department, Stanford University (1995).
- [11] Matthew A. Schickler, Murray S. Mazer, and Charles Brooks: Pan-browser Support for Anotations and Other Meta-information on the World Wide Web,

Computer Networks and ISDN Systems, Vol. 28 (1996).

- [12] 浦本直彦: Semantic Web - 機械のための Web -, 人工知能学会誌, Vol. 16, No. 3, pp. 412-419 (2001).
- [13] Hideo Watanabe, Katashi Nagao, Michael C. McCord, and Arendse Bernth: Improving Natural Language Processing by Linguistic Document Annotation, Proceedings of COLING 2000 Workshop on Semantic Annotation and Intelligent Content, pp. 20-27 (2000).
- [14] World Wide Web Consortium: Extensible Markup Language (XML), <http://www.w3.org/XML/>.

## 3.8 法的推論システムの研究動向と新しい展開

報告者: 新田克己委員

### 3.8.1 はじめに

「法的推論」とは法律家が法律を解釈し、事件に適応させる推論のことである。法律分野はコンピュータの応用としては最も古いものの1つであり、多くの法律エキスパートシステムが作られてきた。AIや認知科学の観点から法的推論が関心をもたれるのは、(1) 法的推論はさまざまな様相をもち、仮説推論や類推や非単調推論などのいろいろな推論方式の応用例題として適していること、(2) 法律の解釈を行うには、言葉の意味を深く解析する必要があり、高度な意味処理技術の課題が多いことなどの理由である。

しかしながら、近年では基礎研究よりも次第に実システムへの研究に向けて、研究動向が変わってきているように見られる。そこで、本報告では法的推論の主要な研究課題の動向を紹介した後で、法的推論技術の新しい応用の可能性について紹介する。

### 3.8.2 法的推論の概要

法律は少ないルールで多様なトラブルに対処するため、抽象的な述語を使ってかかっている。そのため、法律を実際の事件に適用するには「法律の解釈」をし、法律をその事件に適用できるように具体的なルールにする必要がある。法律の解釈には、法目的や市民感情や慣習などのいろいろな価値観のバランスをとる必要があり、機械的に行うことは容易ではない。例えば、

「公園に車は進入禁止」

というルールに対し、急病人を助けるために救急車が公園に入ることはこのルールの例外とすべきであろう。では

「公園に救急車以外の車は進入禁止」

というルールを作ればよいかといえそうではない。帰路に近道をするために公園に入る救急車については進入を認めるべきではないであろう。すると、

### 「公園に緊急のための救急車以外の車は進入禁止」

というルールになる。しかし、訓練のための公園進入は認めるべきだから、このルールにも例外がある。

このように、実際にルールを適用する場合には、そのルールの例外や例外の例外などのいろいろな場面を想定しなければならず、それらをすべて列挙することはできない。上記の問題を解決する現実的な手段は、もとのルールを

### 「公園に車は原則として進入禁止」

と考えておき、「原則」の範囲に関しては、ルールの本来の目的や、社会感情や、慣習や、類似事件の過去の判断との一貫性などを考慮してケースバイケースで決めることである。これは異なる価値観のバランスをとって「妥当な判断」をするということである。

このように考えると、法的推論は、ルールを作ったり、ルールを適用して結論を導いたりするような「論理操作」だけではなく、「価値判断」も重要であることが分かる。価値判断は機械化しにくい部分であり、これが法的推論システムの構築を難しくしている要因である。

また、法的推論には別の側面がある。法的判断が価値判断に依存するということは、価値観によって異なる結論が導かれるということである。法廷で原告と被告の主張が異なるのは、そのためである。価値観の相違は法廷論争の原因であり、法廷論争を行うためには論争や妥協のための推論も必要である。

## 3.8.3 法的推論研究の動向

法的推論システムを実現するための主要な課題である法的知識の記述、法的議論のモデル、価値判断、システム構築方法論について、内容とその研究動向について簡単に説明する。

### 3.8.3.1 法的知識の記述

法律知識には、法令のみならず、判例、学説、社会感情、慣習、政策ほかのさまざまな知識がある。これらをコンピュータ上で表現し利用するための基礎となるのは以下の3つである。

#### (1) ルールベース

法令をプロダクションルールや論理式で表現し、演繹推論で結論を出す最も基本的な推論方式をルールベース推論という。例えば、

$$(\forall X)(殺人を犯す(X) \rightarrow 殺人罪(X))$$

というルールと

殺人を犯す(太郎)

という事実から

殺人罪(太郎)

という結論を導くことができる。

法令文から完全なルールベースが作成できれば、このような演繹で解を導くことが可能である。しかし、前に述べたように実際には「原則として」のようなあいまいな判断がルールに入ることがあり、非単調推論、帰納推論、ファジイ推論などの高度な推論技術を使ってルールを整形したり、常識的な推論を行ったりする必要がある。例えば、

$$(\forall X)(馬(X) \rightarrow 進入禁止(X))$$

なるルールを牛である  $a$  にも適用したいときは、 $\{牛(a)\}$  が正例となるように馬の概念を抽象化して、

$$(\forall X)(家畜(X) \rightarrow 進入禁止(X))$$

というルールを生成する（家畜は馬と牛の共通上位概念であるとする）。これは拡張解釈を実現していることになる。

#### (2) 事例ベース

判例を何らかの形で特徴づけして事例ベースに蓄積しておき、新たな問題（事件）を解決する際に、現事件と類似する判例を引用して結論を導く推論方式は事例ベース推論の典型例である。例えば、Hypo[2]では、

f1a: 殺意をもって暴行した。

f2: 被害者は気絶した。



f3: 死んだものと誤解して放置した。

f4: 放置が原因で死亡した。

判決: 暴行と死亡の間に因果関係を認める。

という 4 つの特徴 (f1a、f2、f3、f4) と結論をもつ判例があるとき、

f1b: 過失で交通事故をおこした。

f2: 被害者は気絶した。

f3: 死んだものと誤解して放置した。

f4: 放置が原因で死亡した。

の性質をもつ事件は、3 つの特徴 (f2、f3、f4) が一致するため、両者が類似すると判断することができ、「交通事故と被害者との間に因果関係を認める (すなわち過失致死罪)」という結論を導くことができる。ただし f1a と f1b の相違点を重視すれば、両者は類似していないと判断されることになるから、この結論は仮定的なものではない。

このほかに、出来事そのものの類似性に着目するのではなく、それが引き起こした効果によって類似性を判定する方法 GDA[8]などが提案されている。

### (3) 義務論理

法律のさまざまな規範を義務論理の論理式で表す方法である。義務論理は命題論理を拡張して、「p は義務である」「p は権利である」「p は禁止されている」をそれぞれ、オペレータ O、P、F、を用いて、

O(p)

P(p)

F(p)

と表現する方法である。例えば、「喫煙が禁止されているなら、喫煙は権利ではない」ことを以下のように表現する。

F(喫煙) → ¬P(喫煙)

法令文を表現するとき、プロダクションルールや述語論理よりも義務論理の方が適しているため、法令解釈の表現手段として研究されている。しかしながら、義務論理は高階の論理であるため、一般的な証明手段はない。そのため、理論的には研究が多くなされているものの、システムとしては限定的な推論機構の開発にとどまっている。

表 3.8-1 International Conference on Artificial Intelligence and Law での発表件数

|       | ルールベース | 事例ベース | 義務論理 | その他 |
|-------|--------|-------|------|-----|
| 1991年 | 20     | 10    | 4    | 5   |
| 2001年 | 8      | 4     | 3    | 15  |

このように、法律の知識表現と推論方式は3つの主要なアプローチがあるが、今までは主にルールベースと事例ベースを組み合わせたハイブリッドなシステムが開発されてきている。

わが国では明治学院大学法学部教授の吉野一を代表とする文部省重点領域研究「法律エキスパートシステム」(1993-1998)によって、ルールベースを基礎として、商取引の国際条約のルールベース化を行い、さらに、帰納推論、アブダクション(仮説生成)、類推、ファジィ推論などの高度な推論機構の法律への応用が図られた。このプロジェクトの終了後は、これらの成果を法学教育に応用する方向で展開が図られている。

海外においては、隔年で開催される「法律と人工知能の国際会議(International Conference on Artificial Intelligence and Law)」での、この3つの記述法の発表件数はこの10年で表3.8-1のように変化しており、記述法の研究自体は全体的に減少傾向にある。

### 3.8.3.2 法的議論のモデル

近年、研究が根強く行われているのは法的議論のモデルである。法的議論に関しては、議論のプロトコルの研究や非単調推論による形式化の研究が行われている。前者においては、CATO[1]のような教育支援システムや Pleadings Game[5]や Room5[7]のような調停システムが代表的なものである。後者の研究の代表的なものは、ルールの優先関係を用いた非単調推論である。例えば、わいせつ物の展示に関して、以下の  $r_2$  と  $r_3$  は矛盾するルールである。そのため、例えば絵画  $a$  がわいせつ物であったとき、ルール  $r_2$  から「展示の自由( $a$ )」という結論が導かれ、ルール  $r_3$  から「一展示の自由( $a$ )」という結論が導かれる。

$r_1$ : わいせつ物( $X$ )  $\rightarrow$  表現( $X$ )

$r_2$ : 表現( $X$ )  $\Rightarrow$  展示の自由( $X$ )

$r_3$ : わいせつ物( $X$ )  $\Rightarrow$  一展示の自由( $X$ )

この矛盾を解消する1つの方法として、

$r_2 < r_3$ . (ルール  $r_3$  はルール  $r_2$  に優先する)

のように、ルールの優先関係を定義する方法がある。優先するルールを含んだ論証の方が、優先しないルールを含んだ論証よりも強い論証である。

このようなルールの優先関係を含んだ非単調推論の形式化は、サーカムスクリプション（極小限定）、拡張論理プログラミング、デフォルト理論のいずれかの拡張によって実現されてきた。優先ルール付きの拡張論理プログラミングの意味論については、近年、[3][6][9][10]ほか、活発に提案がなされている。いろいろな意味論を比較すると、どの意味論にも一長一短があり、法律に適した意味論のための評価基準の検討が必要であると思われる。

拡張論理プログラミングのもとでは、結論に対する反駁（rebutting）と、前提に対する反駁（overcutting）を以下のように形式化する。例えば、

$$\begin{array}{l} \{ \text{鳥(ジョナサン)}. \\ \text{飛ぶ(ジョナサン)} \leftarrow \text{鳥(ジョナサン), not 負傷(ジョナサン)}. \} \end{array}$$

という 2 つの具体化ルールは、「ジョナサンは鳥である。ジョナサンは鳥であって、負傷していることは証明できないから、ジョナサンは飛ぶことができる」という論証を表すとする。not は論理的な否定ではなく、証明失敗としての否定（NAF）である。

これに対する反論として結論に対する反駁は、

$$\begin{array}{l} \{ \text{ペンギン(ジョナサン)}. \\ \text{一飛ぶ(ジョナサン)} \leftarrow \text{ペンギン(ジョナサン)}. \} \end{array}$$

のように反対の結論を持つ論証を相手に提示することであり、前提に対する反駁は、

$$\begin{array}{l} \{ \text{鷹に襲われた(ジョナサン)}. \\ \text{負傷(ジョナサン)} \leftarrow \text{鷹に襲われた(ジョナサン)}. \} \end{array}$$

のように NAF の成立を否定するものである。

論証に対して反論があり、その反論に対してさらに反論がある、……という場合に、結果としてどの論証が優先するのか、については論証枠組み AF[4]が安定モデル、有礎モデルほかのさまざまな意味論を与えている。しかし、AF はルールの優先関係を考慮していない場合しか考慮していないため、ルールの優先関係を扱える新たな論証枠組みの研究が必要である。

#### 3.8.3.3 価値観のバランス

上記の法的な議論のモデル化において、ルールの優先関係を確定することができれば、非単調推論によって、優先する論証を決めることができる。しかし、このルールの優先関係をどのように確定するかは困難な問題である。

ルールの優先関係を客観的に決めるには、価値観のバランスを数理的に測る必要がある。価値観がどのようなファクタからなっているかを分析する研究として、過去の判例を使って統計的な処理を行い、どのようなファクタが判決に強い因果関係があるかを推定する方法がある。わが国では 1970 年代にこのような統計的な研究が行われた。例えば、刑事事件について、被告の年齢、性別、累犯かどうか、被害の程度、被害者の年齢、性別などのさまざまな要素と判決（有罪/無罪の決定や量刑）の間の相関関係を多変量解析の手法で調べ、判決予測や刑事政策などにいかそうとする研究がなされている。しかし、このような統計的手法は、予測結果についての理論的な裏づけがないためか、わが国の法理学の分野では 1980 年以降は低調である。

しかし、SplitUp[11]のように、ニューラルネットの手法を使って、大量の判例から判決予測をする方法が予測精度の点で成功を収めたことが報告されている。判例の統計的手法は今後、もっと注目されるべきだと考える。

#### 3.8.3.4 法的推論システム構築方法論

従来の法的推論システムの開発においては特にシステムの構築方法論が確立していなかった。法律知識のオントロジーの構築に関しても国際的な標準化に関して進んでいるとは言えない。その理由の 1 つは、法律体系が国によって異なるため、概念やルールの統一化が困難だからである。

しかし、近年、エンジニアリング的な観点からのシステム構築論が提案されるようになってきている。これは、法律のオントロジーに基づくドメインモデルと、法律のさまざまな問題解決のパターンからなるタスクモデルを利用している。

ドメインモデルとしては、法律の知識をカテゴリ（法令、判例、世界知識、……）に分類し、カテゴリごとに概念階層やルールベースや事例ベースや義務論理の論理式集合として記述したものをを用いる。

タスクモデルとしては、例えば、相続税を軽減するというゴールを達成するためには、(1) そのゴールを達成するためのサブゴールを列挙し、(2) 特定の観点からその 1 つを選択し、(3) 因果関係を利用して、そのサブゴールを達成するためのプランを策定する。このような作業パターンはビジネスや設計の分野で CommonKADS のようなタスクモデルとしてまとめられており、それを利用することによって作業手順を新たに決める必要から解放されることになる。

### 3.8.4 新しい展開

人工知能と法律の国際会議においては、実験的な研究や理論的な研究が多く、実システムとしての発表は多くなかった。しかしながら、近年、実用システムとしての法的推論システムが有望と思われる分野が現れてきた。

#### 3.8.4.1 有望な分野

##### 3.8.4.1.1 オンラインの紛争解決

紛争があった場合、通常の人が弁護士に相談したり、裁判に訴えたりすることは心理的・経済的が大きいの。そこで、近年、インターネットにより、オンラインで法律相談をするサービスが現れてきた。なかでも裁判外紛争処理 (Alternative Dispute Resolution: ADR) は、今後、法的推論システムの応用分野として有望であると思われる。

ADR とは、紛争を調停または仲裁によって解決しようとするものである。非法律家でも行うことができ、例えば、アメリカでは法学部の学生が行うこともあるようである。日本では、国民生活センター、交通事故紛争処理センター、日本クレジットカウンセリング協会、国際商事仲裁協会、全国貸金業協会連合会などが ADR を行っている。

調停は双方の意見を聞きながら合意に向けた解決案を探すのであるが、仲裁は裁判と同じように双方が対立し、判決と同じように裁決がなされる。ただし、裁判のような 100 対 0 の決着ではなく、互いに妥協を強いるような決着をすることもできる。

オンライン ADR とは、電子メールで申し立てや答弁を行い、ビデオ会議による審問を行い、Web 上で議論をするものである。海外では、1996 年のオンライン仲裁実験プログラム Virtual Magistrate (米)、オンライン調停実験プログラム Online Omnibus Office (米)、1999 年のオンライン調停仲裁実験プログラム Cyber Tribunal (カナダ)、オンライン調停システム e@DR (シンガポール) などが開発されている。

ユーザは気軽に相談でき、遠隔地で利用でき、安価であるという点でオンライン ADR は優れているが、ネットを使った議論では臨場感が出ないという問題点がある。

ADR の別の問題点は相談員の確保である。各地の消費者センターにおいては、相談件数に応じた相談員がいないため、対応に遅延が生じることが少なくない。そこで簡単な事件、典型的な事件については、ADR の相談機能を有するエージェントで代行したり、相談員に助言を与えたりすることが必要になると考えられる。

##### 3.8.4.1.2 紛争の予防

インターネットの普及により、だれでもオンラインによってショッピングができるようになった。しかし、これによって商品購入に関するトラブルが増加しているのも確かである。トラブルで多いのは、金を払ったのに商品を送ってこないケースと、解約しようとし

たら高額な賠償金を請求されたケース、知らない間に国際電話を長時間使っていたケースなどである。これらについては、トラストマークの付与や、解約に関する法律知識を使って、ある程度対応することができる。

しかし、これ以外に問題となるケースで予想以上に多いのは、消費者の法律に関する知識がないため、消費者に否があるのに、消費者がクレームを出すケースである。例えば、内容をよく読まないで YES にクリックしてしまったケースや、クーリングオフ期間の経過後にキャンセルを出すケースなどである。このようなトラブルに対処するには、ショッピング時に、「これをクリックすることにより、どのような法的効果（権利や義務など）が生じるか」をユーザに助言したり、契約の成立後にキャンセル方法についての説明をユーザにしたりすることが有効である。

このように法的知識を有し、ユーザの契約行為を監視して助言を与えるシステムは今後、ニーズが高まると考えられる。特に、海外の業者から購入する際のトラブルに関しては、両国の法律が関連する場合があるので、両方のルールベースを具備するか、あるいは、自国のルールベースを有するエージェントと、相手国のルールベースを有するエージェントの間の情報交換機能を必要とする。

#### 3.8.4.1.3 法学教育

わが国では法科大学院の制度が始まることになり、現在、各大学や弁護士会などで、設立準備が行われている。法科大学院においては、単に講義をするだけでなく、実務に直結するようなトレーニング（例えば模擬裁判）が必要となる。このようなトレーニングは個々の学生ごとに異なる指導をしなくてはならず、しかも教師が長時間束縛されるので、あまり頻繁に行うことができないという問題点が有る。そこで、コンピュータを用いたトレーニング支援システムが必要になると考えられる。具体的には、ビデオ教材の積極的な利用、学生の応答をモニタしながらの講義、オンラインの模擬裁判などの新しい教育システムの開発が期待される。例えば、オンライン模擬裁判においては、原告学生と被告学生の論争の推移をモニタしながら、新しい論点へとナビゲートしたり、論争終了後に論争記録を整理したり、それをもとにして採点の素点をつけたりする機能を自動化する機能が必要となる。

#### 3.8.4.2 課題

上で説明した ADR や法学教育を行うシステムを実現するには、法的推論の技術を開発するだけでは不足である。

ADR に申し込むユーザは、トラブルをかかえているのであり、対応が冷たいと、ユーザ調停や仲裁自体を拒否するかもしれない。そのため、擬人化エージェントを利用して、ユーザに自然なインターフェースを提供する必要がある。

また、現在のオンライン ADR での議論は臨場感がないため、調停や仲裁の効率がよく

ないという問題がある。臨場感を伝えるための新しいマルチモーダルインタフェース技術が必要である。

さらに、調停においては、相手に対して説得する機能も必要とされる。アメリカにおいては、法廷における弁論術や身振りなども教育の対象となる。

模擬裁判においては、教師エージェントは裁判の争点をダイレクトにナビゲートするだけでは学生の教育にならない。ヒントを出して学生に争点を気づかせる工夫が必要である。教育のためのヒントの出し方は教育心理の問題である。

このようにこれらの新しい応用分野で実用的なシステムを開発するには、法律をベースとした推論技術だけでなく、ヒューマン・エージェント・コミュニケーション技術、マルチモーダル情報処理技術、交渉技術、教育技術などが必要となる。

### 3.8.5 まとめ

法的推論研究の動向と新しい展開として有望な分野について述べた。法的推論システムは基礎研究の時代を脱却して実用化に向けた新たな時代に入っている。ここに述べたシステムは法律の知識だけでなく、さまざまな技術を結集しなければ開発することはできない。したがって、これらのシステムはマルチモーダル技術その技術の実証の場を提供することになり、人間主体の情報処理の重要なアプリケーションの1つとなる可能性がある。

#### 参考文献

- [1] Alevan, V.: Evaluating a Learning Environment for Case-Based Argumentation Skills, Proc. IJCAIL97, pp. 170-179 (1997).
- [2] Ashley, K. D.: Modeling Legal Argument: Reasoning with Cases and Hypotheticals, MIT Press (1990).
- [3] Brewka, G.: Well-founded Semantics for Extended Logic Programs with Dynamic Preferences, J. Artificial Intelligence, 4, pp. 19-36 (1996).
- [4] Dung, P.M.: An Argumentation Semantics for Logic Programming with Explicit Negation, Proc. 10th Logic Programming Conference, pp. 616-630 (1993).
- [5] Gordon, T.: The Zeno Argumentation Framework, Proc. ICAIL97, pp. 10-18 (1997).
- [6] Kowalski, R. A., et al.: Abstract Argumentation, Artificial Intelligence and Law, 4, pp. 275-296 (1996).
- [7] Loui, R. P.: Progress on Room5, Proc. ICAIL97, pp. 207-214 (1997).
- [8] Okubo Y., et al.: Construction of Predicate Mapping for Goal-Dependent Abstraction, Algorithmic Learning Theory94, Springer LNAI 872, pp. 516-532

(1994).

- [9] Prakken, H., et al.: Argumant-based Extended Logic Programming with Defeasible Priorities, *J. Applied Non-Classical Logics*, 7-1, pp. 25-75 (1997).
- [10] Sakama C., et al.: Representing Priorities in Logic Programs, *Proc. Int. Conf. and Symp. on Logic Programming*, pp. 82-96 (1996).
- [11] Zeleznikow, J., et al.: Knowledge Discovery in the Split Up Project, *Proc. ICAIL97*, pp. 89-97 (1997).



## 3.9 音楽情報処理研究の新しい潮流

報告者: 平田圭二委員

### 3.9.1 はじめに

本節では、近年の音楽情報処理に関する新しい研究課題の潮流を概観し考察する。その研究課題の特徴は以下の3点にまとめられる。

- (1) 音楽理論を援用し音楽の意味を考慮した処理を実現すること。
- (2) ユーザに提供するタスクが作曲、編曲、演奏という大粒度のものから検索、集約、推論という中粒度でより普遍的なものに変化したこと。
- (3) 応用システムをインターネット/Web上に展開すること。

このような傾向が加速される背景には、実用的な音楽応用に対するニーズの高まり、ユーザのライフスタイル変化に伴う音楽応用の在り方の変化、Webを中心とする計算機環境や状況の変化などが影響している。

以下では、(1)に関連して、計算機への実装に有用な音楽理論を紹介する。この音楽理論は音楽と音楽に対する処理を計算機上のプログラムとして記述できるほどに形式化することが期待される。次に、そのような形式化を前提とすると、音楽知プログラミングという考え方へ発展させることができる。これは(2)に関連して、音楽に対する処理レベルや種類を再考し、より基本的かつ普遍的な枠組みを提供する試みである。最後に、(3)に関連する国際会議を紹介する。

### 3.9.2 計算機への実装に有望な音楽理論

音楽理論とは楽譜に書かれた楽曲を分析、理解するための理論である。10世紀付近よりポリフォニー（多声音楽）<sup>1</sup>の時代が始まるのに合わせて、音楽理論は作曲の技法として誕生、発展してきた。音楽は一般にメロディ、ハーモニー、リズムの三要素から成ると言われており、それぞれを分析理解するための多くの理論がこれまで提案され、それらがその時代ごとの作曲法や音楽の規範としての機能を果たしてきた[10]。代表的なものとして機

---

<sup>1</sup> Polyphony。複数の独立した声部を持つ音楽のこと。

能和声法（18世紀）や対位法（17世紀）などが挙げられる。しかし、これらの音楽理論は当然のことながら計算機上への実装を意図して構築されていないため、暗黙的な知識を前提としていたり、記述があいまいであったり、定義が不完全である。そして、前述したように音楽理論は楽曲の分析手順を記述したものであり、合成手順を記述したものではない。さらに、対象としている音楽ジャンルがクラシックのバロックや古典派に限定されている。よって結果的に、計算機上で実用的な音楽応用を作成する際に役立つとは言い難いものとなっている。

20世紀に入り、より形式的な音楽理論体系の構築を目指す動きが現れる。その先駆けとなったのが Schenker 理論である（1906年）[8]。Schenker 理論は簡約（reduction）という概念を導入し、西洋調性音楽に馴染んだ聴取者が音楽を理解することとは、すべての上手く作曲された楽曲は簡約して行くとある基本形に到達するというものであると主張し、その手順を示した。この Schenker 理論をベースにして、のちに楽曲を階層的なグループとして分析、理解しようとする理論 Generative Theory of Tonal Music (GTTM) [5]と、楽曲に含まれる音どうしの関係をネットワーク的にとらえる Implication-Realization (IR) モデル[6]が提案される。GTTM と IR モデルが共通して主張しているのは、異なる音楽構造は聴取者に異なる認識を与え、その際聴取者に与える認識には、段落感、終止感、進行感、浮遊感、調性感などがあるということである。それぞれの理論では、そのような認識を支配する重要な音やグループの同定方法が提案されている。ここで注目すべきは、GTTM も IR モデルも芸術性に関する感情や感性については触れていないということと、これらの理論が対象としているプロセスは飽くまで聴取であり作曲、編曲、演奏ではないということである。

### 3.9.2.1 GTTM の概要

GTTM は、音楽に関して専門知識のある聴取者の直観を形式的に記述するための理論として Fred Lerdahl と Ray Jackendoff により提唱された。五線譜に記述された楽曲（ホモフォニー）<sup>1</sup>を表層構造として、それに解釈や分析を加え、そこに内在するさまざまな階層構造を深層構造として顕在化させる。Chomsky の変形生成文法の枠組みにならったと言われている。楽曲を分析する理論はこれまで幾つか提案されているが、その中でも GTTM が最も緻密に構築されており計算機への実装に適しているであろうと考えられている。とは言え、後述するように、それでもまだ計算機上への実装には多大な労力が必要とされる。

この理論は、グルーピング構造分析、拍節構造分析、タイムスパン簡約（time-span reduction）、プロロンゲーション簡約（prolongational reduction）という4つのサブ理論から構成されている。グルーピング構造分析は、連続したメロディをより短い部分（フレ

<sup>1</sup> Homophony. 主声部の旋律に対し簡単な和声的な伴奏を付されたもので、単旋律的に解釈できるような様式。

ーズ)に分割することである。長いメロディを歌うときにどこで息継ぎすべきかを見つけるような分析である。拍節構造分析は、楽曲の拍(指揮者がタクトを振るタイミング、あるいは聴取者が曲に合わせて手拍子を打つタイミング)を見出し、さらに4分音符、2分音符、1小節などのレベルにおける強拍と弱拍を同定する分析である。タイムスパン簡約は、あるメロディはそれより長いメロディを簡約化、抽象化することによって得られるという直観を表現したものである。例えば、ある楽曲の全体は最終的にその調の主音に簡約化されると考える。タイムスパン簡約は、グルーピング構造分析と拍節構造分析の結果に基づいて、部分を合わせて全体にまとめあげるという意味でボトムアップに行われる。プロロンゲーション簡約は、曲の進行に関する次の2つの直観を表現したものである。1つは、メロディ中のある音はそれまでの音の繰り返しとしてあるいは次の展開をもたらすために出現しているというものである。もう1つは、メロディの全体や部分は、一般に、緊張と弛緩というパターンを持っているというものである。プロロンゲーション簡約の実行は、タイムスパン簡約の結果に基づいて、トップダウンに行われる。つまり、楽曲中から繰り返しのあるいは展開的に最も重要な音を選び出し、その音を手がかりにして楽曲を分割していく。

楽曲を分析することは、タイムスパン簡約およびプロロンゲーション簡約を行うことである。これらの簡約結果は、各々タイムスパン木およびプロロンゲーション木として表現される。タイムスパン木の例を図3.9-1に示す。

GTTMの規則は、分析を進めて木構造を生成するための構成規則(well-formedness rule)と構成規則適用に関する知識である選好規則(preferecne rule)の2種類から成る。

これらの木が、GTTMの意味で楽曲の深層構造を表現している。人それぞれで楽曲の解釈が異なるように、1つの楽曲には、一般に、複数通りの分析が可能である。これは、GTTMの分析では選好規則による解釈の絞り込みが働いても、タイムスパン木およびプロロンゲーション木が複数通り導かれることに対応する。

Level a

Level b

level c

Level d

Level d

level c

level b

Level a

文献[5] p. 115 より転載

図 3.9-1 タイムスパン木の例

### 3.9.2.2 GTTM 実装の困難さ

GTTM は音楽理論としては比較的形式的に記述されているものの、そのままプログラムとして実装できるほどには形式化されていない。例えばまず、グルーピング構造分析の次の選好規則を考える。

GPR3: 4つの音符 ( $n_1$ ,  $n_2$ ,  $n_3$ ,  $n_4$ ) が連続しているとき、以下の条件が成立すると、 $n_2$  と  $n_3$  の間がグループ境界と認識される。

- (a)  $n_2$ ,  $n_3$  間の音高差が  $n_1$ ,  $n_2$  の音高差および  $n_3$ ,  $n_4$  の音高差よりも大きい。
- (b) 音高に関して  $n_2 \neq n_3$  かつ  $n_1 = n_2$  かつ  $n_3 = n_4$ 。

この選好規則だけならプログラム化は容易であろう。しかし、複数の選好規則間で競合が生じているような場合の解決は難しい。実際、グループ境界を判定する規則は音高だけでなく、音量に関する規定されている。例えば上の例において同時に  $n_3$ ,  $n_4$  の間で音量に関するグループ境界が認識されているとすると、この4つの音符においてグループ境界の候補は  $n_2$ ,  $n_3$  の間と  $n_3$ ,  $n_4$  の間となり競合が生じる。しかし GTTM には選好規則の競合解消に関する規則は不十分（部分的、あいまい）にしか記述されていない。

またグルーピング構造分析には次のような規則もある。

GPR4: GPR2、GPR3 がより多く成立する境界が、上位レベルでもグルーピング境界になる可能性が高い。

GPR5: グループを2等分するグルーピングを優先する。

GPR6: 2つ以上の並列した部分 (parallel) を含むようなグルーピングを優先する。

これらの選好規則には、「可能性が高い」、「優先する」という記述が含まれているが、これらの用語をどう解釈すべきかは記述されておらず、このような規則の適切な実装は一般に困難である。現在このような困難さを克服する技術が囑望されている。

また、GTTM に関しては Alan Ruttenberg による Web 資料[7]も読みやすいので参照されたい。

### 3.9.3 音楽知プログラミング

GTTM や IR モデルなどの音楽理論による音楽の形式化が実現すれば、その上に構築される処理の枠組みにも変化が生じるであろう。これまで音楽というものは一部の専門家だけが作曲、編曲、演奏などに従うことができ、そうでない人々は単に聴いて楽しむことしかできなかった。しかし計算機による支援が得られるのであれば、一般のユーザでも

音楽を自由に処理することが可能となろう。実際、計算機による支援が最も進んでいるのは自然言語の分野であるが、そこでは多くのユーザがさまざまな状況でさまざまなスタイルの文章を自由にやりとりし、コミュニケーションを行っている。音楽に関しても同様の状況が実現できれば、大変有意義なことであると思われる。

さて、ではどのような方針で一般ユーザの音楽活動を計算機で支援すれば良いのだろうか。先に、音楽理論が目標としているのは、聴取者が認識する段落感、終止感、進行感、浮遊感、調性感などと楽曲構造との関連性の解明であると述べた。このレベルの認識に対応するタスクは、検索、集約、推論などであり、専門家が行う高次の作曲、編曲、演奏というタスクではない。ここで、作曲、編曲、演奏といった専門的なタスクを大粒度のタスクと呼び、検索、集約、推論などのタスクを中粒度のタスクと呼ぼう。すると、専門家は、各自の美的センスや芸術的センスに従って中粒度のタスクを組み合わせることで大粒度のタスクを実現していると考えられる。ここで、前節で述べたような音楽理論を援用し、中粒度のタスクを自由に組み合わせる環境を一般のユーザに提供することができれば、一般のユーザも各自のセンスに従って大粒度のタスクを組み上げ、音楽でコミュニケーションを行うことが可能となるだろう。

#### 3.9.3.1 音楽知とは何か

コンピュータ上での形式化を前提とし、音楽理論に基づく構造によって意味が与えられている音楽全般に関する知識や知性の集まりは音楽知と呼ばれる[2]。実際、音楽知にはさまざまな形態や抽象度のものがある。五線譜に書かれている楽曲、音楽理論に基づく楽曲の分析結果、教科書で述べられている知識、音楽学から得られた知見や解明された事実、音楽家の頭の中にものみある暗黙的な知識や知性などである。

音楽知をコンピュータ上で操作する際に考慮すべきは次の3点である。

- (1) 音楽知そのものに含まれる規則や構造。
- (2) 音楽知に含まれるその構造の表現。
- (3) 表現された音楽知に対する演算やアルゴリズム。

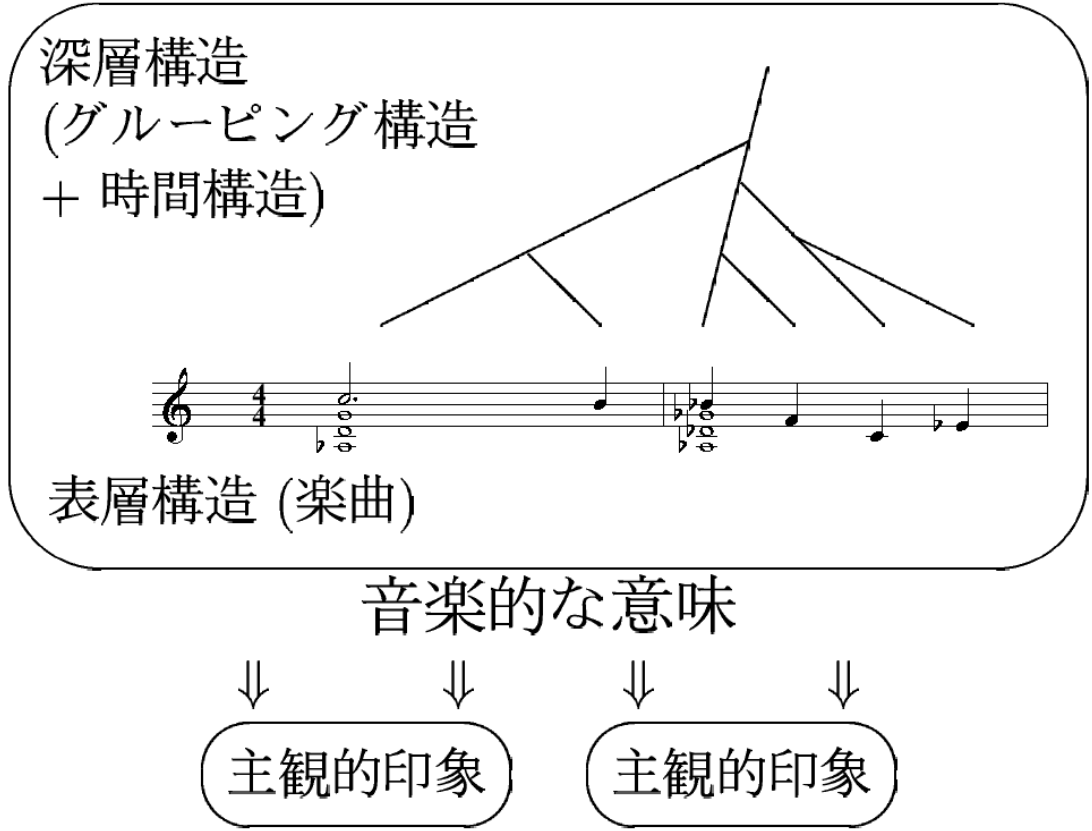


図 3.9-2 表層構造、深層構造、意味、主観的印象

(1) に関しては、前節に述べたような音楽理論に依拠するのが良いだろう。楽曲を五線譜で表す場合、音符や休符などが五線の適切な場所に置かれる。これが人間の目に見える表層のレベルである。例えば和音進行も表層レベルの情報である。しかし五線譜の裏には何らかの音楽的な深層構造が隠されている (図 3.9-2)。

深層構造を顕在化する基本的な考え方は、楽曲に含まれる個々の音に関してある種のまとまりを見い出すことである。まとまりを見い出すことは一般に、グルーピング、体制化、構造化、関係づけなどと呼ばれる。そして楽曲分析を行い楽曲に含まれる音をグルーピングし深層構造を顕在化させることが音楽理論の役割である。

人は無意識的、意識的にかかわらずその深層構造の情報も加味して音楽的活動を行っている。これら深層構造と表層構造を合わせて意味が構成される。したがって、音楽知を取り扱うということは、表層構造だけでなく深層構造も含めてコンピュータ上で取り扱うということである。

主観的印象や情動の取り扱いはまた別の問題である (図 3.9-2)。従来の音楽研究や音楽システム作成では、この主観的印象や情動を音楽知と混同して処理している事例が散見さ

れる。例えば「太郎は駅へ向かった」という文を考えてみよう。この文を解釈することにより、太郎という人がどこかの駅へ向かって移動するという文の意味が理解される。次にこのような認識や理解に基づいてこの文を読んだ人には楽しいという情動や悲しいという情動が喚起あるいは連想されるかも知れない。この情動は現段階で体系的に理解できる対象のように思えない。したがって、自然言語処理では、このように理解された文の意味と喚起・連想された情動は区別して取り扱われる。音楽知を取り扱う場合もこの自然言語処理と同様にまず音楽知とそこから恣意的に喚起・連想される主観的印象や情動を区別し、それぞれの処理手法も区別して考えるべきである。

音楽はそもそも多角的な意味を持っていると考えられるので、複数の音楽理論を用いて多角的に楽曲分析する方がより直観に合致しているように思われる。そうして、より精密な深層構造を顕在化することで、音楽家やユーザの意図を正確に認識できるようになる。

(2) に関しては、人工知能における知識表現技術を応用するのが良いだろう。図 3.9-2 の表層構造や深層構造は、コンピュータ上の記号や項として表現される。このとき、知識表現手法に求められるのは、表現すべきことを簡潔に述べられる記述力と正確さである。1 階述語論理はこの基準を満たす多くの知識表現言語の基礎を成している。

(3) に関しては、音楽知を表現する記号や項に対して適切で意味のある演算や推論体系を設計しなければならない。例えば、整数を対象とすれば、加減乗除の四則演算で実用上十分な関数が構成できる。あるいは、真偽の値を持つ命題を対象とすれば、論理和、論理積、否定の 3 種の演算ですべての論理関数が実現できる。同様に、音楽知を表現した記号や項を対象としたとき、どのような種類の演算や推論を用意すれば、すべてのあるいは実用上十分な音楽的な関数が実現できるのだろうか。これは今後解決すべき重要な課題だと思われる。

### 3.9.3.2 音楽知をプログラミングする

ここまでの検討より、音楽知をコンピュータ上の記号や項として表現する手がかりが得られたものとしよう。次に、現実の音楽知を十分直感的に表現する記号や演算を過不足なく用意し、ユーザがそれらを適切に組み合わせる（プログラミングする）方法論を検討する。音楽知プログラミングとは、音楽知を対象とする推論を思い通りに制御し、所望の音楽タスクを実現することである。このような音楽知プログラミング環境が実現できれば、ユーザは所望の音楽システム、アプリケーション、機能を自由自在にデザインできるようになることが期待される。

音楽的な状況において人が実行しているあらゆる中～大粒度のタスクが音楽知プログラミングの対象となる。例えば、ハーモナイズ、ボイスンク、リハーモナイズ、アドリブ、本番演奏のための練習、試行錯誤と自己評価、異なる形式/様式/形態で獲得した音楽知の統合、記憶、連想、学習、適用、データマイニング、音楽的に合理的な UI や可視化、実時間インタラクションなどである。



音楽知プログラミングは、与えられた基本要素を組み合わせることで所望のソフトウェアを構築するという点では従来のプログラミングと同様である。しかし、計算の対象となる音楽知は、1つの表層構造と何種類かの深層構造が組になっている。表層構造が同じでも異なる深層構造を持っている場合がある。したがって、組になっている深層構造とユーザが付与しようとしている深層構造が同じか異なるか、異なる場合はどのように異なるかをプログラミング中に検査、ユーザ提示する手段が必要となろう。

音楽知プログラミングの特徴は以下のようにまとめられるだろう。

- (1) 音楽知を表現する記号や項には、複数の音楽理論に基づく多角的で一貫した意味が多重に付与されている。
- (2) 音楽的な意味を反映した基本演算を備えておりその意味は直感的に理解できる。
- (3) 楽曲などの表層には (1) で述べたような深層構造 (= 意味) が付与されるが、その (部分的な) 深層構造を最初に (自動的に) 表層構造に付与し、のちにその深層構造を追加、修正、更新する処理が必要である。

ここで注意すべきは、プログラミングという作業が成立するための条件である。その最も重要な条件は、どのようなプログラムを書くときどのような結果や振る舞いが現れるのかをユーザが精度良く予測できることである。音楽知プログラミングはこの条件をクリアすることができるであろうか。もちろん、我々はこの条件をクリアすべく音楽知プログラミングの枠組みを構築しなければならない。上の (3) の特徴が従来の情報処理にはなかった新しい課題を提起するのではないかと思われる。

### 3.9.4 関連する国際会議

前述したように、音楽理論への注目や音楽情報処理の新しい枠組みの提案は、音楽応用に対するニーズの高まりがその原動力となっていた。一方、計算機環境の変化が、次世代で必要とされる技術の流れを変える原動力となる場合もある。本節で紹介する WEDELMUSIC や ISMIR という国際会議は、Internet/Web の出現が引き金となって誕生した音楽関連の研究分野を象徴している。

#### 3.9.4.1 International Conference on Web Delivering of Music (WEDELMUSIC)

その発端は 1998 年に開始された ESPRIT プロジェクトであり、その目標は、音楽を創作、出版する人々と、音楽をあらゆる形式で利用する人々の間をインターネットで結び、安全、安心、自由に音楽をやりとりすること、すなわち Web で音楽配信を行うための標準アーキテクチャを提案することである。技術目標として、XML ベースの統一フォーマッ

トの設計、インターネット上のデータ保護技術の開発、さまざまなツール群の開発が掲げられている。現在、音楽配信、音楽記述、応用などのためのフォーマットは乱立しており、XML に基づくものとそうでないものに大きく分けられる。XML に基づくものとしては MusicXML、WEDELMUSIC などがあり、そうでないものとしては Finale、Sibelius、Score、MML、SMDL、HyTime、MusicTeX、Lilypond、GUIDO、SMF などがある。このような現状を打開する意図も込められている。

会議は 2001 年に第 1 回目が Firenze にて開催され、その実行委員長は Paolo Nesi (Firenze 大学教授) であった[9]。会議およびプロジェクトの活動母体やスポンサーはヨーロッパ系の大学や企業が中心となっているが、一方では IEEE CS TC on Computer Generated Music が後援している。会議の規模は、参加者 40 名程度、発表件数 26 件+ワークショップ 8 件 (以下、「+」の後ろはワークショップでの発表)。発表内容は、セキュリティ、透かし 6 件、楽譜印刷、認識 3 件+7 件、信号処理 7 件、音楽記述 (XML) 3 件+1 件、応用システム作成 7 件であった。初回の会議のためか WEDELMUSIC 規格以外のフォーマットに関する発表も幾つかあった。しかし WEDELMUSIC 規格に準拠したツールの紹介にある程度の時間を割くなど、標準化を意識したセッション構成となっていた。

#### 3.9.4.2 International Symposium on Music Information Retrieval (ISMIR)

近年の情報検索分野は、Web の検索エンジンというキラアアプリケーションも登場し、非常に隆盛である。そのような流れの中で Web 上に多数置いてある楽曲データ (SMF、AIFF、楽譜エディタのファイルなどの形式) の検索エンジンの研究が技術的にもビジネス的にも有意義であるとの認識から、2000 年に第 1 回 (Massachusetts 大学)、2001 年に第 2 回 (Indiana 大学) が開催された[4]。第 1 回の発表は 13 件、招待講演、チュートリアル、パネルなどが 6 件であり、第 2 回の発表は 20 件、ポスタが 17 件という規模である。報告者は実際に会議に参加したことがないので会議自体の雰囲気は不明であるが、発表件数の比率に関して、オーディオデータに基づく音楽検索と記号的な手法に基づく音楽検索では、おおよそ 2 対 3 程度のようなものである。音楽検索以外の発表としては、インデックスに基づく音楽検索、システム実装、標準化などがある。

#### 3.9.4.3 その他

その他関連の深い会議として、International Computer Music Conference (ICMC) [3] では Internet/Web を利用した作曲ツールや合奏システムが発表され、ACM Multimedia[1]では音楽と他メディアとの統合システムやエンタテインメント関連の音楽システムに関する発表が見られる。

## 参考文献

- [1] ACM Multimedia 2002, <http://mm02.eurecom.fr/>
- [2] 平田圭二: 音楽知プログラミング試論, NTT R&D, 電気通信協会, Vol. 50, No. 9, pp. 682-689 (2001).
- [3] ICMC, <http://www.computermusic.org/>
- [4] ISMIR, <http://ciir.cs.umass.edu/music2000/>, <http://ismir2001.indiana.edu/>,  
<http://ismir2002.ircam.fr/>
- [5] Fred Lerdahl and Ray Jackendoff: Generative Theory of Tonal Music, The MIT Press (1983).
- [6] Eugene Narmour: The Analysis and Cognition of Basic Melodic Structures, The University of Chicago Press (1990)
- [7] Alan Ruttenger: Review and Discussion of A Generative Theory of Tonal Music, <http://alanr.www.media.mit.edu/people/alanr/Jackendoff&LerdahlFinal.html>
- [8] SchenkerGUIDE, <http://www.schenkerguide.com/>
- [9] WEDELMUSIC, <http://www.wedelmusic.org/>
- [10] 矢向正人: 言語ゲームとしての音楽, 勁草書房 (2001).

## 3.10 インタラクティブグラフィックスの動向

報告者: 宮田一乗委員

### 3.10.1 はじめに

CG 技術のいちばん顕著な応用の場としては、映画での特撮やアニメーション制作が挙げられる。一方で、Web ブラウザ上でのコンテンツ、PC や家庭用ビデオゲーム機でのリアルタイム CG 映像など、インタラクティブな CG 映像も広く市場に浸透している。

本報告では、インターネット上でのグラフィックス技術、リアルタイム CG を支えるハードウェア技術とその具体的な展開例である家庭用ビデオゲーム機の現状、さらに、インタラクティブグラフィックスの応用として VR や AR、インスタレーションアートなどの事例について述べ、インタラクティブグラフィックスの将来像を考察する。

### 3.10.2 インターネット上のグラフィックス

インターネットを介して、さまざまな情報がやり取りされるようになった。テキストや画像、映像、音情報の配信と比較して、3次元形状データの配信は、それほど熱心には勧められていないようである。その理由としては、形状データの容量や受け手の再生機能の問題などが挙げられる。本節では、インターネット上での3次元グラフィックスデータの取り扱いに関する、いくつかの試みについて述べる。

#### 3.10.2.1 ポリゴンリダクション

ポリゴンリダクションとは、表示したい物体の画面に表示される大きさや重要度などに応じて、3次元形状データの詳細度を変化させる手法である。この手法は、例えば、インターネット上でのデータ配信時に回線の速度に合わせて適用したり、リアルタイムでの形状データの表示（描画処理速度に合わせる）へ適用したりすることが考えられる。

#### 3.10.2.2 WEB3D

WEB3D[1]は、数年前から CG の新しい応用分野として話題になってきた技術であり、特にインターネットショッピングの利用などに期待されている。インターネットブラウザ上での3次元グラフィックスの表示技術として VRML が一時期話題になったが、その後の展開が伸び悩んでいる。WEB3D が VRML 以上に注目を浴びている大きな理由として

は、以下のような点が挙げられる。

(1) 少ないデータ量

VRML ではポリゴンメッシュで形状を表現するため、対象形状に曲面が含まれる場合にはデータ量が増大する。一方、WEB3D では曲面 (NURBS パッチ) を直接扱うことができるので、データ量が少なくすむという利点がある。

(2) 表現の質感向上

従来のように多面体で曲面を近似して表現するよりも、曲面を直接表示したほうが滑らかな物体を表現できる。人体ばかりでなく、車体などのインダストリアルデザインでは自由曲面を多用しているので、WEB3D でのデータ構造の方が表現の質感向上につながる。また、WEB3D では、平行光線だけではなく点光源や物体表面への映り込みを表現することができる。例えば車体への風景の映り込みを施すことにより、金属の質感を表現することができ、車体のリアル感が向上する。

以上の 2 点が技術的な大きな理由であるが、これ以外に、マッキントッシュ版の開発が遅れたこと (半数のデザイナーがマッキントッシュを使っている)、VRML コンソーシアムの体質 (非営利団体の舵取りで意見の集約が困難であった) など、VRML が予想以上に浸透しなかった理由として挙げられる。

### 3.10.2.3 情報の埋め込み

素材の収集がネットワークを介して簡単に行えるようになったことで、ネット上で入手した素材を加工し組み合わせて作品を仕上げることも可能になった。しかも、でき上がりの作品から、素材のオリジナリティを判別することは困難である。画像や音楽への電子透かし技術は数年前から研究されてきているが、3 次元の形状モデルへの情報埋め込みの技術は始まったばかりである。このような著作権保護のための技術は、動画やモーショングレードータなどの多様な素材への拡大が予想される。今後は、シングルソース・マルチユースの考えのもとに、コンテンツの再利用に注目が集まると考えられ、このような著作権保護のための技術が重要度を増すと考える。

### 3.10.3 リアルタイム CG のハードウェア技術

リアルタイム CG を支えるハードウェア技術の大部分は、搭載されているグラフィックボード (ビデオカード) の技術と言える。GUI の発展に伴い、グラフィックを画面に高速に描画する機能を持つグラフィックチップが多数開発されてきた。グラフィックカードは PC の中で果たす役割が年々増大してきており、その心臓部のグラフィックチップは、CPU

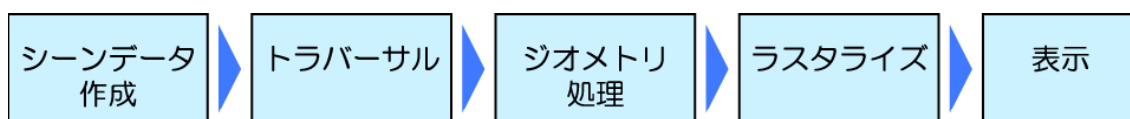


図 3.10-1 3次元グラフィックスの処理の流れ

の進化スピードに匹敵する勢いで進歩している。本節では、3次元グラフィックスの処理の流れと、高速化の手法について述べる。

#### 3.10.3.1 処理の流れ

図 3.10-1 は、一般的な 3次元グラフィックスの処理の流れを示している。各段階の処理の内容は次の通りである。

- (1) シーンデータ作成  
表示するモデルデータなどの作成やカメラおよびライトの設定。
- (2) トラバーサル  
シーンデータを解釈して必要な描画オブジェクトと描画命令を次段階へ渡す。
- (3) ジオメトリ処理  
モデルの座標変換、光源計算、投影変換、クリッピング処理などを行う。
- (4) ラスタライズ  
ジオメトリデータのラスタ化やシェーディング処理、Zバッファ処理、テクスチャ処理などを行う。
- (5) 表示  
ビデオメモリの値を読み出し表示する

#### 3.10.3.2 高速化の手法

図 3.10-1 でのラスタライズと表示の部分は、初期のグラフィックカードからハードウェアで実装されている。現在の 3次元グラフィックカードでは、特にジオメトリ処理の部分をハードウェアで実装し、高速描画を実現している。3次元グラフィックスの描画の高速化は、以下のようなさまざまなアプローチがある。

##### 3.10.3.2.1 ジオメトリ処理の高速化

ジオメトリ処理には、CPUで行う方法（ソフトウェアによるアプローチ）と、グラフィックプロセッサで行う方法（ハードウェアによるアプローチ）がある。

ソフトウェアによるアプローチでは、Pentium の SSE 命令を使用する方法と、AMD の 3DNow!を使用する方法などがある。

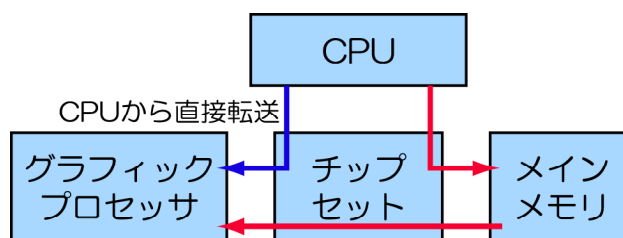


図 3.10-2 データ転送の高速化

一方、ジオメトリエンジンと呼ばれるジオメトリ処理を専用に行うチップがハードウェアとして実装され、最近の3次元グラフィックボードのほとんどもに搭載されるまでになっている。CPUの演算速度が十分速ければ、ジオメトリエンジンは不要であるが、現状ではジオメトリエンジンの処理能力がCPUを上回っている。

### 3.10.3.2.2 グラフィックプロセッサの並列化

グラフィックプロセッサを並列搭載して描画性能を向上させる手法であり、ハイエンドのOpenGL用グラフィックアクセラレータのほとんどが採用している。ラスタライズの分散処理では、ラスタライズプロセッサの数だけ画面を細分割して、各スレッドの処理を割り当てている。

### 3.10.3.2.3 データ転送の高速化

図 3.10-2 に示すように、一般的にCPUで処理されたモデルのデータは、PCのメインメモリに保存されてからグラフィックプロセッサに転送される。そこで、メインメモリにデータを転送せずに、CPUから直接グラフィックプロセッサにデータを転送すれば、データ転送の高速化が可能になる。

### 3.10.3.2.4 nfiniteFX エンジン

nfiniteFX エンジン[2]は、nVIDIA社がグラフィックプロセッサ「GeFORCE3」用に設計したアーキテクチャであり、現在では2世代目のnfiniteFX IIが発表されている。VertexシェーダとPixelシェーダを構成要素としており、この2つはともにプログラマブルである。Vertexシェーダでは、独自のライティングモデルを定義したり、物体表面への凹凸づけのような手続き型のジオメトリ操作も可能である。Pixelシェーダでは、ピクセル単位でのライティングやテクスチャづけなどを定義できる。Microsoft社のDirectX8[3]にも以上の2つのシェーダ機能が含まれているが、これは、XBOX[4]のグラフィックエンジンを開発したnVIDIA社が、Microsoft社にライセンスしたものである。両者が密接に結びつくことにより、DirectXのAPIを介して、ハードウェア機能を用いた高品質なリアルタイムグラフィックスを実現している。

表 3.10-1 家庭用ビデオゲーム機の世代交代

| 世代              | CPU    | 主メモリ  | 画面                      | 補助記憶            | 3次元<br>グラフィックス     | サウンド           |
|-----------------|--------|-------|-------------------------|-----------------|--------------------|----------------|
| 第1世代<br>(72-93) | 8bit   | 2KB   | 256×2404<br>bits/pixel  | ROM             | NA                 | ビープ音<br>モノラル   |
| 第2世代<br>(90-96) | 16bit  | 128KB | 512×44815<br>bits/pixel | ROM             | NA                 | FM音源<br>ステレオ   |
| 第3世代<br>(93-99) | 32bit  | 2MB   | 640×48024<br>bits/pixel | ROM<br>CDROM    | ジオメトリ +<br>ラスタエンジン | FM音源<br>CD並み   |
| 第4世代<br>(99- )  | 128bit | 32MB  | 640×48024<br>bits/pixel | DVDROM<br>CDROM | ジオメトリ +<br>ラスタエンジン | CD並み<br>サンプリング |

### 3.10.4 家庭用ビデオゲームの現状

初めての家庭用ビデオゲーム機「オデッセイ」(Magnavox 社)が登場したのが、札幌冬季オリンピックの開催された1972年であった。その後30年の年月を経た家庭用ビデオゲーム機の現状を本節で簡単に述べる。

#### 3.10.4.1 家庭用ビデオゲーム機の世代交代

任天堂の「ファミリーコンピュータ」が登場(1983年)する以前から、いくつかの家庭用ゲーム機が発売されていた。家庭用ゲーム機の世代交代を表3.10-1にまとめたが、この分類はハードウェアのモデルチェンジにあわせたものである。後述するが、ゲームの外見の質が劇的に変わったのは、3次元グラフィックス機能が搭載され、ゲームの提供に大容量の補助記憶メディアが使われるようになった、第3世代のビデオゲーム機である。



### 3.10.4.2 PS2・ゲームキューブ・XBOX

現在の家庭用ビデオゲーム機は、PS2（ソニー、1999年）、ゲームキューブ（任天堂、2001年）、XBOX（マイクロソフト、2001年）の3つが主流と考えられている。主要な機能の比較を表 3.10-2 に挙げる。

3機種の中では、XBOX がいちばん PC に近いアーキテクチャとなっており、ハードディスクや LAN ポートまで備えている。PS2 では、コアとなる Emotion Engine にジオメトリエンジンが含まれており、このメディアプロセッサを中心に構成されたアーキテクチャとなっている。一方ゲームキューブでは、CPU とメインメモリ以外の要素は、ほぼ Flipper と呼ばれる 1 チップに集積されている。

XBOX では、グラフィックスの API に DirectX を用いており、PC ゲームの開発者が移植しやすいようになっている。一方、PS2 では、LINUX による開発環境をとっており、ミドルウェアの開発が重要な位置を占めている。

描画性能は後発の XBOX が 3 機種の中でいちばん優れており、毎秒 125 メガポリゴン、すなわち 30 フレーム/秒の場合に、1 画面約 400 万ポリゴンで構成されるシーンを表示できる。

### 3.10.4.3 スプライトからポリゴンへ

第 3 世代のビデオゲーム機から、3 次元グラフィックスの表示機能が搭載されるようになった。これにともない、ゲーム制作のスタイルが一変することになる。第 2 世代のビデオゲーム機までは、スプライトアニメーションと呼ばれる 2 次元の画像（通称：ドット絵）を動かすことで、ゲームの画面を構成していた。一方、3 次元グラフィックスの表示機能が搭載されることにより、以下のようなメリットが生じる。

表 3.10-2 機能比較

|                  | PS2                                 | ゲームキューブ                       | XBOX                    |
|------------------|-------------------------------------|-------------------------------|-------------------------|
| CPU              | MIPS ベース Emotion Engine<br>(295MHz) | PowerPC ベース Gekko<br>(485MHz) | Pentium III<br>(733MHz) |
| 主メモリ             | 32MB                                | 43MB                          | 64MB                    |
| グラフィックス          | Graphics Synth (147MHz)             | Flipper (162 MHz)             | NVIDIA 製<br>(250MHz)    |
| 表示性能<br>(ポリゴン/秒) | 66 メガ                               | 6-12 メガ                       | 125 メガ                  |

- 表現の自由度の拡大  
3次元の物体を扱えることにより、より複雑で自由な空間を表現できるようになった。特に、物体の表面にテクスチャを張り込むことにより、映像の質は格段に向上した。
- 滑らかなアニメーション  
スプライトアニメーションでは、例えば回転する物体を表示する場合、それぞれの回転角に応じた2次元画像をあらかじめ制作して保持しておく必要がある。3次元グラフィックスでは、ポリゴン（3次元の面片）をベースに物体を表現し、それに直接動きを与えて表示できるので、スプライトアニメーションよりも滑らかな動きが表現できる。
- 多彩なカメラアングル  
3次元空間内で自由にカメラアングルを設定できるので、例えば、レーシングゲームでは、俯瞰視点とドライバ視点とを自由に切り替えることが可能である。
- データ量の削減  
スプライトアニメーションでは、すべての画像を制作して保持する必要があるが、3次元グラフィックスでは、物体のポリゴンデータとカメラアングル、動きなどのパラメータを保持していれば、さまざまなアニメーションを表現できる。データ容量は、同じ物体を表現する場合には、2次元の画像データの方が圧倒的に多くなり、かつ制作の作業量も増大する。

#### 3.10.5 インタラクティブグラフィックスの応用事例

ビデオゲーム以外のインタラクティブグラフィックスの応用事例として、バーチャルリアリティ（VR）や拡張現実感（AR: Augmented Reality）、芸術への応用としてのインスタレーションアートなどが挙げられる。本節では、これらのいくつかの例を挙げる。

##### 3.10.5.1 VR / AR の例

VR や AR の例としては、医療への応用が比較的多い。医療現場での経験や訓練を積むためのツールとして研究開発されたものがいくつかあり、例えば、国立がんセンターでの脳腫瘍手術シミュレーションシステム[5]や、米国ノースキャロライナ大学での乳ガン手術のサポートシステム[6]などが挙げられる。後者の例では、頭部につけたトラッキング装置で術者の頭の位置を計測しながら、超音波で撮った病巣の部位をリアルタイムにCGで生成する。術者は、生成されたCG画像をヘッドマウントディスプレイ上で実際に見ている患者の身体と重ね合わせて見ることができ、病巣にうまく治療用の針を刺せるようなサポートが受けられる。

一方、米国MITの石井グループ[7]では、「タンジブル・ビット」というコンセプトで、

GUIに替わる新しいユーザインタフェースの研究を行っている。このコンセプトでは、電子的なデジタルの情報に物理的な形を与えることで、インタフェースの実体感が向上し、かつ実世界と仮想世界とがシームレスにつながると考えている。例えば、「I/O Bulb」では、テーブル上に置かれた建築の模型にある操作を施すと、その建物の影が計算されてテーブル上に表示される。ディスプレイ上の仮想空間内でのモデルに対してではなく、実体として存在する模型に影が投影されることで、より直感的に理解できるインタフェースとなっている。

上述の応用以外に、スーパーマーケットの内装や陳列の検証、オーダーメイドのキッチンの使い勝手の検証、オートバイの運転シミュレータによる事故回避の学習など数々の試みがなされている。

### 3.10.5.2 インスタレーションアートの例

インスタレーションアートとは、1970年の後半ごろから、絵画や彫刻のようなファインアートの範疇で括れない作品に対して多く用いられるようになった用語である。現在では、コンピュータを用いたインタラクティブな作品が多く見受けられるようになり、日本人の作家では岩井俊雄[8]が著名である。彼の作品では、音と映像のインタラクションをテーマにしたものが多く、坂本龍一とのコラボレーションである「Piano」や、「レゾナンス・オブ・フォー〜4つの共鳴」などが代表作品として挙げられる。これらの作品では、演者（もしくは観客）が打鍵したり、テーブル上に設置されたダイヤルを回したりすると、コンピュータがそれらのイベントに従ってリアルタイムで映像を生成し、大スクリーンに投影する仕組みになっている。

海外では、SIGGRAPH ([www.siggraph.org](http://www.siggraph.org)) や Ars Electronica ([www.aec.at](http://www.aec.at))、Milia ([www.milia.com](http://www.milia.com)) などが、インスタレーションアート作品のアカデミックな発表の場として有名である。

### 3.10.6 おわりに

以上、インタラクティブなCGの構成要素と応用例などを簡単に述べた。日本は、家庭用ビデオゲームの分野において常に世界のトップの座に位置しているが、PCゲームの展開を目論むXBOXの登場により、今後はゲーム市場の混戦が予想される。さらに、インターネットを巻き込んで、家庭用ビデオゲーム機は単なるビデオゲーム機ではなく、家庭におけるメディアセンタとしての地位を確立するようになるかもしれない。一方で、急速に普及したモバイル端末も、CGの新たな応用分野として考えられる。近い将来、すべての情報端末がシームレスに接合し、インタフェースとしてのCGが果たす役割はますます重要になると予想する。

#### 参考文献

- [1] <http://www.web3d.org/>
- [2] [http://www.nvidia.com/view.asp?IO=feature\\_nfinitefx](http://www.nvidia.com/view.asp?IO=feature_nfinitefx)
- [3] <http://www.microsoft.com/japan/directx/default.asp>
- [4] <http://xbox.jp/>
- [5] 青野雅樹ほか: 仮想環境における脳手術シミュレーションシステム, 情報処理学会グラフィクスと CAD 研究会, Vol. 95, No. 47, pp. 41-46 (May 1995).
- [6] Andrei State et al.: Technologies for Augmented Reality Systems: Realizing Ultrasound-guided Needle Biopsies, Proceedings of SIGGRAPH 96, pp. 439-446.
- [7] <http://tangible.media.mit.edu/>
- [8] <http://www.iamas.ac.jp/~iwai/>

## 3.11 検索履歴を用いた情報検索サポート

報告者: 山名早人委員

### 3.11.1 はじめに

インターネットの普及により、WWW 上に蓄積される情報が急増している。このように急増する多種多様な情報を検索するために Yahoo![1]や Google[2]など数々の情報検索システムが存在する。

情報検索システム（サーチエンジン）とは、検索語をユーザに入力してもらい、あらかじめ生成しておいたインデックスに基づいて検索を実行し、結果をユーザに返すシステムである。しかし、その検索精度<sup>1</sup>や再現率<sup>2</sup>には問題があることが指摘されており、品質向上のためにさまざまな工夫が試みられている。その1つとして、Google では HTML のリンク構造をランキング<sup>3</sup>に利用する PageRank 法[3][4][5]を提案し利用している。PageRank 法は、「重要なサイトからの被リンク数<sup>4</sup>が多いほど重要な情報である」という経験則に基づき検索結果をランキングする手法である。すなわち、多くの人が良い Web ページであると評価しているページを検索結果の上位にランキングする。Web 上には膨大な情報があるため、このようなリンク構造を用いたランキングは、膨大な情報の中から有用な情報を上位にランキングするための1つの有効な手段となっている。

しかし、すばらしいランキング手法を用いたとしても、ユーザが的確な検索語を入力しなくては、情報検索システムも的確な検索結果を返すことができない。一般に、情報検索システムに入力される検索語は 1~2 語であるため、的確な検索結果を返すことが非常に困難となっている。例えば、検索語として「コンピュータ」が入力されたとする。この場合、ユーザが「コンピュータの機種一覧を知りたい」のか、「コンピュータの操作方法を知りたい」のか、情報検索システムにはわからない。従って、「コンピュータ」という検索語での的確な検索結果を返すことは、一般的には不可能である。

上記のような問題に対する解決策として、いわゆるポータルサイトなどでは、ユーザに嗜好などの情報を登録させ、ユーザの嗜好情報に合致した情報を提供する仕組みを提供し

---

<sup>1</sup> 検索結果の内、何%が検索語に対する検索結果として正しいかを示す割合。

<sup>2</sup> 本来検索結果として出力すべき Web ページの内、何%を出力できたかを示す割合。一般に Web を対象とした検索システムでは、全 Web 空間を定義できないため、再現率を正確に算出できない。

<sup>3</sup> 順位づけすること。

<sup>4</sup> リンクされている数。

ている。このような仕組みは、一般的に情報フィルタリングと呼ばれる。例えば、Yahoo! は、1996年7月から Web 上で初めての大規模なパーソナライズを開始した[6]。Yahoo! では、My Yahoo! と呼ばれるページでユーザの嗜好に合わせて自分専用の Web ページを構築できる。例えば、ユーザがパーソナライズしたショートタイムの Web ページを閲覧していて、直後の検索で郵便番号を検索語として入力した場合は、映画館を検索結果として出力する[6]。

情報フィルタリングの手法は、情報検索においては、query clustering[7][8]として応用されている。query clustering では、同一の検索結果にたどりついた検索語群を同一のクラスタに属するとみなして、検索語と検索結果を対応づけておく。そして、ユーザが分類済のクラスタに属する検索語を入力した場合、入力された検索語に対応づけられた検索結果を返すことにより、ユーザが入力する検索語の情報不足を補う。

情報検索においては、ほかに、query transformation、query routing などの手法が提案されている。これらは、広義の意味においてユーザの検索意図を判断する手法に分類できる。

以下では、ユーザの意図を判断するための手法として、これまで提案されている情報フィルタリング、query clustering、query transformation、query routing の手法を概説するとともに、新しい手法として、同一ユーザが検索を行う際の連続する検索語の出現パターンから、ユーザの検索意図を発見する手法について紹介する。

#### 3.11.2 情報フィルタリング

情報フィルタリングは、大きく 2 つに分類することができる。1 つは、content-based filtering であり、もう 1 つは、collaborative filtering[9][10]である。

content-based filtering では、ユーザのプロファイル（嗜好情報）と、コンテンツとの情報とのマッチング<sup>1</sup>を行い、ユーザに合ったコンテンツを推薦する。しかし、最初に登録したプロファイルが正確でなかったり、情報提供者がプロファイル上で想定している分類とユーザが考える分類とが異なっていたりした場合は、うまく機能しない。

collaborative filtering は、content-based filtering が、該当するユーザの情報のみを用いてコンテンツの推薦を行うのに対し、他の同じ嗜好を持つユーザの情報を使って、コンテンツの推薦を行う仕組みである。つまり、あるユーザ A とユーザ B の嗜好が同じであれば、ユーザ A が高い評価を下したコンテンツは、ユーザ B も高い評価を下すと仮定する。例えば、A～E 君までの 5 人がそれぞれ好きな映画タイトル情報があったとする。<sup>2</sup>

<sup>1</sup> 一般には、キーワードマッチングを行う。

<sup>2</sup> 文献[8]を参考に例を作成した。

- A 君: 王様と私、スターウォーズ
- B 君: 王様と私、スターウォーズ、風の谷のナウシカ
- C 君: バットマン、スターウォーズ
- D 君: バットマン、スターウォーズ、ランボー
- E 君: 風の谷のナウシカ、ウィスパーズ

今、F 君が「王様と私」が好きだったとすると、F 君はほかに何が好きであろうか。「王様と私」が好きな A 君と B 君は「スターウォーズ」が好きなので、F 君も「スターウォーズ」が好きであると判断できる。しかし、一方で「スターウォーズ」はほとんどの人が好きな映画であるので、推薦はできるがあまり有効な情報は言えない。そこで、「王様と私」が好きな B 君が好きな「風の谷のナウシカ」を推薦することもできる。さらに、もう 1 段推論を進めると E 君の情報から、「風の谷のナウシカ」が好きであれば「ウィスパーズ」も好きであると推測できる。このような手法を **collaborative filtering** と言う。

しかし、**collaborative filtering** にも問題がある。例えば、新規にコンテンツが登録されても、新規コンテンツがユーザに推薦されにくいといった問題がある。さらに、ユーザが興味を持つコンテンツであったとしても、実際に、そのコンテンツにアクセスする機会がなければ推薦できない<sup>1</sup>。このような問題を解決する手法として、**content-based filtering** と **collaborative filtering** を組み合わせたハイブリッドな手法が提案されている[11]。文献[11]では、コンテンツ提供者によって静的に定義されたカテゴリに加えて、動的に類似コンテンツ集合からなるカテゴリを作成し、動的カテゴリを導出している。これによって、ユーザの興味や関心を動的に反映したカテゴリ作成が可能となる。

### 3.11.3 ユーザごとの情報検索サポート

本節では、従来提案されているユーザごとの情報検索サポート手法について概説する。

#### 3.11.3.1 Query Clustering

3.11.2 節で紹介した情報フィルタリングの手法は、情報検索においては、**query clustering**[7][8]として応用されている。

クラスタリングにおいては、2 つの検索語（各検索語は複数の語で構成されているとす

---

<sup>1</sup> 映画の例では、観れば好きな映画と判断できても、実際に観なければ推薦されないことを意味する。

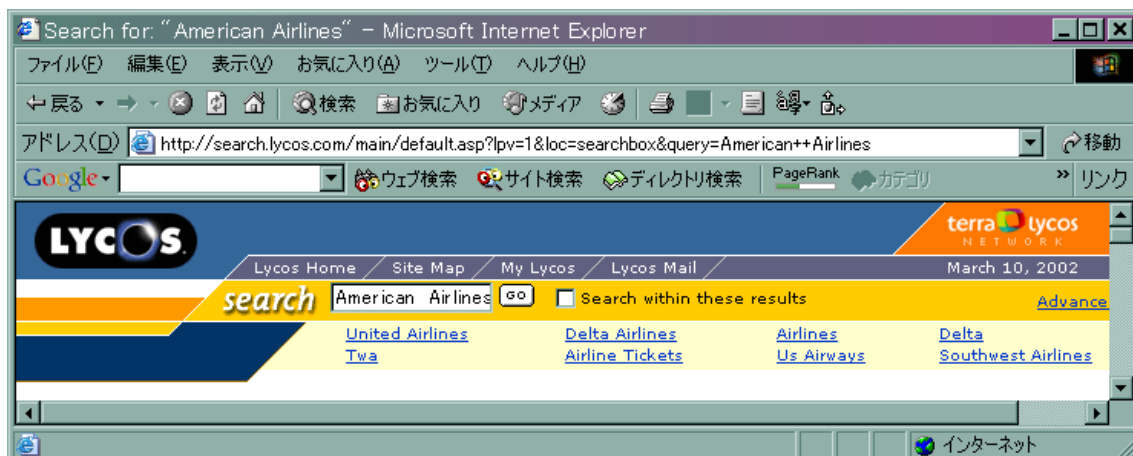


図 3.11-1 Lycos での検索語クラスタリング

る)があった時、これらの検索語の間に類似性<sup>1</sup>があれば、同一もしくは類似した情報を検索していると判断する。しかし、情報検索システムに入力される検索語は一般的に 1~2 語であり、検索語間の類似性を検索語のみから判断することは難しい。そこで、情報検索分野で使われているクラスタリング手法が用いられる。情報検索においては、伝統的に「類似した文書は同じ検索語により検索される傾向がある」ことがわかっているため、この関係を逆に利用し、「検索履歴を用いて、最終的に同一の検索結果にたどりつく検索語群」を同一のクラスタに分類している。

そして、ユーザが分類済のクラスタに属する検索語を入力した場合、入力された検索語に対応づけられた検索結果を返すことにより、ユーザがあいまいな検索語を入力した場合でも的確な検索結果を返すことを可能にする。ただし、必ずしも 1つの検索語が 1つのクラスタに属するとは限らないため、入力された検索語が複数のクラスタに属する場合は、検索結果の上位に異なるクラスタの検索結果を表示するという手法がとられる。この場合、ユーザが、「similar」<sup>2</sup>などのボタンをクリックすることによって、情報検索システムは、クラスタを特定することができ、最終的にユーザが望んでいる検索結果を返すことが可能になる。

一方、検索語のみのクラスタリングも可能であり、Lycos[12]では、検索語と同一のクラスタに属する語をユーザに提示する仕組みを持つ。例えば、American Airlines で検索すると United Airlines や Delta などを表示する (図 3.11-1)。さらに、文献[13]では、同一のユーザの連続する検索履歴をもとに、同一ユーザが入力した検索語群に対してクラスタ

<sup>1</sup> 同一のキーワードを含んでいるなど。

<sup>2</sup> 検索結果の各 Web ページに付加的につけられたボタンで、「similar」をクリックすると該当の Web ページに類似する検索結果のみが表示される。



リングを行う仕組みを提案している。例えば、「ホームセンター」で検索した場合に、同一のクラスタに属する「ディスカウントショップ」での検索を並行して行う仕組みを提案している。

### 3.11.3.2 Query Transformation

Query Transformation は、検索語とインデックス化された Web ページとの間の語句のミスマッチングをなくすための変換手法であり、検索語として自由文が入力された際に効果を発揮する。例えば、「ハードディスクって何?」という検索語が入力された場合、形態素解析<sup>1</sup>によって検索語を分割すると「ハードディスク」「って」「何」「?」となる。一般的に、「って」「?」は stop word<sup>2</sup>として無視されるため、最終的には、「ハードディスク」AND「何」で検索が実行される。ところが、Web ページとして探し出したいページに「何」という語句が存在するかどうかは疑問である。例えば、このような例の場合、「ハードディスクとは」で始まる文章<sup>3</sup>を持つ Web ページを検索することにより、ユーザの検索意図にあった Web ページを見つけ出すことができる。すなわち、「ハードディスクって何?」という検索語を「ハードディスクとは」に変換し検索を行うことを Query Transformation と呼ぶ。

Query Transformation に関連する研究は、TREC[14] (Text Retrieval Evaluation Conference) の Question-Answering Track において多くの研究が実施されている。TREC においては、50~250 バイトの解答文書 (英文) をあらかじめ用意された question セットにより検索する課題が出されるが、実際の Web 上で適用する場合には、さらに一般化が必要となる。例えば、AskJeeves[15] (図 3.11-2) では、これらの変換に対してあらかじめ変換テーブルを用意し、ユーザが入力した質問形式の検索語を回答の書き方に合わせて変換する。

図 3.11-2 では、「Who is the president of USA?」に対してホワイトハウスの Web ページが検索結果として出力されている。

文献[16]では、「what is ×」のような質問に対して「× is」「× refers to」という解答を含む文で現れる語句に変換する仕組みを利用して複数の情報検索システムに Query Transformation 語の検索語を送り検索させるシステムを提案している。さらに、文献[17]では、これらの変換テーブルをタグ付きの音声データから自動取得する方式を提案している。

---

<sup>1</sup> 形態素 (動詞、名詞、……) に分割すること。

<sup>2</sup> 検索語として適さない語であり、検索の際に無視される語句。

<sup>3</sup> 一般的に語句を説明する文章は「○○とは」で始まることが多いため。

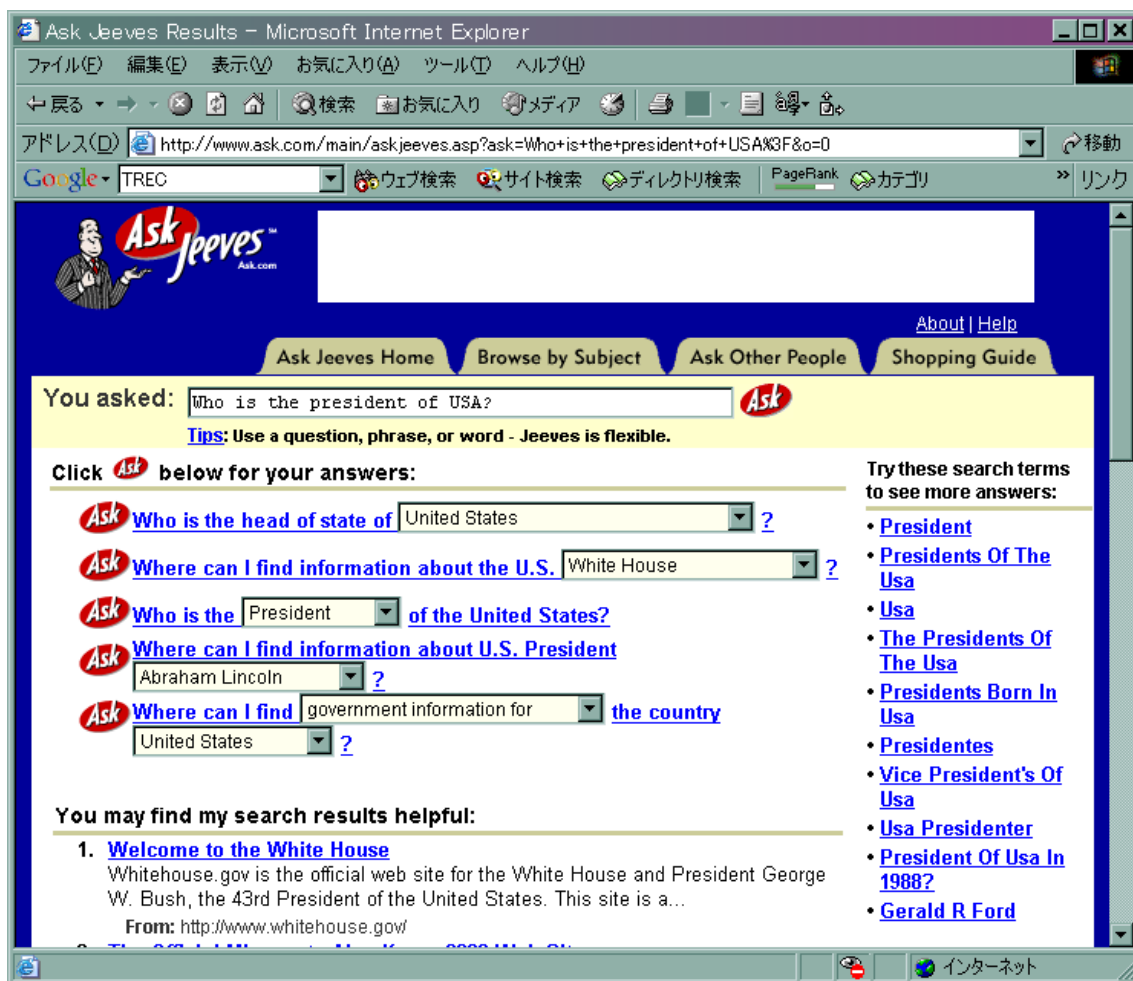


図 3.11-2 AskJeeves

### 3.11.3.3 Query Routing

複数の情報検索システムの中から、入力された検索語のトピックに適した情報検索システムを探し出すことを **query routing** と呼ぶ。近年、さまざまな専用化された情報検索システムが登場してきており、**query routing** の役割は今後ますます重要となると考えられる。

Query Routing は、その手法を 3 つに分類することができる[18]。

- (1) 手動ルーティング。
- (2) 各データベースが持つコンテンツ情報に基づいて自動ルーティングする手法。
- (3) 各データベースのコンテンツ情報を使用せずに自動ルーティングする手法。

手動ルーティングは、AllSearchEngines.com のようにあらかじめ各情報検索システム

をカテゴリ化しておき、検索時にユーザに使用する情報検索システムを選択させる手法である。各データベースのコンテンツ情報に基づいて自動ルーティングする手法は、各データベースがコンテンツとして持つ情報からキーワードを作成し、検索語とのマッチングをとることにより、入力された検索語に適した情報検索システムを選択する手法である。しかし、Webを対象とした場合、内部のデータベースにアクセスすることは困難であるため、Webを対象とした Query Routing では、3番目の手法が一般的に用いられる。

### 3.11.4 連続するクエリの出現パターンの分類とユーザの検索意図の推定

本節では、従来の検索語の内容に依存した情報検索サポートではなく、検索語の出現パターンのみからユーザの検索意図の一部が解析可能であることを示すために行った実験[19]について紹介する。実験では、同一ユーザの連続する検索語の時系列をモデル化し、各モデルに対して検索意図の推定を行った。

#### 3.11.4.1 実験手法

実験にあたっては Microsoft Accessibility[20]および Microsoft Product Support Service のサポート技術情報検索[21]に入力された検索履歴を使用した。使用する検索履歴は個人情報を含んでおらず、検索語と検索が行われたタイムスタンプのみである。このため、まず、同一ユーザの連続する検索を同定する必要がある。以下に解析手順を示す。

##### (1) 同一ユーザによる連続する検索の同定

入力された各々の検索語は独立の検索としてログ上に記録されているため、まず、同一ユーザによる一連の検索を同定する必要がある。本報告では、10分以内の連続する検索であり、かつ、複数の検索語を用いて検索している場合は1番目に入力されている検索語が同一である場合、同一ユーザによる検索であると判断した。したがって、同一ユーザによる検索であっても、1番目の検索語を変更した場合は、本実験では同一ユーザとして判断できていない<sup>1</sup>。

##### (2) 検索のパターン分類

(1) によって得られた同一ユーザによる連続する検索語入力について、検索語の増減、変化を分類する。各分類（パターン）の発生頻度を求める。

---

<sup>1</sup> 同一ユーザによる連続する検索を同定するためには、Cookie を用いるのが一般的である。しかし、Cookie の利用は個人情報取得に関係するため、今回の実験ではタイムスタンプと検索語のみを用いて同一ユーザの同定を行った。同定においては、検索語の最初の語が等しいことを前提としたが、これは、「一般的に最初の語は変更しない場合が多い」という経験則に基づいたものである。

## 3.11.4.2 解析結果

表 3.11-1 同一ユーザによる連続検索回数

| 連続検索回数 | 人数    | 割合      |
|--------|-------|---------|
| 2      | 8698  | 78.52%  |
| 3      | 1770  | 15.98%  |
| 4      | 399   | 3.60%   |
| 5      | 129   | 1.16%   |
| 6      | 49    | 0.44%   |
| 7      | 21    | 0.19%   |
| 8      | 3     | 0.03%   |
| 9      | 4     | 0.04%   |
| 10 回以上 | 4     | 0.04%   |
| 合計     | 11077 | 100.00% |

表 3.11-2 検索パターンの分類と割合

| No. | パターン     | 件数    | 割合     |
|-----|----------|-------|--------|
| 1   | A-AB     | 3939  | 35.6%  |
| 2   | AB-AC    | 2030  | 18.3%  |
| 3   | AB-A     | 2729  | 24.6%  |
| 4   | A-AB-AC  | 335   | 3.0%   |
| 5   | A-AB-A   | 95    | 0.9%   |
| 6   | A-ABC-AB | 521   | 4.7%   |
| 7   | AB-A-AC  | 151   | 1.4%   |
| 8   | AB-A-AB  | 15    | 0.1%   |
| 9   | AB-AC-AB | 266   | 2.4%   |
|     | その他      | 996   | 9.0%   |
| 合計  |          | 11077 | 100.0% |

A、B、Cはそれぞれ異なる検索語を示す。  
 A-ABは最初の検索で「A」を、2回目の検索で「A and B」を検索語として入力したことを示す。

複数回（2回以上）検索しているユーザの検索回数を表 3.11-1 に示す。この表より 4 回以上の検索を行っているユーザは全体の 5%であり、残り 95%の検索は 3 回以内の検索で完了していることがわかる。以上から、連続する 3 回までの検索について頻出する検索パターンを分類した。検索パターンの分類とその割合を表 3.11-2 に示す。各検索パターンの説明を表 3.11-3 に、各検索パターンにおける再検索の推測理由を表 3.11-4 に示す。

表 3.11-3 検索パターンの説明

| No. | パターン     | 説明                                                   | 例                                                                                              |
|-----|----------|------------------------------------------------------|------------------------------------------------------------------------------------------------|
| 1   | A-AB     | 最初の検索語から検索語の数を増やす場合。                                 | 1 回目: "excel"<br>2 回目: "excel"、"ダウンロード"                                                        |
| 2   | AB-AC    | 先頭の検索語は変えずに後ろの検索語を他の検索語に変える場合。                       | 1 回目: "IP アドレス"、"プリンタ"<br>2 回目: "IP アドレス"、"プリントサーバー"                                           |
| 3   | AB-A     | 1 回目に複数個の検索語で検索を行ったが、2 回目で検索語を減らす場合                  | 1 回目: "windows2000"、"ファイルキャッシュ"<br>2 回目: "windows2000"                                         |
| 4   | A-AB-AC  | 2 回目の検索で検索語を増やし、3 回目の検索語で後ろの検索語を変えている場合。             | 1 回目: "フォント"<br>2 回目: "フォント"、"追加"<br>3 回目: "フォント"、"ダウンロード"                                     |
| 5   | A-AB-A   | 2 回目の検索で検索語の数を増やし、3 回目の検索で再び検索語を減らし元に戻す場合。           | 1 回目: "キーボード"<br>2 回目: "キーボード"、"OSK"<br>3 回目: "キーボード"                                          |
| 6   | A-ABC-AB | 2 回目の検索で 2 個増やしたが、3 回目でも再び 1 個減らす場合。                 | 1 回目: "ユーザインターフェイス"<br>2 回目: "ユーザインターフェイス"、"ガイドライン"、"Windows"<br>3 回目: "ユーザインターフェイス"、"Windows" |
| 7   | AB-A-AC  | 1 回目の検索で複数の検索語を入れて検索し、2 回目で検索語を増やし、3 回目で再び検索語を増やす場合。 | 1 回目: "ProxyServer"、"フィルタ"<br>2 回目: "ProxyServer"<br>3 回目: "ProxyServer"、"設定"                  |
| 8   | AB-A-AB  | 検索語を減らしたあと、再び同じ検索語を増やす場合                             | 1 回目: "ユーザー"、"切り替え"<br>2 回目: "ユーザー"<br>3 回目: "ユーザー"、"切り替え"                                     |
| 9   | AB-AC-AB | 検索語を変えた後、再び同じ検索語に戻す場合                                | 1 回目: "ユーザー"、"切り替え"<br>2 回目: "ユーザー"、"切替"<br>3 回目: "ユーザー"、"切り替え"                                |

A、B、C はそれぞれ異なる検索語を示す。

### 3 研究開発の新しい展開と内外の動向

表 3.11-2 より、2 回の検索（パターン 1～3）が約 80%を占め、3 回までの検索を含めると全検索の 91%を占めることがわかる。なお、表 3.11-1 では 3 回までの検索回数が全体の約 95%を占めているのに対し、表 3.11-2 では 91%となっているのは、表 3.11-2 のパターン分類に分類されない 3 回以内の検索が含まれているからである。

表 3.11-4 各検索パターンにおける再検索の推測理由

| No. | パターン   | 検索結果が多すぎた場合 | 検索結果が少なすぎた場合（0 の場合も含む） | 検索結果を見たが、欲しい情報がなかった場合 | 実際にページを閲覧した結果、欲しい情報がなかった場合 | 1 回目の検索結果の方がふさわしいと判断した場合 | 表記のゆれを訂正 |
|-----|--------|-------------|------------------------|-----------------------|----------------------------|--------------------------|----------|
| 1   | A-AB   | ○           |                        | ○                     | ○                          |                          |          |
| 2   | AB-AC  | ○           | ○                      | ○                     | ○                          |                          | ○        |
| 3   | AB-A   |             | ○                      | ○                     | ○                          |                          |          |
| 4-a | A-AB   | ○           |                        | ○                     | ○                          |                          |          |
| 4-b | AB-AC  | ○           | ○                      | ○                     | ○                          |                          | ○        |
| 5-a | A-AB   | ○           |                        | ○                     | ○                          |                          |          |
| 5-b | AB-A   |             |                        |                       |                            | ○                        |          |
| 6-a | A-ABC  | ○           |                        | ○                     | ○                          |                          |          |
| 6-b | ABC-AB |             | ○                      |                       |                            |                          |          |
| 7-a | AB-A   |             | ○                      | ○                     | ○                          |                          |          |
| 7-b | A-AC   | ○           |                        | ○                     | ○                          |                          |          |
| 8-a | AB-A   |             | ○                      | ○                     | ○                          |                          |          |
| 8-b | A-AB   |             |                        |                       |                            | ○                        |          |
| 9-a | AB-AC  | ○           | ○                      | ○                     | ○                          |                          | ○        |
| 9-b | AC-AB  |             |                        |                       |                            | ○                        |          |

4-a とはパターン 4 の 1 回目、4-b とはパターン 4 の 2 回目をあらわす。

### 3.11.4.3 検索パターンに基づいた情報検索サポート

表 3.11-5 各パターン別の情報検索サポート

| No. | パターン    | 検索語入力 of 意図 (推定)                           | 情報検索システムでの検索サポート (例)                           |
|-----|---------|--------------------------------------------|------------------------------------------------|
| 1   | A→AB    | (1) A の検索において、検索結果上位に求める結果が存在しない。          | 特になし。                                          |
| 2   | AB → AC | (2) B での絞り込みに失敗。                           | (2') AC の検索において、B を含まないものを上位にランキング。            |
| 3   | AB→A    | (3) AB では検索結果が絞り込まれすぎたが、B はまったく関係ない語ではない。  | (3') A の検索において、B および B に関連する語を含むページを上位にランキング。  |
| 4-a | A→AB    | (1) と同様。                                   | 特になし。                                          |
| 4-b | AB → AC | (2) と同様。                                   | (2') と同様。                                      |
| 5-a | A→AB    | (1) と同様。                                   | 特になし。                                          |
| 5-b | AB→A    | (3) と同様。                                   | (3') と同様。                                      |
| 6-a | A → ABC | (1) と同様。                                   | 特になし。                                          |
| 6-b | ABC→AB  | (4) ABC では検索結果が絞り込まれすぎたが、C はまったく関係ない語ではない。 | (4') AB の検索において、C および C に関連する語を含むページを上位にランキング。 |
| 7-a | AB→A    | (3) と同様。                                   | (3') と同様。                                      |
| 7-b | A→AC    | (5) AB では絞り込まれすぎ、A では検索結果上位に求める結果が存在しない。   | (5') AC の検索において、B および B に関連する後を上位にランキング。       |
| 8-a | AB→A    | (3) と同様。                                   | (3') と同様。                                      |
| 8-b | A→AB    | (6) A よりも AB の検索結果の方がよい。しかし、AB では絞り込まれすぎる。 | (6') AB の検索において B を B の関連語に拡張して検索。             |
| 9-a | AB → AC | (2) と同様。                                   | (2') と同様。                                      |
| 9-b | AC → AB | (7) AC よりも AB の検索結果の方がよい。                  | (7') AB の検索において、C を含む語を上位にランキング。               |

各検索パターンに対する再検索理由（表 3.11-4）をもとに、各検索パターンに対して表 3.11-5 に示すような情報検索サポートが可能であると考えられる。現時点では、これらの情報を用いた実験は未実施であるが、3.11.3 で示したような既存の手法とあわせて用いることで個人の検索意図にあった検索サポートが期待できると考えられる。

#### 3.11.5 おわりに

本節では、情報検索において、ユーザの検索意図を検索結果に反映させるための仕組みについて、従来の研究を概説するとともに、新しい手法として、同一ユーザが検索を行う際の連続する検索語の出現パターンから、ユーザの検索意図を発見する手法を紹介した。世界中のコンピュータがインターネットにより接続される現在、ネットワークにより接続された各種情報は、Knowledge Grid と呼ぶべきものであり、膨大なインターネット上の情報の中から必要な情報を効率よく取り出し活用する仕組みが今後ますます重要になっていくものと考えられる。

#### 参考文献

- [1] Yahoo!: <http://www.yahoo.com/>
- [2] Google: <http://www.google.com/>
- [3] L. Page, S. Brin, R. Motwani, and T. Winograd: The PageRank Citation Ranking: Bringing Order to the Web, Stanford Digital Library Technologies, Working Paper SIDL-WP-1999-0120 (1998).
- [4] M. Henzinger: Link Analysis in Web Information Retrieval, IEEE Bull. of the Tech. Committee on Data Engineering, Vol. 23, No. 3, pp. 3-8 (2000).
- [5] T. H. Haveliwala: Efficient Computation of PageRank, Stanford Digital Library Technologies, Technical Rep. 1999-31 (1999).
- [6] [6] U. Manber, A. Patel and J. Robison: Experience with Personalization on Yahoo!, Communications of the ACM, Vol. 43, No. 8, pp. 35-39 (2000).
- [7] [7] D. Beeferman and A. Berger: Agglomerative Clustering of a Search Engine Query Log, KDD2000, pp. 407-416 (2000).
- [8] J. R. Wen, J. Y. Nie, and H. J. Zhang: Clustering User Queries of a Search Engine, Proc. of 10th World Wide Web Conference, Hong Kong (2001).
- [9] U. Shardanand and P. Maes: Social Information Filtering: Algorithms for Automating 'word of mouth.', Proc. of the ACM CHI Conf. (CHI95) (1995).



- [10] L. H. Ungar and D .P. Foster: Clustering Methods for Collaborative Filtering, Proc. of the Workshop on Recommendation System, AAAI Press, Menlo Park California (1998).
- [11] M. Okamoto, J. Kamahara, S. Shimojo, and H. Miyahara: Automatic Production of Personalized Contents with Dynamic Scenario, Proc. of 2001 IEEE Pacific Rim Conf. on Communications, Computers and Signal Processing (PACRIM01), Vol. 1, pp. 91-94 (2001).
- [12] 大浦, 喜連川: Web アクセスログのクラスタリングによる問い合わせ拡張に関する検討: iタウンページ上での実験, 第13回データ工学研究会 (DEWS2002) (2002).
- [13] Lycos: <http://www.lycos.com/>
- [14] TREC: <http://trec.nist.gov/>
- [15] AskJeeves: <http://www.ask.com/>
- [16] S. Lawrence and C. L. Giles: Context and Page Analysis for Improved Web Search, IEEE Internet Computing, Vol. 2, No. 4, pp. 38-46 (1998).
- [17] E. Agichtein, S. Lawrence and L. Gravano: Learning Search Engine Specific Query Transformations for Question Answering, Proc. of 10th World Wide Conference, Hong Kong (2002).
- [18] A. Sugiura and O. Etzioni: Query Routing for Web Search Engines: Architecture and Experiments, Proc. of 8th World Wide Web Conference (1999).
- [19] 鈴木俊輔, 山名早人: 時間間隔を用いた検索履歴のモデル化, 情報処理学会デジタルドキュメント研究会 (Mar. 2002).
- [20] Microsoft Accessibility: <http://www.microsoft.com/japan/enable/>
- [21] Microsoft Product Support Service: <http://support.microsoft.com/>

## 3.12 Web マイニングとその将来像

報告者: 津田宏講師

### 3.12.1 はじめに

Web は、1989 年に当時 CERN (Conseil European pour la Recherche Nucleaire、欧州合同素粒子原子核研究機構) に勤務していた Tim Berners-Lee が情報共有のために発明し、1990 年から実験的にサーバが稼働していたとされる。しかし、その後の伸びが目覚ましいのは周知の通りである。ページ数も、次のように天文学的な数字になりつつある (ただし、Web は総ページ数をどこかで管理しているわけではなく、動的なページをどう扱うかで数値も変わり、これらは推計方法もまちまちであるため、一列に並べるのに問題はある。ただし、このようなオーダーのページ数があるというのは、ほぼ正しいだろうと思われる)。

- 1998 年 4 月: 3.2 億ページ (NEC 北米研)。
- 1999 年 7 月: 8 億ページで、15TB のデータと 1.8 億の画像を含む (NEC 北米研)。
- 2000 年 1 月: 10 億ページを突破 (Inktomi 社ほか)。
- 2000 年 7 月: 21 億ページで、2001 年にはその倍。

これだけの量の文書があり、なおかつ Google をはじめとする優秀な検索エンジンのおかげで、さすがに体系的な知識の入手という点では書籍に劣るものの、特定のトピックやちょっとした情報であれば、最も簡単に素早く入手できる環境ができています。

こうした膨大な Web を単に調べるだけでなく、その中から必要な知識をうまく取り出し、研究やビジネスに活用していきこうという流れが「Web マイニング」(Web からの知識発掘) である。データマイニング (大量生データからの知識発掘) やテキストマイニング (大量文書からの知識発掘) と違い、Web には文書だけでなく、リンクやログといった色々なデータがある。そのため、Web マイニングに有用性はあるものの、固有技術はあまりなく、従来技術を Web に活用していったという研究が多く、全体としての体系化は遅れている。

以下、3.12.2 節では、Web マイニングの概要を実例も含めて紹介する。3.12.3 節では、Web マイニング技術の限界と次世代 Web における展開を述べる。

表 3.12-1 Web マイニング技術の概要

|               | 対象                                                                         | 抽出する知識例                                                                     | 技術                                                                   | コミュニティ                                          | 応用                    |
|---------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------|----------------------------------------------------------------------|-------------------------------------------------|-----------------------|
| データ<br>マイニング  | 生データ<br>(POS データ<br>など)、<br>VLDB                                           | アイテム間の<br>相関ルール、<br>予測                                                      | 相関ルール抽出、<br>クラスタリング、<br>分類、視覚化                                       | DB                                              | 意思<br>決定、<br>効率<br>向上 |
| テキスト<br>マイニング | テキスト (mail、<br>新聞、論文、<br>コールセンター<br>ログなど)                                  | 単語間関係、<br>文書間関係<br>(分類)                                                     | テキストからの<br>データ化<br>(形態素解析、<br>情報抽出)                                  | DB、<br>NL、<br>IR、<br>AI                         | CRM、<br>KM            |
| Web<br>マイニング  | Web<br>(HTML:<br>テキスト +<br>URL +<br>タグ +<br>ハイパーリンク +<br>レイアウト、<br>アクセスログ) | ページメタデータ<br>(人気度、分類<br>など)、<br>ページ集合<br>(コミュニティ)<br>発見、<br>DB 知識<br>(XML 化) | タグ構造からの<br>情報抽出<br>(wrapper)、<br>リンク解析、<br>効率的収集<br>(アクティブ<br>マイニング) | DB、<br>NL、<br>IR、<br>AI、<br>KA、<br>Web 系、<br>…… | CRM、<br>KM、<br>……     |

AI: artificial intelligence (人工知能)

CRM: customer relationship management (顧客関係管理)

DB: database (データベース)

IR: information retrieval (情報検索)

KA: knowledge acquisition (知識獲得)

KM: knowledge management (知識管理)

NL: natural language processing (自然言語処理)

### 3.12.2 Web マイニングとその概要

#### 3.12.2.1 データマイニングから Web マイニングへ

表 3.12-1 に、データマイニング技術、テキストマイニング技術、および Web マイニング技術の関係の概要をまとめた。「対象」の列は、何から知識を取り出すかを示す。「抽出する知識例」の列は、対象から取り出される知識を示す。「技術」の列は、そこで必要となる技術の例を示す。「コミュニティ」の列は、その技術が発表される領域を示す。Web 系には、WWW Conference などがある。「応用」の列は、ビジネス上の用途を示す。

こうして見るとわかるように、Web マイニングは、とりわけ多くのコミュニティにまたがっている。これは、対象とする Web 情報が、文書 (テキスト)、データ (ログ、リンク) など、さまざまな側面を持っていることに由来する。

#### 3.12.2.2 Kosala による分類

Kosala らは、Web マイニングを、次のように処理対象データの側面から分類している[9]。

##### (1) Web Content Mining

Web ページの文書としてのコンテンツから知識を取り出す。

##### (1a) IR (information retrieval、情報検索) 的観点

Web ページを構造化されていないテキストと見る。

IR や自然言語処理技術を利用。

応用としては、ページ分類やテキスト中のパターン情報抽出など。

##### (1b) DB (database、データベース) 的観点

Web ページを (タグにより) 半構造化されたデータとして見る。

タグを利用した情報抽出 (wrapper) などの技術を利用。

応用としては、ページの繰り返しパターンの発見など。

##### (2) Web Structure Mining

ページ間のリンク関係から情報を取り出す。

応用としては、分類 (コミュニティ発見) や、人気度の獲得などがある。

##### (3) Web Usage Mining

Web の利用ログの分析である。サーバのアクセスログや、検索エンジンのログなどがある。

応用としては、Web サイトの構築/改築支援、マーケティングなどがある。ただし、ログ分析単体では、どちらかというとなるデータマイニングの範ちゅうに入ると思われる。

(1) と (2) の実例については後述する。

### 3.12.2.3 山西による事例紹介

山西は、Web マイニングを、表 3.12-2 のように事例の側面から分類している[17]。

「意見分析」は、Content Mining の例であり、Web ページや掲示板の情報から、利用者が製品についてどのような意見（好意的とか、けなしているとかなど）を持っているかを弁別するもので、Web を使ったマーケティング支援の例である。

「コミュニティ発見」は、Structure Mining の例であり、リンク関係を利用して類似したページを集めるというものである。

「最適広告配信」以下の4つは、いずれも Usage Mining の例である。さまざまな技術が Web のログに適用されている。

表 3.12-2 山西による事例紹介

| Function     | Input          | Output            | Model + Algorithm              |
|--------------|----------------|-------------------|--------------------------------|
| 意見分析         | コンテンツ          | 特徴語リスト            | 情報抽出、テキスト分類、確率的コンプレキシティ        |
| Web コミュニティ発見 | リンク            | コミュニティページリスト      | グラフ理論、平均場理論、ベイジアンネット           |
| 最適広告配信       | ログ<br>(クリック履歴) | バナー広告             | ニューラルネット、情報量基準、Hitchcock 型輸送問題 |
| リコメンデーション    | ログ、クッキー、コンテンツ  | 推薦コンテンツ           | 相関係数法、逐次型 2 項関係学習法             |
| アクセスパタン分析    | アクセスログ         | アクセスパタン<br>(クラスタ) | マルコフモデル、有限混合モデル                |
| 異常ログ検出       | アクセスログ         | 異常ログリスト           | 統計的外れ値検出、オンライン忘却型アルゴリズム        |

#### 3.12.2.4 Webマイニングのプロセス

Kosalaによると、Webマイニングは以下のプロセスから成る[9]。

(1) Resource Finding

対象とするリソースの同定、収集。

(2) Information Selection, pre-processing

データのクレンジング（ゴミデータの除去）、形態素解析などのテキスト処理、情報抽出などの特定パターン取り出しなど、対象を処理できる情報に変換したり、データをきれいにしたりするプロセスである。

(3) Generalization

マイニングプロセス。ルールやパターンの取り出し、分類などマイニングの本体である。

(4) Analysis

得られた結果を人手でチェックする。

KDD (knowledge discovery in database、データベースからの知識発見)、いわゆるデータマイニングでも、同様のプロセスが必要とされている[Usama 1996]。ただし、データマイニングやテキストマイニングではあまり重要ではなかったが、Webマイニングで重要となるプロセスが(1)のResource Findingである。というのも、3.12.1節で述べたように、Webは膨大で常に拡大しているため、そのすべてをWebロボット（自動収集プログラム、Webクローラとも言う）で集めるのは膨大な時間がかかり、また集めたときには対象となるWebは変化してしまっている。このように、対象が常に変化する膨大なデータから、当面必要となる情報をいかにうまく取り出すかということが、Webマイニングでは必要になる（このような性質のデータからの知識発掘は、「アクティブマイニング」と呼ばれはじめている[11]）。

#### 3.12.2.5 Resource Finding

Webロボットは1995年くらいから行われてきており、Webの収集においては、深さ限定で幅優先探索が有効とされている。ただし、そのような固定した収集方略では、まんべんなく大量にとる場合は良いが、特定のページを収集するということとはできない。そのため、目的に応じて必要なページを選択的に収集する技術が提案されている。Webの拡大と共に、そのように効率的に収集する手法への期待は大きい。

東芝で（現在は独立した会社で）サービスが行われている、フレッシュアイという検索エンジンサービスでは、ページの更新頻度を学習し、良く更新されるページを優先的に収集することが特徴である。

Focused Crawling[2]は、ハイパーリンク関係にある文書は内容的にも近いので、文書

のコンテンツ類似度から見て、離れそうになったら探索を打ち切るという考え方である。

**Efficient Crawling**[3]は、リンク先を表している表現（HTMLであればアンカータグの内容）から、次にそのリンク先の文書を収集するかどうか決定するという考え方を使っている。リンク関係を利用し、次に可能性がありそうなものを予想することも行っている。つまり、収集済みページからの

- ・ 被リンク数が多い
- ・ ランク重みが高い
- ・ 共参照関係（一緒にリンクされることが多い）にある

ものは関連度が高いというヒューリスティックスを利用し、次に集めるのが有効そうなページを判定する。例えば、被リンク数の多い文書は人気度が高いと考えられる。収集済み文書群から被リンク数の多い順に文書を収集することで、重要度の高い文書を優先して収集できる。

津田らは、特定ジャンルの少数のページから、リンク関係および共参照関係にあるページをブートストラッピングにより順次集めることで、意味的に近いページを効率良く集める手法を提案している[14]。

ただし、検索や分類と違い、このような収集系に関しては、決まった正解セットに対して精度を競うというレベルにまだ成熟した学問領域ではなく、個々の論文で独自に効果を測定している段階である。各論文では良いとは言っているが、それらは基本的には追試不可能である。単純な幅優先の方がかえって良かったという論文も出ており[12]、公平な測定手法が求められている。

### 3.12.2.6 Web Content Mining

**Web Content Mining** の例として、**Web** ページからの情報抽出の実例を紹介しよう。**Web** に公開されている分散したカタログ情報から、製品名、価格などのレコード情報を自動で取り出し、横断検索を行うというアイデアは**Web** の初期のころから提案されていた。例えば、**ShopBot**[5]などがある。また、**Web** を知識ベースとして扱おうという **WebKB**[4]も、同様の試みである。

こうした **Web** からの情報抽出では、従来のテキストからの情報抽出のようにテキスト中の文字パターンからの取り出しも可能である。坂元らは、**Web** からの近接語相関パターン的高速発見の手法を述べている[13]。また、**Web** ならではの技術として、HTML のタグ情報を利用した、**Wrapper Induction**[10]が有名である。HTML ソースと、取り出したいレコードの対データ（正解セット）を与えると、値を取り出すようなタグのパターンを学習する機械学習手法を提案している。表形式など、個々の値が<b>タグなどでくくられているようなケースには、有効である。

#### 3.12.2.7 Web Structure Mining

Web のハイパーリンクを解析することで、ページの人気度を算出したり、特定コミュニティを取り出したりする研究は、97 年くらいから数多く行われている。

Brin らは、検索エンジン Google の元となった PageRank アルゴリズムを提案している [1]。これは、ページを見ている人々の推移確率をモデル化した人気度算出手法で、

- ・ 被リンクが多いのは人気
- ・ 人気ページからリンクされているページは人気

という再帰的な計算を行い、不動点としてページの人気度（重み）を決定している。

Kleinberg は、HITS (Hyperlink-Induced Topic Search) アルゴリズムを提案している [8]。これは、特定ジャンルの、ハブページ（優良リンク集）とオーソリティページ（優良コンテンツ）を取り出すものである。オーソリティページは、有用な情報を多く含み、ハブページから良くリンクされる。ハブページは、オーソリティページを良くリンクする。このアルゴリズムでは、検索エンジンの結果からリンク、被リンクを 1 または 2 レベルたどった探索空間を設定し、繰り返し計算による不動点を求める。

Gibson らは、HITS アルゴリズムを応用して、コミュニティを、オーソリティページを中心にハブページが取り囲む構造として、取り出している [7]。

津田らは、Yahoo のような Web ディレクトリを自動構築する際の、リンク解析技術について述べている [14][15]。Web ディレクトリでは、新鮮で、人気度が高く、情報の入り口となるページが望ましい。[14]では、リンク解析による人気度付与と、ページタイプ（メニュー、リンク集、コンテンツ）付与を行っている。

津田らは、リンク解析による人気度付与の時系列解析により、最近人気が出ているページなど、Web の動きを視覚化する研究を行っている [15]。図 3.12-1 は、toto（サッカーくじ）のホームページの人気度（重みでソートした順位）の 2000 年 2 月から 12 月までの時系列的動きをプロットしたものである。3 月の開始時点では、マスコミで取り上げられたこともあり、(ニュースサイトなどから)同サイトへのリンクも増え人気度が上昇している。しかし、夏くらいから頭打ちになり、段々人気度が降下していることがわかる。これは、人気度を順位でプロットしており、Web の拡大速度についていけなくなると、徐々に順位は落ちていくためである。





## <http://www.toto-dream.com/> の分析結果

- [タイトル] サッカーくじtoto official web site
- [説明] (なし)
- [キーワード] (なし)
- [ディレクトリ登録済] ([団体,不明](#)) ([toto,不明](#))
- [収集日,最終更新日] 20010925, 20010924
- [ジャンル] 団体, toto
- [地域] 不明,
- [順位] 10096
- [一次格納ディレクトリ] 20010925/219/1944#1 **cache**
- [内部ID] 35509
- [ランク重み] 5962, [被参照数] 121(うちリモート 121)
- [ランク変化度大] 傾き42, 切片4755, 変化度8

### 1. 人気度ランキングの変動履歴

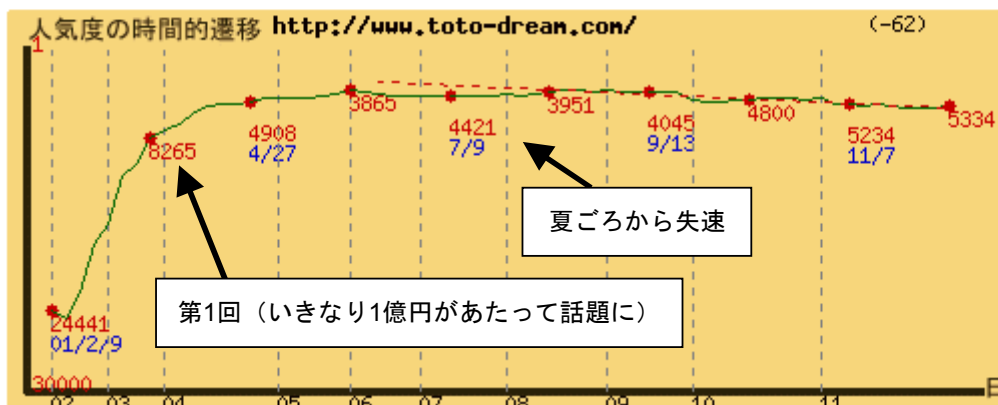


図 3.12-1 toto (サッカーくじ) ホームページの人気度推移

図 3.12-2 は、2000 年 10 月くらいから人気度が伸びている、厚生労働省の狂牛病（牛海綿状脳症、BSE: bovine spongiform encephalopathy）関係ページの人気度推移の例である。世間で問題になったのとほぼ時期を同じくして、同サイトへのリンクが増えていることがわかる。



<http://www.mhlw.go.jp/kinkyu/bse.html>

- [タイトル]「牛海綿状脳症(BSE)関係」ホームページ
- [説明] (なし)
- [キーワード] (なし)
- [ディレクトリ登録済] ([厚生労働省](#)) ( )
- [収集日,最終更新日] 20020116,
- [ジャンル] 厚生労働省,
- [地域] ,,
- [順位] 21348
- [一次格納ディレクトリ] 20020112/185/18545#7 **cache**
- [内部ID] 128304472
- [ランク重み] 5594 , [参照数] 53(うちリモート 7), [被参照数] 57(うちリモート 57)

## 1. 人気度ランキングの変動履歴

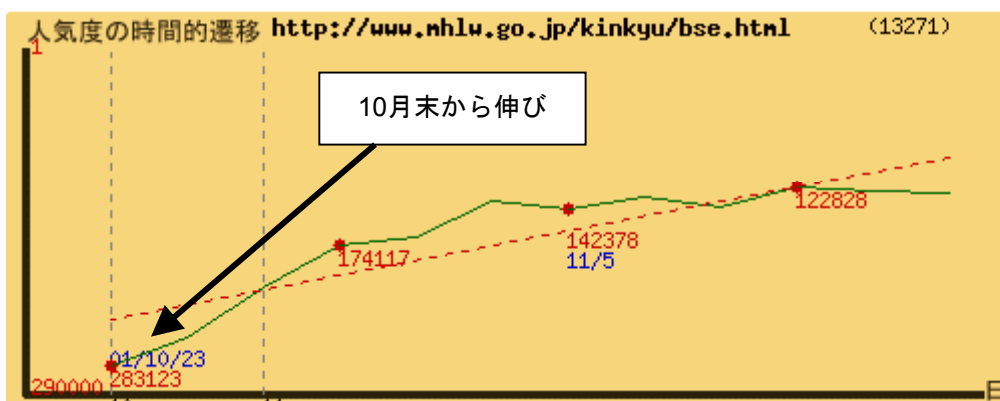


図 3.12-2 厚生労働省狂牛病関係ページの人気度推移

### 3.12.3 おわりに～次世代 Web に向けて

#### 3.12.3.1 Web マイニングの限界

3.12.1 節でも述べたように、このまま Web が拡大していくことで、人類は有史以来最も充実した知識の宝庫を持つ可能性があるが、それには次の問題点がある。

- Web は雑多で整理されていない: ある程度、3.12.2 節で述べたような技術や、Yahoo や ODP (Open Directory Project) [20]といったプロジェクトにより、人手をかけて

整理するという外づけサービスはあるものの、Web 自体にはそういった整理するという機構がない。

- 情報の信頼性が保証されない: ページ内の情報が信頼できるかどうかわからない。
- 得られた情報を有効活用できない: 現在の Web は「ブラウザで人が見る」ことに特化されているため、あるページ内の情報を利用して他の処理に自動的に連携、というような処理の自動化ができない。

Web マイニングのこれまでの事例で見たように、Web マイニング自体は完ぺきな技術ではない。他の（データ、テキスト）マイニングがそうであるように、知識の素となるような情報を取り出し、最終的には人間がチェックして、本当に役立つ知識として活用される。

Web Content Mining においては、HTML における記述の自由度（HTML は基本的には文書の意味的構造はあまり制限せず、表示主体である）や、多様な著者による多様な表現があるため、情報抽出や自動分類といった技術で精度を上げるのは一般に難しい。

Web Structure Mining では、Web のリンク解析は検索においてはそこそこ有効であると考えられるが、どういう意図でリンクしているかの情報がないため、必ずしも、リンクベースで人気度のあるページが正しいことを言っているページであるという保証はない。例えば、ある病院が事故を起こして、さまざまなニュースページからのリンクが増えると、その病院のページの人気度重みは高くなるものの、それは必ずしもその病院が良い病院ということにはならない。Web Usage Mining でも同様のことが言える。多くの人が見るからと言って、それは内容が正しいことを意味するわけではない。

## 3.12.3.2 次世代 Web

こうした背景のもと、次世代 Web として、SemanticWeb[21]や Web サービス (Web Services) が提案されている。これらは、メタデータやメタ的な情報を決まった形式で記述し、機械処理による流通・交換・処理を可能とするものである。SemanticWeb と Web サービスの特徴を、表 3.12-3 にまとめる。

表 3.12-3 次世代 Web: SemanticWeb、Web サービス

|             | Web                        | 次世代 Web                                            |                                     |
|-------------|----------------------------|----------------------------------------------------|-------------------------------------|
|             |                            | SemanticWeb                                        | Web サービス                            |
| 目的          | 人が見るため                     | ソフトウェアが意味的な処理をするため                                 | Web 上のアプリケーション連携                    |
| 記述方法        | HTML、画像、動画、flash、xhtml、……  | RDF/XML<br>(メタデータ: URI で指されるリソースについてのデータ)、<br>DAML | XML (データ)、<br>DAML-s<br>(サービスメタデータ) |
| オントロジー・構造定義 | 自由 (SHOE、お酒の箆笥など、独自の試みはある) | RDF Schema/XML、<br>DAML + OIL                      | DTD                                 |
| 利用手段        | ブラウザ                       | エージェント (http?)                                     | 統合サービス (SOAP、WSDL、UDDI)             |
| 認証、セキュリティ   | 独自                         | Dsig (電子署名)                                        | Passport (Microsoft) など             |
| 主な活動        | W3C、IETF、MS、Netscape ほか多数  | W3C、DAML (米国 DARPA)、<br>OIL (欧州)                   | IBM、MS .Net、<br>Sun ONE             |

## 3.12.3.3 Web サービス

Web サービスは、現在まちまちな形式で行われている Web 上のアプリケーション (ショッピングなど) を、統合的に連携するためのモデルである[16]。以下のような規格が提案されており、多くのメーカーがコミットしている。

- UDDI (Universal Description, Discovery, and Integration) : サービスを検索するた

めのディレクトリ。

- WSDL (Web Service Description Language)：サービスの入出力の記述言語 (XML)。いわばサービスのメタデータに相当する。
- SOAP (Simple Object Access Protocol)：低レベルの手続き呼び出しのためのプロトコル。

このような共通基盤で記述された Web サービスは、複数のものを連携することが可能であり、例えば出張に合わせて鉄道、ホテル、会議の予約を連携して行うなどのワンストップサービスが可能になる。Web サービスの具体例として、Microsoft による.NET などが有名である。

### 3.12.3.4 SemanticWeb

Web サービスがアプリケーションの統一を目指したのに対して、SemanticWeb は、Web ページなどのリソースへのメタデータ付与を提案している。SemanticWeb は、Web の発明者である Tim Berners-Lee によるコンセプトであり、人間向けで雑多になってきた Web に対して、ソフトウェアが意味的に処理できる Web 世界も加えていこうということを、長期目標としている。現在の Web を補完する考えで、Web が SemanticWeb に取って代わられるものではない。基本的には、Web リソース (URL など) に統一的なメタデータを付与することで実現する。

SemanticWeb に関連する規格の階層を図 3.12-3 に示す。現在、次のような状況にある。

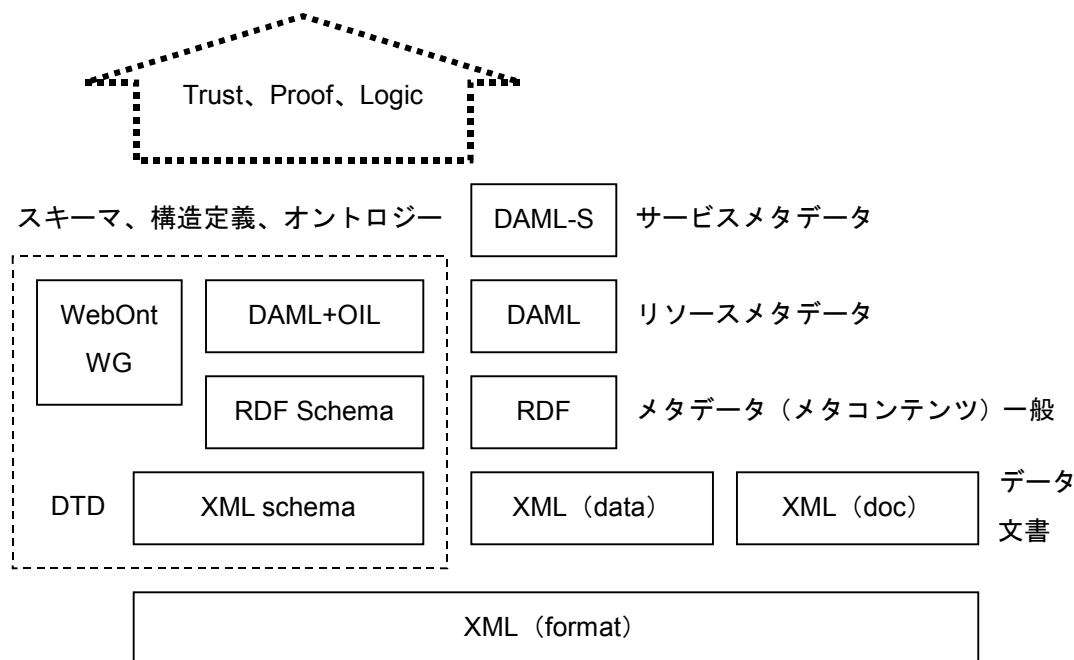


図 3.12-3 SemanticWeb の規格の階層

- データを記述する枠組みとしての RDF (Resource Description Framework) が勧告 (W3C の規格として制定される) されている。
- オントロジーについては、WebOnt WG により、オントロジー記述言語が制定されつつある。

また、SemanticWeb を利用したプロジェクトとしては、次のようなものがある。

- DAML (DARPA Agent Markup Language) [18]: 2000 年 8 月～2005 年、米国。DARPA が資金提供 (5 年で 7000 万ドル、約 8 億円)。エージェントが Web 上の情報を効率よく探索するためのマークアップ言語の設計、プロトタイプを実装する。
- OIL (Ontology Inference Layer/Language) [19]: 2000 年～、欧州。On-To-Knowledge。予算規模は 250 万ユーロ (約 2.8 億円)。オントロジーの表現言語の設計、プロトツール作成を目的とする。当初理論寄りとみられていたが、2000 年 12 月には DARPA との共同プロジェクト DAML+OIL もスタートし、現実問題への適用が計画されている。

#### 3.12.3.5 おわりに、提言

こうした次世代 Web に関しては、残念ながら日本では、INTAP (Interoperability Technology Association for Information Processing、情報処理相互運用技術協会) がタスクフォースやワーキンググループを立ち上げて SemanticWeb に関する調査を行っている程度である。ただし、まだ具体的なプロジェクトは立ち上がっておらず、欧米に比べて遅れているのが実状である。

本報告で述べたような Web マイニング技術は、このような SemanticWeb や Web サービスの世界では、メタデータやオントロジーの構築を補助し、人手コストを減らす技術として今後は有効に活用されると考えられる。

SemanticWeb のようにリソースにメタデータを付与するアプローチは、海外では電子政府系の公開情報を中心にすでに行われつつある。わが国もこうした次世代 Web に遅れをとらないように、早期に研究プロジェクトを立ち上げ、研究者の裾野を広げるとともに、積極的にこのような規格にも絡んでいく必要があるだろう。さもないと、日本の社会事情に則したようなメタデータやオントロジーの記述で、制限を受けるなどの問題も起こりかねない。

## 参考文献

- [1] Sergey Brin and Lawrence Page: The Anatomy of a Large-Scale Hypertextual Web Search Engine, WWW7 Conference (1998).
- [2] Soumen Chakrabarti, Martin van den Berg, and Byron Dom: Focused Crawling: A New Approach to Topic-specific Web Resource Discovery, WWW8 Conference (1999).
- [3] Junghoo Cho, Hector Garcia-Molina, and Lawrence Page, Efficient Crawling Through URL Ordering, WWW7 Conference (1998).
- [4] Mark Craven, Dan Dipasquo, Dayne Freitag, Andrew McCallum, Tom Mitchell, Kamal Nigam, and Sean Slattery: Learning to Construct Knowledge Bases from the World Wide Web, Artificial Intelligence, Vol. 118, pp. 69-113 (Apr. 2000).
- [5] Robert B. Doorenbos, Oren Etzioni, and D. S. Weld: A Scalable Comparison-shopping Agent for the World-Wide Web, Autonomous Agents (1997). (<http://www.cs.washington.edu/homes/etzioni/>)
- [6] Usama M. Fayyad, etc.: Advances in Knowledge Discovery and Data Mining, MIT Press (1996)
- [7] David Gibson, Jon Kleinberg, and Prabhakar Raghavan: Inferring Web Communities from Link Topology, Proceedings of the 9th ACM Conference on Hypertext and Hypermedia (1998).
- [8] J. M. Kleinberg: Authoritative Sources in a Hyperlinked Environment, Proceedings of 9th ACM-SIAM Symposium on Discrete Algorithms (1998).
- [9] R. Kosala and H. Blockee: Web Mining Research: A Survey, ACM SIGKDD Exploration, Issue 2 (June 2000). (<http://www.acm.org/sigs/sigkdd/explorations/issue2-1/contents.htm>)
- [10] Nicholas Kushmerick: Wrapper Induction: Efficiency and Expressiveness, Artificial Intelligence, Vol. 118, pp. 15-68 (Apr. 2000). (同アイデアは IJICAI97 で最初に発表されている)
- [11] 元田浩: 情報洪水時代におけるアクティブマイニングの実現, 科学研究補助金「特定領域研究 (B)」. ([www.ar.sanken.osaka-u.ac.jp/activemining/](http://www.ar.sanken.osaka-u.ac.jp/activemining/))
- [12] Marc Najork and Janet L. Wiener: Breadth-first Search Crawling Yields High-quality Pages, WWW10 Conference (2001).
- [13] 坂元, 有村: Web マイニング, 人工知能学会論文誌, Vol. 16, No. 2, pp. 233-238.
- [14] Hiroshi Tsuda, Takanori Ugai, and Kazuo Misue: Link-based Acquisition of Web Metadata for Domain-specific Directories, PKAW2000, pp. 317-324 (2000).
- [15] 津田, 鵜飼, 三末: Web ディレクトリのためのページメタデータの自動付与の試み,

情報学シンポジウム 2002, pp. 17-24 (2002).

- [16] 浦本: Semantic Web – 機械のための Web –, 人工知能学会, Vol. 16, No. 3, pp. 412-419 (2001).
- [17] 山西: Web マイニングと情報論的学習理論, 情報学シンポジウム 2002, pp. 9-16 (2002).
- [18] DARPA Agent Markup Language, <http://www.daml.org/>
- [19] Ontology Inference Layer/Language,  
<http://www.ontoknowledge.org/oil/index.shtml>
- [20] Open Directory Project, <http://dmoz.org/>
- [21] W3C, Semantic Web, <http://www.w3.org/2001/sw/>



## 3.13 3次元視覚システムの技術動向と展望

報告者: 富田文明講師

### 3.13.1 コンピュータビジョンの歴史

研究初期の 1970 年代の研究は、パターン認識、画像処理、シーン解析がキーワードとなる。パターン認識の成果として、文字認識などができるようになった。画像処理は、製造業の大量生産工程で広く使われるようになった。シーン解析は、3次元視覚処理の走りである。当時は、積み木の世界しか対象にできないと批判されたが、現在製品化されているレンジファインダを生み出した。現在利用されている画像処理アルゴリズムのほとんどが 1970 年代に考案され、その後 SPIDER という画像処理のライブラリにまとめられ、国内外で広く流布している。このように、1970 年代には、実用的にも成果が生まれた。

1980 年代には、人工知能 (AI: artificial intelligence) の分野でエキスパートシステムが注目され、その影響を受けた画像理解がキーワードとなる。豊富な知識を使って画像を理解させようとしたが、知識の定義のあいまいさの問題に直面する。その反動もあって、撮像条件を非常に限定して、2次元の画像に投影されて縮退した情報から、どうやって3次元の世界を復元するかという、数理的な研究が主流となった。それらは、Shape-from-X と総称され、X には Shading、Texture、(Short) Motion などがある。理論的な研究は 1990 年代の中ごろまで続くが、あまり実用的な成果が生まれず、結果として、学界と産業界が乖離してしまった。

一方、計算機の高性能化を受け、正則化原理と呼ばれる処理は単純だが計算量の多いシミュレーション的な手法が一時期流行した。Snake と呼ばれるシステムが最初で、口のまわりの境界線を、エネルギー最小原理に基づく処理の繰り返しによって、だんだんと正確に検出する。変形するやわらかい対象物には有効な手法の 1 つである。

1980 年代後半になって、自動車に眼をもたせて道路を走らせる自律走行車 (ALV: Autonomous Land Vehicle) が、ドイツのベンツ社とミュンヘン連邦軍大学との共同で始まった。あとから始めたカーネギー・メロン大学の NAVLAB の方が、大陸横断するなどして有名になった。自動車会社を含めて研究は活発化したが、技術的には、条件の良い高速道路などの走行に限定され、道路の端や白線を見つけて走行制御する程度にとどまっており、実用化にはまだ遠いことが認識され、現在は下火になっている。高度道路交通システム (ITS: Intelligent Transport Systems) のプロジェクトの課題のひとつとなっているが、プロジェクトの主体は、運転を補助する道路などのインフラストラクチャの整備にある。

1990年代には、欧米が中心となって SPIDER の国際版にあたる画像理解用標準ソフトウェア (IUE: Image Understanding Environment) のプロジェクトが始まったが、1980年代に技術的に大きな進歩がなかったためか、結果的にはしりつぼみとなっている。このころからコンピュータビジョン (CV: computer vision) の研究者は、応用分野として仮想現実感や複合現実感に眼を向け始め、顔認識、ジェスチャ認識といったヒューマンインタフェースに関連する研究が多くなっている。最近では、ウェアラブルコンピュータの影響を受け、ウェアラブルビジョンと称して、眼鏡的に小型のカメラとディスプレイを装着し、インターネット風に、注目している部分をクリックして関連する情報を取り出すといったシステムも実験的に開発され始めている。

1990年代の後半になって、ようやく3次元視覚への関心が向き始める。3～数台のカメラを用いて相関法によって実時間で距離計測するステレオカメラシステムが市販されるようになってきた。精度は粗いが、物体の有無の判断程度には利用できる。理論の方でも、Shape-from- (Long) Motion、すなわち、1台のカメラの移動によって得られる時系列画像から3次元形状を復元する有効な因子分解法が提案され、簡易に屋外環境の3次元データを取得する方法 (セルフキャリブレーション) として期待される。

2000年代になって、さらに何十台ものカメラを使うマルチカメラシステムが出現している。カメラをドーム状に配置し、真中の人間や物体の距離を、全方向から時系列的に同期計測した後、任意視点からの映像を出力する。まだ技術的な課題もあるが、いずれは仮想現実感などに応用されていくであろう。

マルチカメラが1点を集中して撮影するのに対して、分散視覚は多数台のカメラの視野を補うように配置し、例えば、動く物体を、その情報を引き渡ししながら、順番に追いかけていく。マルチカメラも分散視覚も、背景には PC クラスタの出現という、計算機の進歩が後押ししている。

現状の研究をまとめると、技術的にはあまり大きな進歩はないが、応用志向が強まり、計算機の進歩もあって、多様な分野に適用しようとしている。しかし、必要な視覚機能が充分でないと、いずれは限界に達することはすでに経験していることに注意したい。

#### 3.13.2 汎用システムの必要性

従来のシステムは、あるタスクが与えられた場合に、コストと高速性に工夫をこらした専用システムを専門家が構築していた。この方法は、例えば、製造業の大量生産工程には有効であったが、現在のような多品種少量生産には限界がきている。製造の現場では、いろいろな作業に適用できる汎用型のシステムが、本質的に求められている。

汎用システムの開発には、共通的な「機能」と専用の「知識」を分離し、必要な機能を体系的にモジュール化してそろえることが必要である。従来の専用システムは、機能の

不足を知識で補っていた。そうすると、知識はタスクごとに異なるから、まったく独立のシステムとなってしまふ。機能と知識が混在していると、他のシステムに流用できない。

しかし、今でも、汎用システムは役に立たないという、コンピュータが高価で性能が低い時代の古い迷信が根強く残っている。しかし、汎用システムであっても、専用システムと同様に、タスクに応じた効率的な処理ができなければならないのは事実である。十分な機能を用意しておけば、知識を入れかえるだけで、専用システムができると考える。専門家でない利用者が、プログラミングをすることなく、対話的に必要な機能を選択し、パラメータを調整しながら、専門家が作るシステムと同等以上の性能をもった専用システムを生成できるタスク指向化システムを開発する必要はある。

このようなシステムに対して、対照的な 2 つの疑問が投げかけられる。専門家でない一般の多くの人には、すでに汎用システムができていると思われている。工場などで視覚を使ったシステムがたくさん実用化されているが、専用システムでは、対象となる物体や観測の環境が非常に限定されていることが専門家でないと見た目にはわかりにくい。

逆に、専門家の間では、汎用システムはすぐには実現できないと思われている。これに対しては、実際に汎用型システムを作り、その有効性を実証するしかない。

### 3.13.3 3次元視覚システム

産業技術総合研究所（旧電子技術総合研究所）では、多分野で多目的に利用できる高機能 3次元視覚システム VVV (Versatile Volumetric Vision) [1]を 10年以上にわたって体系的に開発しており、現在も進化している。

3次元視覚の基本プロセスを図 3.13-1 に示す。まず、3次元視覚センサによって、観測している環境（以下、シーン）の 3次元距離を計測する。次に、この距離データからシーンに存在する物体の面、境界線、頂点などの特徴を抽出し、その 3次元形状を計測する。そして、この形状データとあらかじめ登録している物体の形状モデルとの照合を行い、その 3次元の位置と姿勢を決定する。いわゆる物体認識である。さらに、物体が動いている場合は、照合処理を継続的に繰り返すことによって、その 3次元運動を追跡する。物体のモデルの登録には、3次元 CAD によって合成する方法と、自然物など CAD では合成が難しい場合には、実物の 3次元形状データを計測し、それを形状モデルとして利用することができる。

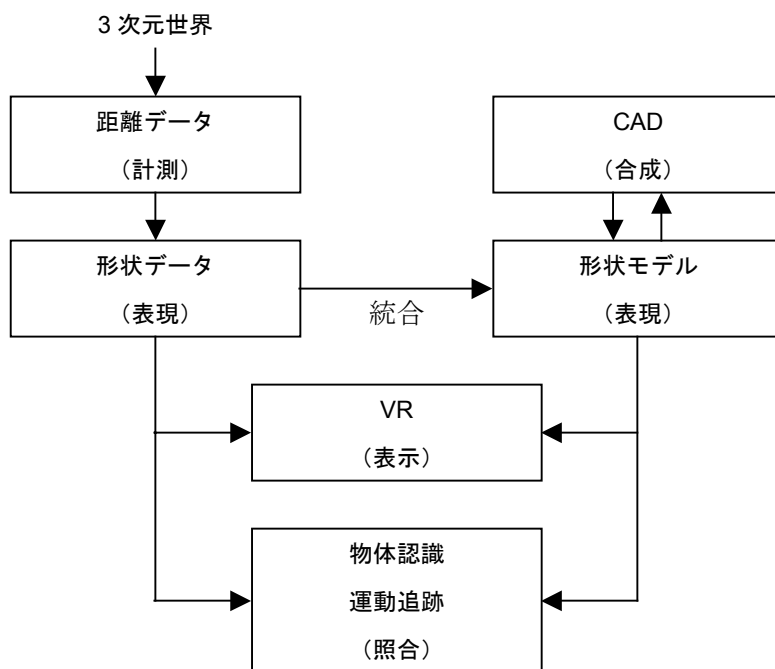


図 3.13-1 3次元視覚プロセス

### 3.13.3.1 距離計測、形状計測

距離を計測するための3次元視覚センサの基本はステレオビジョンであり、装置として複数台のカメラと画像入力ボードを用意すればよいが、用途や対象に応じて、いくつかのシステムを開発し、利用している。図 3.13-2 のアクティブ3眼ステレオカメラシステムは、3台のカメラを用いる最も基本的なシステムである。注視点を変えることができるように、ヘッドのパンとチルト（首振り）、左右カメラの輻輳角と中央カメラの俯角（眼球運動）、そして、各カメラのズーム、フォーカス、アイリスをコンピュータ制御することができる。さらに、レーザ光の光源を取りつければ、物体に光を当てて距離を計測するレーザレンジファインダとなる。

カメラが2台か3台かでは、大きな違いがある。物体までの距離は、3角測定の原理によって測定する[2]。しかし、2台のカメラを水平に置いた場合、水平線の距離が計測できない。人間も本来は同じであるが、豊富な知識や経験を使って距離を推定している。2台のカメラでも技術的にある程度は可能であるとしても、まだそれ程賢くないコンピュータに推定させて間違える危険を冒すよりも、垂直方向にもう1台のカメラを用意する方が妥当と考える。3台以上は、何台であっても大差はない。

また、カメラ 1 台でも移動して撮像すれば、セルフキャリブレーションにより、ステレオカメラと同等の距離計測が可能である。

単一方向から観測では、対象物の一面のデータしか得られない。したがって、実物を使ってそのモデルを作成する場合は、対象物を回転台の上に置き、回転させながら距離を測定することにより、全周の 3 次元データが得られる。動物のように静止していない対象物の場合には、マルチカメラシステムによって同期的に計測することができる。

このようにして得られた距離データは単なる 3 次元の点の集まりなので、これから物体の面、境界線、頂点などの特徴を抽出して、3 次元 CAD と相互変換可能な幾何形状データに変換する。

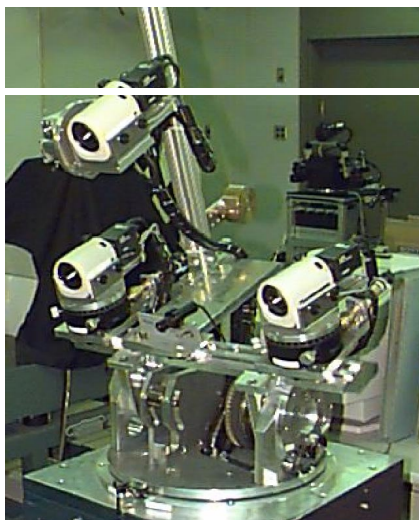


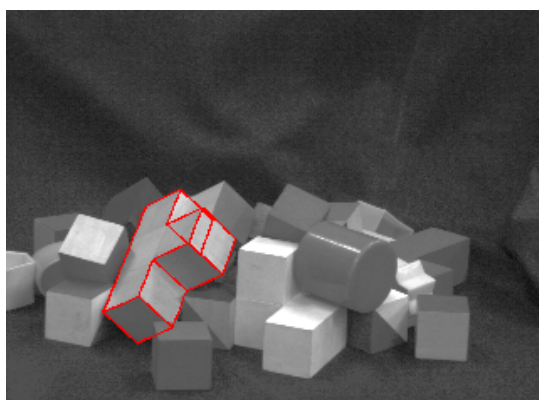
図 3.13-2 アクティブ 3 眼ステレオカメラシステム

### 3.13.3.2 物体認識、運動追跡

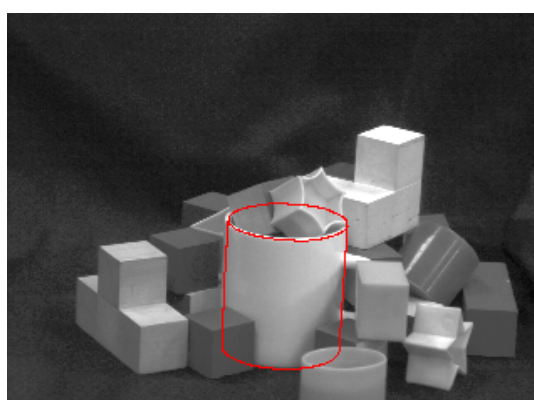
物体の認識とは、既知物体の 3 次元モデルと観測シーンの 3 次元データを照合して、その物体の存在と 3 次元 6 自由度の位置・姿勢を計測することである[3]。このとき、汎用性の必要条件として、観測条件や対象をできるだけ限定しないことが要求される。すなわち、任意の 3 次元形状を対象とし、表面に模様や光沢があってもよい。対象物がどんな位置、姿勢にあってもよく、不特定の背景に不特定の物体と混在していてもよい。部分的に隠れていてもよい、などの「無条件」を克服しなければならない(図 3.13-3)。技術的には、3 次元の距離データを扱い、構造的に解析して、パターン照合する 3 次元視覚が必要である。従来の、2 次元の画像データを扱い、統計的に解析して、パターン分類する 2 次元視覚だけでは対処できない。2 次元視覚では、絵に描いた餅が絵(平面)かどうかもわからない

のである。

対象物が動き、その動きを追跡する場合には、連続的に入力される時系列ステレオ画像間の時空間照合により、物体の3次元運動パラメータ（3次元位置・姿勢の変化）の計測を繰り返す[4]。物体認識と同じ照合処理ではあるが、時系列画像間では対象物は接近しているため探索範囲が小さく、通常の計算機で物体認識が秒単位であるのに対して、運動追跡はミリ秒単位の処理が可能となる。



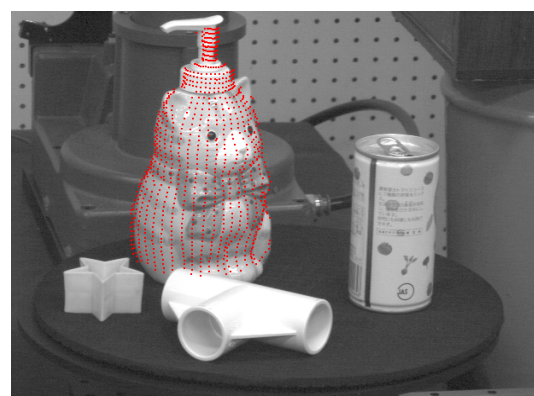
(a) 他の物体と混在する積み木



(b) 部分的な隠れのあるコップ



(c) 模様のある空き缶



(d) 自由曲面体の陶器

赤線または赤点が物体のモデルを表し、その物体の3次元の位置と姿勢が認識されたことを示す。

図 3.13-3 物体認識

### 3.13.4 モデルの階層

視覚の対象のモデルには、幾何モデル、物理モデル、共通モデルの3階層がある。

モデルの基礎は幾何モデルであり、個々の対象を定量的に正確に表現する。3次元CADで標準的な境界表現（b-rep）が基本データ構造である。曲面を対象とする場合には、曲面形状を表すネットモデルが付加される。幾何モデルは観測から直接得られる表現であり、同一サイズ、同一形状の対象を認識する場合や、CAD、VRのための3次元データを生成する場合に利用する。

物理（作業）モデルは、外界に作用する機械（マニピュレータ、自律走行車など）との協調に用いられる。例えば、幾何モデルには、把持可能な部位、走行可能な面など、対象の物理的特徴に関する情報がないので、これらを表現する。

共通（概念）モデルは、対象を定性的に表現する。例えば、幾何モデルを用いて自動車を認識する場合には、世界中の車種のモデルを用意しておかなければならない。これに対して、共通モデルは、自動車を包括的に表現するもので、一度も見たことのない自動車であっても自動車と認識することができるもので、CVの最大のオープンプロブレムである。この共通モデルは、パターン情報の記号化、言語化へと一般化される。パターンと自然言語（名詞、動詞、形容詞、副詞）との結びつきは、自然言語に実体を与え、人との対話をより自然なものとすることができる。

### 3.13.5 応用システム

3次元視覚システムは本来、人間の眼が必要とされる多くの作業や機械に共通的に利用でき、その自動化を促進することが期待される。前述のモデルの階層に対応して、3次元視覚の典型的な応用システムとして、製造分野におけるハンドアイロボットシステム、交通分野における自律走行システム、福祉分野における視覚障害者用視覚代行システムなどを開発している。

#### 3.13.5.1 ハンドアイロボットシステム

視覚とマニピュレータが協調して行う各種物体の各種操作の基本動作を体系化してライブラリ化することにより、タスク毎のプログラミングを必要としないハンドアイシステム TORCS (Task-Oriented Robot Control System) [5]を開発している (図 3.13-4)。以下は、これまでに試作した実験システムの例である。

- ・ 各種のブロックを積み上げるシステム。
- ・ 多くの物の中から空き缶を認識してゴミ箱に捨てるシステム。
- ・ 自由曲面体であるバナナを認識してお皿に置くシステム。
- ・ 移動するコンベア上にランダムに置かれた鋳造品の動的 Pick & Place システム。
- ・ 部品を穴に挿入するときに、穴が動いても穴を追いかけて挿入できるシステム。
- ・ 操作物体の移動軌道の教示システム。

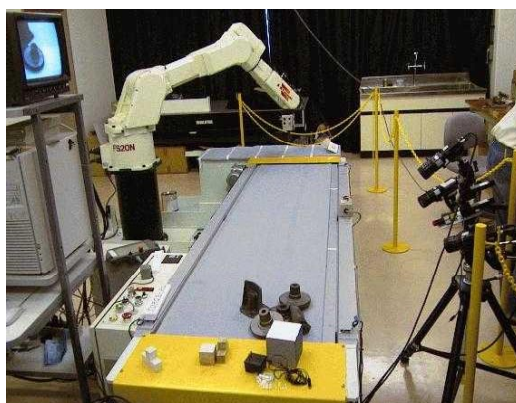


図 3.13-4 ハンドアイロボットシステム

#### 3.13.5.2 自律走行システム

外部に補助的な軌道施設や位置検出装置などを必要とすることなく、屋内や屋外を自律走行するシステムを開発している[5]。実験用車両として、バッテリー駆動のゴルフカートを利用し、3次元視覚センサとして、アクティブ3眼ステレオカメラシステムを搭載している (図 3.13-5)。



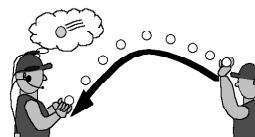


図 3.13-5 自律走行車

これまでに、廊下を自律走行するための基本機能として、走行路を認識し、障害物を回避しながら走行制御し、右折や左折時にターニングポイントを検出してターンし、また、自動的に車庫の出入ができるシステムを開発した。現在、簡単な地図として、ターニングポイント間の距離情報を与えるだけで、目標地点まで自動的に移動することが可能となっている。

### 3.13.5.3 視覚障害者用視覚代行システム

視覚障害者をユーザとして、図 3.13-6 のように、小型ステレオカメラ（映像入力）と骨伝導ヘッドフォン（音響出力）を装着することにより、歩行やスポーツなどの日常活動ができるように、ユーザが必要とする各種（特に、3次元）の視覚情報を、オンライン、実時間で、有効かつ適切に知覚できる聴覚情報に変換し、伝達するウェアラブルシステムを開発している[6]。



スポーツ支援

図 3.13-6 視覚障害者用視覚代行システム

## 3.13.6 将来への展望

視覚やロボットの研究は、もともと広い意味での AI の一環として始まったが、現在では記号処理としての AI から独立している。しかし、本来ひとつのシステムとして統合すべきであるし、それができる時期にきている。人間が扱っている情報は 80%以上が視覚情報と言われており、AI やロボットが視覚を使わないことには無理がある。今後の AI やロボットの進展は、視覚の成熟にかかっているかもしれない。

例えば、現在目標とするロボットは、一般的な環境で、自ら状況を判断しながら自律的に行動する知能ロボットである。しかし、専門家のプログラミングに頼っている産業用ロボットだけでなく、今話題の人間型ロボット ASIMO にしても、眼がついているが、遠隔操縦が基本であり、ペットロボット AIBO はピンクのボールにだけ反応し、パーソナルロボット PaPeRo は人間の顔を統計的に認識する程度であるため、できる行動が限定される。

AI も 1980 年代のエキスパートシステムの後低迷している。情報が完全である問題のみ解決可能な従来の演繹推論に替わり、情報が不完全な実用問題を解決するために、帰納推論の方法が模索されている。また、抽象的な記号を処理するのではなく、物理的な実体を持つこと-身体性が必要とされている。

今後は、CV だけではなく、AI、ロボット、知識を検索するデータベース、ヒューマンインタフェースなどを含めた、統合的な情報処理システムを目指している。やはり、システムは汎用であるが、機能はモジュール化されており、データベースの知識を入れかえれば、専用システムとしてさまざまに応用できるようにしたい。システム構成として、個別の機能の一部だけを使ってもよいし、すべてを一体化して自律ロボットにしてもよい。各機能には、信号のレベルから自然言語のレベルまで、いくつかの階層がある。情報の交換やフィードバックは、異なるサブシステムの間だけでなく、同じ機能の各階層の間でも行わなければならない。

## 参考文献

- [1] 富田文明: 高機能 3 次元視覚システム VVV, 情報処理, Vol. 42, No. 4, pp. 370-375 (2001).
- [2] 河井良浩, 植芝俊夫, 石山豊, 角保志, 富田文明: セグメントベースステレオにおける連結性と対応評価, 情報処理学会論文誌, Vol. 40, No. 8, pp. 3219-3229 (1999).
- [3] Y. Sumi, Y. Kawai, T. Yoshimi and F. Tomita: 3D Object Recognition in Cluttered Environments by Segment-Based Stereo Vision, International Journal of Computer Vision, Vol. 46, No. 1, pp. 5-23 (2002).
- [4] 石山豊, 角保志, 富田文明: ステレオビジョンによる 3 次元物体の 3 次元運動追跡, 日本ロボット学会誌, Vol. 18, No. 2, pp. 213-220 (2000).

- [5] 平成 10 年度地域コンソーシアム研究開発事業「3 次元視覚システム VVV を用いたタスク指向型・ロボット制御システム TORCS の研究開発」成果報告書, 新エネルギー・産業技術総合開発機構 (NEDO) (2000).
- [6] 河井良浩, 小林真, 皆川洋喜, 宮川正弘, 富田文明: 3 次元仮想音響による視覚障害者用支援システム, 電気学会論文誌 C (電子・情報・システム部門誌), Vol. 120-C, No. 5, pp. 648-655 (2000).

## 付属資料 1 海外調査報告

報告者: 大須賀昭彦委員

人間主体の知的情報技術調査ワーキンググループ活動の一環として、Autonomous Agents 2001 への参加、米国におけるエージェント研究開発動向の調査などを行ったので報告する。

日程: 2001 年 5 月 29 日から 6 月 6 日まで

調査員: 大須賀昭彦委員

主な調査先:

- (1) Autonomous Agents 2001
- (2) Carnegie Mellon 大学
- (3) SRI International

### 1 Autonomous Agents 2001

International Conference on Autonomous Agents は ACM 他が主催する、自律的エージェント技術に関する国際会議。MA (International Conference on Mobile Agents) や ICMAS (International Conference on Multi-Agent Systems) 等と並び、エージェント技術の国際会議としては最大規模のもの。2001 年はカナダのモントリオールで開催され、248 件の論文投稿中 57 件がフルペーパーとして採択された。エージェント分野の著名な研究者が多数出席し、参加者は約 500 名となった。

## (1) 招待講演

招待講演は4件行われた。以下に講演概要を紹介する。

Microsoft 主席研究員 Eric Horvitz の"Attention, Intention, and Initiative: Toward Insightful Interactive Agents"では、ユーザの心理を考慮した種々のツールに関する研究が紹介された。特に、ユーザの注目点を検知して GUI の制御を行うデモを数多く見せていた。Office アシスタントは失敗であったとの見解から、このような GUI に至ったとのことである。

Pennsylvania 大学教授 Norman Badler の"Embodied Agents with Behavioral Consistency"の講演は、コンピュータグラフィックスの作成に関するものであった。バーチャルキャラクターの CG 作成において、人間の挙動の一貫性（体の各部の動きに矛盾がないこと）に注目して、よりリアルな挙動を実現する技術が紹介された。

Brandeis 大学教授 Jordan Pollack の講演は、ロボットへの学習機能の搭載に関するものであった。人間を模するロボットにおいて、すべての機能をプログラムで実現するのは事実上困難であり、機械学習が必要になるとの主張が展開された。

Carnegie Mellon 大学準教授 Tuomas Sandholm の講演は、Autonomous Agents Award の受賞記念講演で、エージェント技術を用いたオークションに関するものであった。複数の商品を扱うオークションは NP 完全となり、近似的にも多項式時間で解けないので、ここに人工知能技術を適用して解決しようという研究が紹介された。



会場の Marriott ホテル

(2) 一般講演

一般講演は、3トラックの並列で、18セッションが行われた。以下にセッションの内訳を示す。

(a) 学習、推論、プランニングに関するセッション

- Action Selection & Planning
- Adaptation and Learning (2件)

(b) マルチエージェントに関するセッション

- Coordinating Multiple Agents
- Multiagent Terms
- Agent Communication Language
- Agent Societies
- Multiagent Communication & Collaboration

(c) 擬人化、人間対エージェントに関するセッション

- Collaboration between Humans and Agents
- Models of Emotion and Personality
- Human-like Qualities of Synthetic Agents

(d) e-コマースに関するセッション

- Artificial Market Systems & E-Commerce (2件)

(e) ソフトウェア工学に関するセッション

- Software Prototypes (3件)
- Agent Architectures
- Agent-Based Software Engineering

学習関連は3セッションが開かれ、プランニングの拡張や、BDI (Beliefs, Desires and Intentions) に関する発表が多く見受けられた。例えば、J. Broersen (オランダ Vrije Univ. de Boelelaan) らの"The BOID Architecture"の発表では、BDI アーキテクチャを

obligation の概念で拡張する BOID アーキテクチャが提案された。これにより、エージェント間の競合解消などが容易に実現できるとのことである。

マルチエージェントは 5 セッション開かれ、標準化が進むエージェント間コミュニケーション言語に、いかに意味を形式化して導入するかといった話題が多数あった。例えば、J. Pitt (英国 Imperial College) らの "Interaction Patterns and Observable Commitments in a Multi-Agent Trading Scenario"、および "Denotational Semantics for Agent Communication Languages" の発表では、AUML (Agent UML)、および FIPA (Foundation for Intelligent Physical Agents) ACL (Agent Communication Language) に形式的意味論を与えることにより、厳密な分析を可能にする手法が提案された。

擬人化関連は 3 セッションで、感情の導入や人間との対話などに関するモデルの形式化、試作結果の報告が行われた。

e-コマースは 2 セッションの他に、パネルディスカッションも開催され、いずれも多数の参加者を集めていた。

ソフトウェア工学関連は 5 セッションで、エージェントソフトウェアの要求仕様獲得、設計方法論、開発支援環境などが報告された。例えば、A. Perini (イタリア ITC 研究所) らの "A Knowledge Level Software Engineering Methodology for Agent Oriented Programming" の発表では、エージェントアプリケーション開発における早期の要求定義を支援する開発方法論 Tropos が提案された。この方法論により、従来のエージェント指向開発方法論に欠けていた早期の要求定義が可能になるとのことである。エージェント分野でのソフトウェア工学は、ここ 1、2 年で研究が活発化しており、ソフトウェア工学分野で最大規模の会議である ICSE (International Conference on Software Engineering) でも、エージェント技術に関する発表が増加している。

### (3) ワークショップ

ワークショップは full day workshop 9 件、half day workshop 5 件が開催された。

#### (a) Full Day Workshops

- Agent Oriented Information Systems
- Infrastructure for Scalable Multi-Agent Systems
- The Second Workshop on Agent-Oriented Software Engineering
- Ontologies in Agent Systems
- Learning Agents
- Fourth Workshop on Deception, Fraud, and Trust in Agent Societies
- Norms and Institutions in MAS
- Representing and Annotating Non-Verbal and Verbal Communicative

Acts to Achieve Contextual Embodied Agents

- Security in Mobile Multiagent Systems

(b) Half Day Workshops

- Agents and Internet Learning,
- Autonomous and Cooperative Agents for Industry
- Agent-Oriented Software Architectures for B2B
- Agent-Supported Cooperative Work
- Autonomy-Oriented Computation

以下で、2つのワークショップの概要を紹介する。

(a) Agent Oriented Information Systems

(エージェント指向情報システムワークショップ)

このワークショップはエージェント技術の情報システムへの応用に関するもので、約 40 名が参加し、招待講演 1 件を含め、8 件の発表があった。M. P. Singh (米国 North Carolina 大学教授) の招待講演 "Commitment Machine" は、柔軟なエージェントの動作記述を commitment (公約) の概念を用いて実現する手法を提案したものであった。BDI などの同様な手法よりも容易に利用できるとの印象を受けた。その他、要求定義、モバイルエージェント、オークション、オントロジー、情報収集などの話題についての発表があった。

(b) The Second Workshop on Agent-Oriented Software Engineering

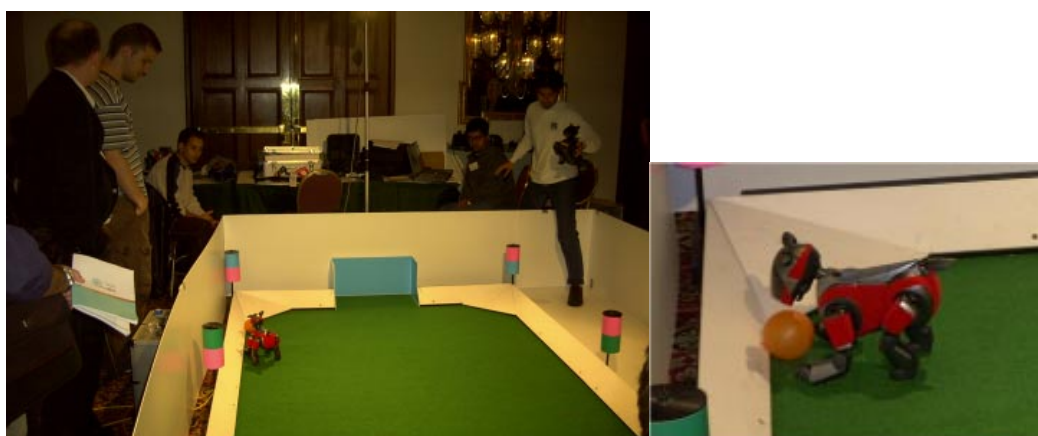
(エージェント指向ソフトウェア工学ワークショップ)

本ワークショップはエージェントシステムにおけるソフトウェア工学に関するもので、約 50 名が参加し、14 件の発表とパネル討論があった。多くの発表は、UML (The Unified Modeling Language) のエージェントアプリケーション開発への適用に関するもので、パネル討論のテーマも "UML and Agents" であった。

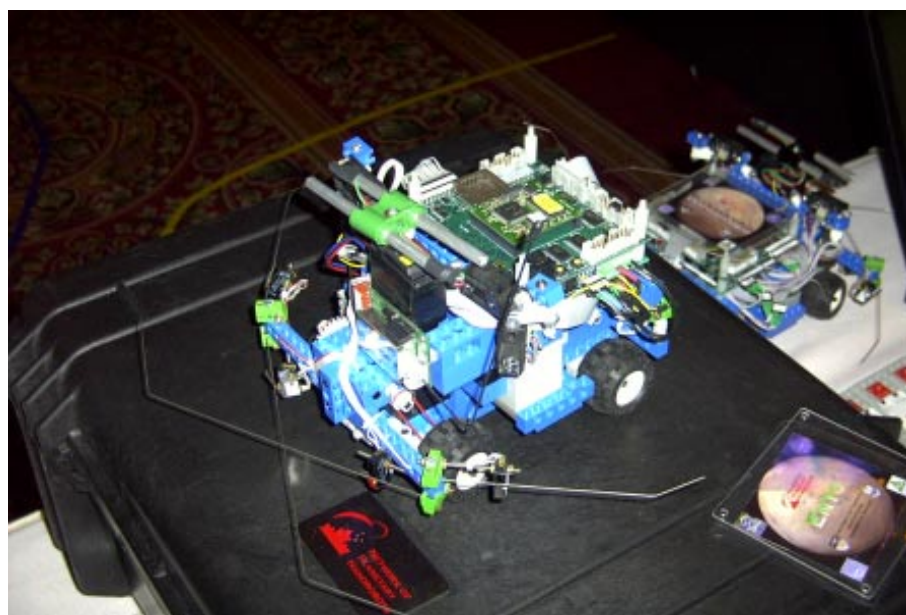


#### (4) その他のイベント

ロボットイベントでは、AIBO サッカーリーグ、ロボカップジュニアなどが行われた。Toronto 大学の RISE (Robotics in Space Exploration) プロジェクトは、宇宙探索をロボットで行うというものであった。6 台の惑星探査車両のそれぞれが CPU を持ち、ある程度は自律的に動作しながら、互いに通信することによって協調してタスクを遂行するデモが行われた。



AIBO サッカーリーグの様子



RISE における惑星探査車両

デモンストレーション会場では、SemanticWeb とモバイルエージェントに関するデモが多く見受けられた。FIPA 標準を盛り込んだ組み込み向けモバイルエージェント LEAP による、PDA と PC 混在環境上でのスケジュール調整のデモなどが行われた。他にはインタフェースエージェントに関するもの、DAML (DARPA Agent Markup Language) に関するものへの関心が高く、見学者を多く集めていた。



デモンストレーション会場の様子

パネル討論"Agents in E-Commerce"では、EC におけるエージェント技術の現状がパネルから紹介された後、社会的問題やその他の Web 技術などとの関連について討論が行われた。

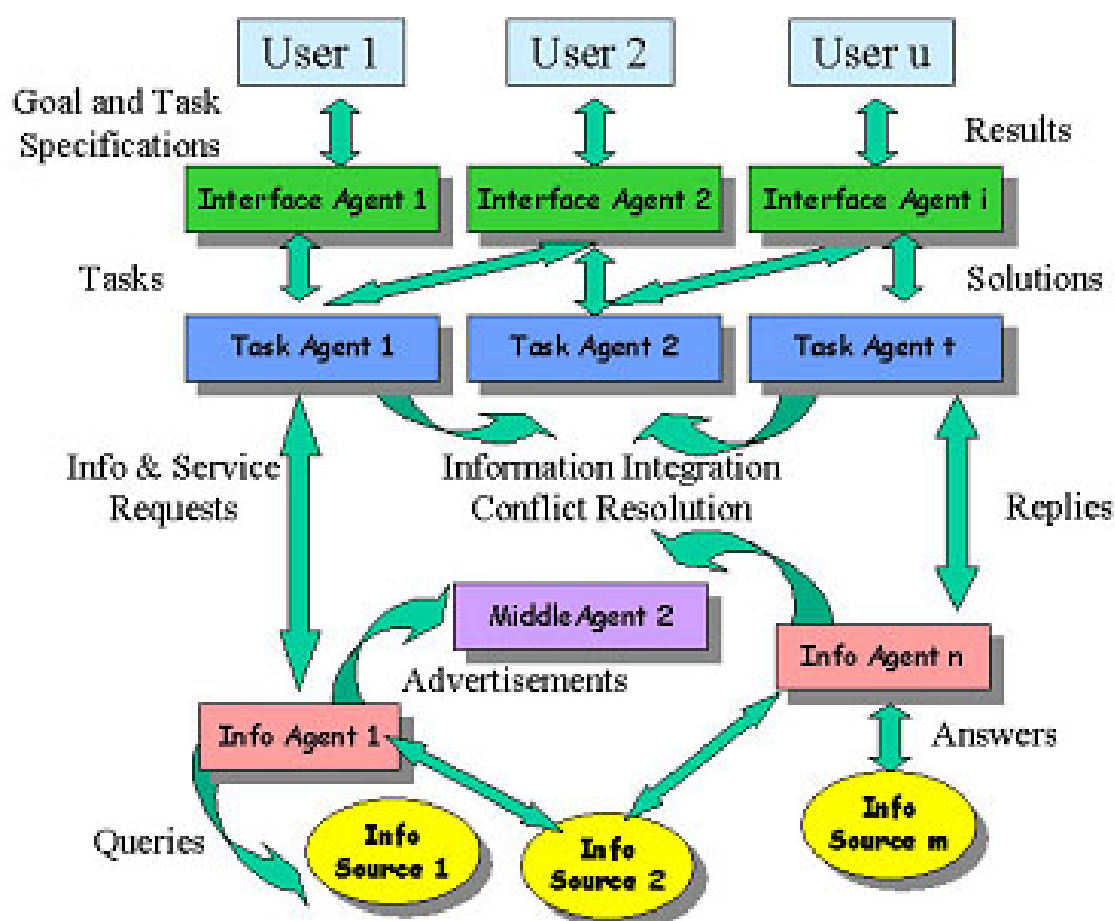
## 2 Carnegie Mellon 大学 Katia Sycara 教授訪問

Sycara 教授の研究グループは 10 名前後のメンバーで知的エージェントの研究に取り組んでいる。DARPA からの出資や企業との提携により多数のプロジェクトが推進されているが、特に DAML やユビキタスコンピューティングに関する研究が多い。以下に現在のプロジェクトリストを示す。

- e-Commerce
- Middle Agents
- WebMate
- Language for Agent Advertisement & Request
- Retsina MultiAgent Infrastructure
- Agent-Based Aircraft Maintenance
- Agent-Based Team Aiding
- Agent-Based Route Planning
- MultiAgent Learning
- Agent Interoperability
- MultiAgent System Interoperability
- MultiAgent System Security
- Intelligent Agent Matching System



Sycara 教授と研究グループメンバ、報告者



Retsina MultiAgent Infrastructure の構成

- MultiAgent DeMining System
- MultiAgent Calendar Scheduling
- Financial Planning Agents

この中で、以下の2つのプロジェクトについて議論した。

Retsina MultiAgent Infrastructure は異種エージェントのコミュニティを支援するオープンなマルチエージェント基盤である。この基盤上のエージェントはピアツーピアで相互に作用しながらコミュニティを組織する。このコミュニティの構造は外部環境やネットワーク構造などの影響を受けず、エージェント間の関係から柔軟に構築される。これにより、Retsina ではマルチエージェントの集中管理が不要となり、エージェントによる分散サービスの提供が可能となる。

Intelligent Agent Matching System は、登録されたサービスの意味的な検索を行うサービスマッチメイキング技術である。オントロジーをチェックするオントロジーフィルタ、サービスの入出力の型チェックなどを行うシンタックスフィルタ、サービスの入出力条件の論理的な整合性をチェックするセマンティクスフィルタが用意されている。サービスマ



David Martin と報告者

タッチメイキングは、Web サービスにおけるサービスディレクトリ規約 UDDI (Universal Description, Discovery, and Integration) の出現によってその重要性が広く認識されつつあるが、この研究は、Web サービスが広まる前から進められてきた。これまでは、独自のサービス記述言語を用いていたが、現在はこれを DAML へ適用する試みが行われている。

### 3 SRI International David Martin 氏訪問

David Martin は DAML-S 開発の中心的な人物で、Carnegie Mellon 大学の Sycara 教授の研究グループとも共同して活動している。現在 Web サービス分野においてサービス仕様記述言語の標準となりつつある WSDL (Web Service Description Language) は、記述力不足の面があると言われている。このため、Marin は DARPA の支援によって主流になりつつある DAML を基にした、新たな仕様記述体系 DAML-S の研究を進めている。DAML-S ではサービス仕様をプロファイル、サービス・モデル、サービス・グラウンディングの 3 種の情報で記述する。プロファイルではセキュリティレベルなどを記述する。サービス・モデルではサービスを利用するためのプロセスを記述するが、WSDL よりも言語としての記述力を高めている。サービス・グラウンディングではサービスへ接続するための情報を記述する。

今のところ、DAML-S は SRI や大学主体で進められており、産業界への普及見込みは不明である。しかし、サービス仕様記述言語の記述力向上は今後の重要課題であると思われる、DAML-S の成果が何らかの形で影響をおよぼす可能性が高い。

## 付属資料 2 自然言語処理の最近の動向 — 「言語理解と行動制御」を中心に—

講演者: 田中穂積講師

採録者: 幹事

### 1 自然言語処理研究の概観

自然言語理解の研究は古く、1970年代にはタイプライタでロボットへの指示を入力していた。近年はコンピュータグラフィックス技術や音声認識技術の著しい進展の影響を受けている。

自然言語理解と行動を結びつける研究は、まだ数少ない。知る限りでは、研究者はStanford大学のBarbara Hayes-Rothくらいである。

このような研究が重要となった背景には、人間を取り巻く環境において、現実空間よりも仮想空間のなかで過ごす時間が増大していることがある。人間にとって使いやすいインタフェースを提供するためには、言葉の取り扱いが不可欠であり、そのための鍵となる技術が自然言語理解である。

自然言語理解は、従来は機械翻訳を中心に適用され、商品を生み出してきた。この場合の理解は、「浅い」水準といえる。

一方、「深い」理解においては、「これ」のような照応、語の省略、あいまい性などを取り扱う必要がある。文脈、状況、言語行為などを考慮しないと、「暑い」という発話の裏にある「水をくれ」という要求は認識できない。発話と行動が密接に関連しており、課題は難易度が高く、基礎研究の段階である。

### 2 自然言語理解システム「傀儡」

自然言語理解のプロトタイプシステム「傀儡」は、言葉による指示によって、コンピュータグラフィックスで生成したアニメーションの制御を行う。音声認識には、日本アイ・ビー・エムのViaVoiceを使用している。計量言語学、コンピュータグラフィックス、ロボット工学、認知科学などの学際領域の研究である。

エージェントとして、雪だるま、ウマ、ニワトリ、カメラの4種類が登場する。カメラ自身がエージェントである点は、特徴のひとつである。「それを押して」といった照応、「も

うちよつと」といった省略、「左」と命じたときに話者からみて左なのかエージェントの左なのかといった相対的な位置関係、などを取り扱うことができる。

「傀儡」の課題としては、エージェントが自律していないため、個性をもっていないことがある。例えば、ウマはニワトリを押せるがニワトリはウマを押せないことは、エージェントの個性ではなく、全世界を知っている神のような存在の知識として記述されている。また、カメラに対して「左」と命じたときは、ある対象を画面に入れたいという意図が考えられるが、現在は考慮されていない。

イントネーションなど発話付随音声現象や、身振り手振りといった非言語的意思疎通の考慮、さらにエージェント自身に発話させることも、今後の課題である。哲学の研究者からは、アイコンタクトをしないことが不自然であるとの指摘を受けた。音声認識、自然言語処理とも、文法的に不適格な表現を扱うため、計算量が膨大となり、課題の解決は容易ではない。

### 3 学術創成研究「言語理解と行動制御」

日本学術振興会の科学研究費補助金による学術創成研究「言語理解と行動制御」は、2001年度から2005年度までの5年間のプロジェクトである。大学の研究者25人が参加しており、3つの研究グループに分かれている。

台本にしたがってアニメーションを制作するデモンストレーションを紹介する。動作の立案は2段階に行われる。巨視的な立案によって、「部屋から出る」といった言語表現が、「扉のところに行く」、「鍵を開ける」、「取っ手を回す」といった基本動作の並びに分割される。微視的な立案によって、基本動作がアニメーションに対応づけられる。基本動作とアニメーションの対応は、一意ではない。学生に「カップをつかめ」と言ったところ、実にさまざまなつかみかたをした。類似の研究には、日本放送協会放送技術研究所の林正樹による、テレビ番組記述言語 TVML (TV program making language) がある。

自然言語理解と行動を結びつける研究には、コンピュータグラフィックスから実際のロボットまで、さまざまな応用が考えられる。医療システムに応用し、胃カメラを遠隔操作すると面白い。また、将来はすべて「耳」の機能をもつであろう、情報家電への応用も考えられる。「切れろ」と言ったとき、すべての家電のスイッチが切れては困る。

講師がプログラム言語 Scheme で作成したデモンストレーションを紹介する。登場するエージェントは、種類は「傀儡」と同じであるが、自律しており、自分にできる動作を知っている。命令はすべてのエージェントに伝えられ、各エージェントが自分に対するものかどうかを判断する。デモンストレーションでは、音声認識と自然言語処理は別に行われたものと仮定しており、「対象」と「動作」を入力する。例えば、ウマを対象に「走れ」と入力すると、ウマが走り、雪だるまとニワトリはそれを眺め、カメラは全体を撮影しつつ



ける、といった動作が、各エージェント独自の判断によって行われる。「対象」は、指定しないときには省略する。「走れ」とだけ入力すると、ウマとニワトリが走る。雪だるまは、走れないという属性をもたされているため、自分に対する命令ではないと解釈する。

奈良先端科学技術大学院大学の受付案内ロボット Aska は、身振りを行う。来客のほうを向いてアイコンタクトを行い、「後ろです」と答えるときは、後ろを指さす。

#### 4 質疑応答および意見交換

ロボットが対話している相手の位置を知る方法には、2本のマイクを使って音だけで定位する方法と、画像認識を併用する方法がある。マイクはロボットに装着されないことが多い。自分の動作音をひろってしまい、その除去が難しいためである。ソニーの AIBO のようなペット型ロボットにあまり高度な音声認識ができないのは、この理由による。日本電気の PaPeRo は、最もうまくマイクから音をひろうが、1メートル以内で話さないといけない。

音声認識にしる、自然言語理解にしる、何をもって「理解」と判断するかは難しい。例えば、音声認識において、音声を文書に変換しても、理解したことにはならない。理解の結果として行動を起こさせることによって、理解したとかどうかが直感的にわかりやすくなった。

受付案内ロボットのように、言語と行動を併用して回答する場合、両者の適切なバランスを動的に変化させることは、非常に難しい課題である。

想定される質問や命令に対しては、ロボットはある程度まともな対応ができるようになってきた。しかし、予想外の事態に対しては、デフォルトの反応を用意するぐらいしかできない。既知の事態と照合し、どれが近いかを判断するような機構も必要であろう。

ロボットの回答を相手の人間がどれくらい理解したかを判断し、それに応じて回答を変えるような、利用者モデルの研究も行われている。

本書の全部あるいは一部を断りなく転載または複写（コピー）することは、  
著作権・出版権の侵害となる場合がありますのでご注意ください。

## 人間主体の知的情報技術に関する調査研究Ⅴ

© 平成 14 年 3 月発行

発行所 財団法人 日本情報処理開発協会  
先端情報技術研究所

東京都港区芝 2 丁目 3 番 3 号  
芝東京海上ビルディング 4 階

TEL(03)3456-2511