

資 料

人間主体の知的情報技術  
に関する調査研究 VI

平成 15 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 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 主任研究官
委員	杉本 重男	筑波大学 図書館情報学系 教授
委員	長尾 確	名古屋大学 情報メディア教育センター 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 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 メディア情報研究部 特別研究員
委員	宮田 一乗	北陸先端科学技術大学院大学 知識科学教育研究センター 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	牧村 信之	(財)日本情報処理開発協会 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VI

## 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
1.1.1 研究調査の概要.....	2
1.1.2 調査方針.....	2
1.1.3 調査対象 .....	2
1.1.4 調査活動の方針 .....	4
1.1.5 調査結果の要約 .....	5
1.2 研究調査活動.....	6
1.2.1 委員会.....	6
1.2.2 米国政府支援研究開発の調査.....	8
1.2.3 各委員による研究調査 .....	9
1.2.4 海外調査 .....	10
1.3 人間主体の知的情報技術の研究調査の概要 .....	11
1.3.1 アウェアネス技術の動向の概要.....	11
1.3.2 次世代 Web とエージェントの概要 .....	11
1.3.3 IT 技術の法への応用の概要.....	11
1.3.4 WWWサーチエンジンの最新技術の概要 .....	12
1.3.5 デジタルライブラリとメタデータ - 概要と課題.....	12
1.3.6 セマンティック・トランスコーディング -より実用的な“Semantic Web”に向けて-の概要.....	13
1.3.7 多感覚情報のデータベース化と情報サービスへの応用の概要 .....	13
1.3.8 感性のモデル化技術とEコマースへの応用の概要.....	14
1.3.9 テレイマージョンとビジュアルデータマイニングの概要.....	15
1.3.10 昨今の音楽情報処理における研究プロジェクトについての概要.....	15
1.3.11 次世代音声認識の手法と実用化の動向の概要 .....	16
1.3.12 アクティブオーデイションを利用した ヒューマンロボットインタラクションの高度化の概要.....	16
2. 米国政府支援 SDP(Software Design and Productivity)研究開発の動向 .....	19
2.1 SDP 発足の経緯と PITAC の提言 .....	20

2.2	NSF SDP ワークショップ	22
2.2.1	SDP ワークショップの狙い	22
2.2.2	各検討グループからの提言	23
2.2.3	SDP ワークショップ提案のまとめ	29
2.3	2003 年度 SDP の活動計画	31
2.3.1	取り組み方針	31
2.3.2	2003 年度の SDP の主要な研究課題	32
2.3.3	2003 年度における各機関の代表的活動	33
2.3.4	2003 年度の予算	34
2.4	SDP の目指すものに関連した技術、システムの動向	36
2.4.1	モデリング、特定領域記述言語(DSL: Domain Specific Language)	36
2.4.2	多面的ソフトウェア	42
2.4.3	科学、工学的な設計	44
2.4.4	オープンな協調開発環境	49
2.4.5	コンポーネントウェア	54
2.4.6	既存ソフトウェア資産を生かす技術	58
3.	人間主体の知的情報処理に関する研究調査	65
3.1	アウェアネス技術の動向	65
3.1.1	あらまし	65
3.1.2	CSCW2002 にみるアウェアネス技術	66
3.1.3	まとめ	76
3.2	次世代 Web とエージェント	80
3.2.1	はじめに	80
3.2.2	エージェント技術に関わる最近の動向	80
3.2.3	Web サービス向けマッチメーカー	82
3.3	IT 技術の法への応用	89
3.3.1	はじめに	89
3.3.2	教育支援システム	89
3.3.3	紛争解決支援システム	93
3.3.4	関連会議など	95
3.4	WWWサーチエンジンの最新技術	99
3.4.1	はじめに	99
3.4.2	インターネットの発展	99

3.4.3	効率的な収集とインデックス化に関する研究動向.....	104
3.4.4	検索結果のランキングに関する研究動向 .....	113
3.4.5	まとめ.....	116
3.5	デジタルライブラリとメタデータ - .....	119
3.5.1	はじめに .....	119
3.5.2	デジタルライブラリの研究開発の活動を振り返る .....	120
3.5.3.	メタデータ .....	125
3.5.4	デジタルライブラリを取り巻く環境 ー情報環境の変化とそこで求められる技術.....	127
3.5.5	おわりに .....	133
3.6	セマンティック・トランスコーディング ーより実用的な“Semantic Web”に向けてー.....	135
3.6.1	はじめに .....	135
3.6.2	Semantic Web のためにやるべきこと .....	135
3.6.3	アノテーション：デジタルコンテンツの階層化 .....	137
3.6.4	トランスコーディング：デジタルコンテンツの個別化 .....	138
3.6.5	トランスコーディングの仕組み.....	139
3.6.6	提言.....	141
3.6.7	マルチメディアコンテンツのアノテーションとその応用.....	142
3.6.8	アノテーションに基づくビデオ検索システム.....	146
3.6.9	まとめと今後の課題 .....	152
3.7	多感覚情報のデータベース化と情報サービスへの応用.....	155
3.7.1	はじめに .....	155
3.7.2	感性の工学的なモデル化の新しい視点.....	155
3.7.3	感性の工学的なモデルの枠組み.....	156
3.7.4	多感覚感性データベースの技術的ニーズ .....	161
3.7.5	海外の研究動向 .....	162
3.7.6	多感覚感性データベースの応用分野.....	162
3.8	感性のモデル化技術とEコマースへの応用 .....	163
3.8.1	はじめに .....	163
3.8.2	BtoC 型のEコマースの問題点と解決のアイデア .....	163
3.8.3	感性的レコメンデーションシステムのシステム構成 .....	164
3.8.4	感性的レコメンデーションシステムの技術課題 .....	166

3.8.5	期待される波及効果 .....	168
3.9	テレマージョンとビジュアルデータマイニング .....	169
3.9.1	はじめに .....	169
3.9.2	テレビ電話とウェアネス .....	169
3.9.3	テレマージョンとメディア技術 .....	170
3.9.4	ビジュアリゼーションとビジュアルデータマイニング .....	174
3.9.5	遠隔コラボレーション .....	176
3.9.6	おわりに .....	178
3.10	昨今の音楽情報処理における研究プロジェクトについて .....	179
3.10.1	はじめに .....	179
3.10.2	OMRAS (Online Music Recognition and Searching) .....	180
3.10.3	CUIDADO (Content-based Unified Interfaces and Descriptors for Audio/music Databases available Online) .....	183
3.10.4	WEDELMUSIC (Web Delivering of Music) .....	185
3.10.5	MUSICNETWORK .....	188
3.10.6	SALIERI/GUIDO .....	190
3.11	次世代音声認識の手法と実用化の動向 .....	193
3.11.1	はじめに .....	193
3.11.2	次世代音声認識手法の動向 .....	193
3.11.3	音声処理技術の実用化の動向 .....	198
3.11.4	おわりに .....	201
3.12	アクティブオーディションを利用した ヒューマンロボットインタラクションの高度化 .....	203
3.12.1	はじめに .....	203
3.12.2	ロボット聴覚システム .....	204
3.12.3	視聴覚統合による実時間複数人物追跡 .....	205
3.12.4	アクティブ方向通過型フィルタによる音源分離 .....	207
3.12.5	分離音の音声認識 .....	212
3.12.6	実験と評価 .....	214
3.12.7	結論 .....	216
付属資料	海外調査報告 .....	219
1.	はじめに .....	219
2.	DEXA2002 の概要 .....	219

3. DEXA2002.....	220
4. ECWeb2002.....	222
5. まとめ.....	224



# 1. 総論

## 1.1 調査ワーキンググループの活動方針

奥乃 博 主査

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きた。このIT革命とも呼ばれる大きな変革の波も、実際的な応用が明確にならないまま不景気とともに急激に萎んでいる。IT革命の中身が実現されることなく、システム投資だけに明け暮れたITバブルがはじけた現在、IT革命を真に成功させ、国民の幸福に結びつけるためには、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。

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、人間が主役である『**人間主体(Human-Centered)システム**』の構築が焦眉の急である。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な課題であるが、それに偏重することは、IT革命を推進する上で片手落ちである。

人間主体システムの設計・開発では、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)、つまり、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」に、より一層着目する必要がある。従来のコンピュータシステムが、ハードウェアだけでソフトウェアがなければ、ただの箱であったように、コンテンツも単なるデータの集積である一次情報のアーカイブだけであり、書誌情報や索引などの二次情報がなければ、ただの紙くずと変わらなくなる。

人間主体システムのもう一つの課題は、子供、日本語非母国語人、身体障害者、老人などの社会的な弱者にも配慮したインタフェースの高度化が不可欠である。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、社会的弱者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・検討を引き続き行うことにした。6年目となる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さらに集積して、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。

### 1.1.1 研究調査の概要

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。

### 1.1.2 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。

### 1.1.3 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。

### 1.1.4 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 2003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  
加藤委員による DEXA 参加。
- (2) 第 1 回会議  
活動方針の討議。  
高野講師による「連想に基づく情報空間との対話技術」の講演。  
加藤委員による「DEXA 紹介、感性情報処理のビジネス応用」の報告。
- (3) 第 2 回会議  
中臺講師による「アクティブオーディションを利用した  
ヒューマンロボットインタラクションの高度化」の講演。  
長尾委員による「セマンティック・トランスコーディングと Semantic Web」  
の報告。
- (4) 第 3 回会議  
田中講師による「Meme Media and Meme Pools for Re-editing and  
Redistributing Intellectual Assets」の講演。  
山名委員による「WWW サーチャエンジンの最新技術—Google を例にとって—」  
の報告。

### 1.1.5 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り、MPEG-7 やダブリンコアなどの標準化を追求していく必要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には、基となるデータが編集された信頼性の高いデータが不可欠であること。
- (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、日本がいち早く注目し、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) 知的インタフェースの高度化には、従来にない環境認識技術、たとえば、環境アウェアネスの機能が不可欠であり、そのような情報をインタフェースで容易に利用可能とするための要素技術の開発が不可欠である。

## 1.2 研究調査活動

### 1.2.1 委員会

今年度は3回の委員会を開催した。委員会では、外部講師、及び委員から、以下に示すテーマについて、研究動向、研究成果を発表して頂き、議論を行った。

#### [第1回(9月27日)]

##### (1) 高野教授講演「連想に基づく情報空間との対話技術」

連想計算(文書群同士、単語群同士、文書群と単語群間の類似性関連計算)を用いて、キーワードではなく、文書(群)を入力して類似の文書を検索したり、文書(群)の内容把握を助ける特徴語グラフを出力したりできるシステム DualDAVI とその汎用連想計算エンジン GETA) を紹介し、そのデモを行った。

##### (2) 加藤委員講演「DEXA 紹介、感性情報処理のビジネス応用」

DEXA2002(データベースとエキスパートシステムの応用に関する会議)とそれと同時開催された EC-Web などの会議の報告、および、感性情報処理のビジネス応用として、個人の好みなど感性情報を匿名とするレコメンデーションシステム、顔画像のモニタリングによって顧客を特定しその顧客の感性データベース情報に基づき顧客好みの商品を紹介するシステムを紹介した。

DEXA2002 の詳細は付属資料に示し、感性情報処理のビジネス応用についての概要は、1.3.7、1.3.8 に示し、その詳細は、3.7、3.8 に示す。

#### [第2回(10月25日)]

##### (1) 中臺講師講演「アクティブオーディションを利用したヒューマンロボットインタラクションの高度化」

[北野共生システムプロジェクトの紹介]

共生系の概念を適用した研究として、生命現象の理解を目的とするシステムバイオロジー、および知能の実現を目的とする共生系知能の研究を実施している。

[ロボットの視聴覚の研究]

日常的環境(人と音声で、音源が見えなくても、雑音下)でも使えるロボットを目指し、①視聴覚の統合による人物追跡、②アクティブ方向通過型フィルタによる音源分離、③複数の音響モデルを利用した音声認識、に取り組んでいる。

①視聴覚の統合による人物追跡では、両耳間の位相差・強度差、倍音構造に注目した音源分離による音認識、リアルタイムな複数の顔の検出による顔認識、およびステレオビジョンからのストリーム(時系列イベント)を統合処理することにより人物追跡を実現し、二人の話者が同時に発話した時に、話者の切り替えを抽出することに成功した。

②アクティブ方向通過型フィルタによる音源分離では、聴覚中心窩ベースの通過帯域制御により、3人の話者が同時に発話した時、それぞれの話者の発話に分離することに成功した。

③複数の音響モデルを利用した音声認識では、①、②の技術に音声認識システム（特定話者で150語を認識）を追加して、ロボットからの質問に3人が同時に答える実験を行った結果、3人の答えを認識することができた。

本内容の概要を1.3.12に示し、詳細を3.12に示す。

### (2) 長尾委員講演「セマンティック・トランスコーディングと Semantic Web」

情報をさらに高度に活用し再利用するために、コンテンツに意味的内容（アノテーション）を付与し、それを活用することによって、コンテンツをコンテンツ利用者の好み・環境に応じて加工（トランスコーディング）することができるようにした。

アノテーションとしては、言語的アノテーション（文書構造、文・句・品詞情報、語義属性）、およびマルチメディアアノテーション（アルチメディアに対するトランスクリプト、動画に対するシーン分割、音声・映像のオブジェクト記述、MPEG7情報、言語的アノテーション）をコンテンツから自動生成し、人がそれに修正・追加することができるようにした。

トランスコーディングでは、多言語翻訳、テキスト要約、音声化、画像変換、ビデオ要約、ビデオフォーマット変換等ができるようにした。

Semantic Webは、意味的情報の枠組みだけであるが、本システムは具体的に意味的情報の活用方法を示したことに意義がある。

本内容の概要を1.3.6に示し、詳細を3.6に示す。

### 【第3回（11月29日）】

#### (1) 田中教授講演「Meme Media and Meme Pools for Re-editing and Redistributing Intellectual Assets」

遺伝子が進化していくと同様に、社会において、知識を広め、進化させていくには、Meme（文化遺伝子：情報、知識、知恵という意味）を外在化、再編集する「Meme Media」、Memeを出版し、再利用し、再配布する「Meme Pool」、Memeをビジネス活動と結びつける「Meme Market」、およびMemeを組織化する「Meme Management」が必要である。

Meme MediaとしてIntelligentPad、IntelligentBoxを開発した。IntelligentPadは、Padと呼ぶドキュメントを貼り合わせ、スロットと呼ぶ外部インタフェースを介してドキュメントを機能合成し、複雑な合成ドキュメントを再編集できるシステムである。IntelligentBoxは、その三次元版である。Meme Poolとして、再利用し、再配布ができる

Piazza Server & Client を開発し、Meme Market として、課金の仕組みを考えた。

現在は、Meme が増えなかったこと、スロット情報などがわかりにくくて貼り合わせがむずかしかったことを踏まえて、Web コンテンツを Meme にし、再編集、再流通させるシステムを開発している。

## (2) 山名委員講演「WWWサーチエンジンの最新技術 — Google を例にとって —」

インターネットの全世界の Web ページ数は年々増加の一途を辿り、2002 年 5 月時点で 70 億ページと推定される。これに伴い、サーチエンジンのカバー率は 10%近くまで下がったが、サーチエンジン Google のおかげで 35%位まで持ち直している。

サーチエンジンの技術課題は、いかに効率よく Web ページを収集し、高速にインデックスを作成し、効果的に検索結果のランキングを行うかである。

収集は、重要ページから収集する戦略がとられ PageRank、BackLink の高い順等から判断して収集され、Web ページの Freshness を上げるように Web ページ更新を行い、分散処理で高速化をはかる工夫がなされている。

インデックス作成はソフトウェアパイプライン方式で高速処理されている。

ランキングは、有用な情報を含む Authority、Authority ページへのリンクをもつ Hub から求める HITS 法、ページ A からページ B にリンクがあるとき、B が A に一票投じていると考え、A の重要度によって一票の重みを変える PageRank 法 (Google の手法) がある。

Google は、2002 年 8 月現在、15000 台の PC で 30 億ページを収集し、1.5 億/日の検索要求をこなしている。最近の動向として、Google Search API を公開したり、Google ToolBar に分散コンピューティング機能を追加したり、モバイル端末に力を入れたりしている。

本内容の概要を 1.3.4 に示し、詳細を 3.4 に示す。

### 1.2.2 米国政府支援研究開発の調査

牧村幹事が、米国政府の情報技術研究開発政策の柱である「ネットワーキング及び情報技術研究開発」(NITRD : Networking and Information Technology Research and Productivity) の七つの研究領域の一つである「ソフトウェアの設計と生産性」(SDP : Software Design and Productivity) についての調査を行った。

その概要を以下に示し、詳細を 2 章に示す。

SDP は、1999 年の PITAC(IT に関する大統領諮問委員会)の提言により発足し、ソフトウェア生産性の劇的向上を目指し、基礎的な問題にチャレンジしようとするものである。

これに向けて、どのようなソフトウェア研究をすべきかについて、2001年4月にワークショップを開催して議論された。その結果、SDPが目指す方向性が以下のように示された。

### ①有用な抽象化を含むモデリング

- 特定領域のモデル記述
- 設計の意図の反映
- 多面的ソフトウェア
- 科学・工学に基づいた手法

### ②オープンな協調開発環境

- Webベースのオープンで協調できる開発プラットフォーム
- コンポーネント等のソフトウェアの宝庫

### ③既存ソフトウェアを生かす仕組み

- システムを徐々に現代化する仕組み
- 新しいものと古いものを統合する仕組み

この方向性で研究が進められているSDPの研究テーマを列挙する。

また、SDPが目指す方向を向いている注目すべき最近の技術、システムとして、Simulink、Ptolemy、MDA/UML2、BPEL4WS、アスペクト指向、軽い形式的枠組み、Semantic Web、オープンソースソフトプロセス、エクストリーム・プログラミング、デザイン・パターン、Webサービス、アプリケーションインテグレーション技術、CBS (COTS-Based Systems Initiative) を紹介する。

### 1.2.3 各委員による研究調査

人間主体の知的情報処理、ヒューマン・インタフェースに於ける各委員が専門とするそれぞれの分野について、各委員が以下のテーマについて研究調査活動を行った。

- 國藤委員：アウェアネス技術の動向
- 大須賀委員：次世代Webとエージェント
- 新田委員：IT技術の法への応用
- 山名委員：WWWサーチエンジンの最新技術
- 杉本委員：デジタルライブラリとメタデータ
- 長尾委員：セマンティック・トランスコーディング
- 加藤委員：多感覚情報のデータベース化と情報サービスへの応用
- 加藤委員：感性のモデル化技術とEコマースへの応用
- 宮田委員：トレイマージョンとビジュアルデータマイニング
- 平田委員：昨今の音楽情報処理における研究プロジェクトについて

児島委員：次世代音声認識の手法と実用化の動向

これらの研究調査の概要を 1.3 に、詳細を 3 章に示す。

#### 1.2.4 海外調査

加藤委員がフランスで開催された DEXA (International Conference on Database and Expert Systems Applications) 2002, DaWak(International Conference on Data Warehousing and Knowledge Discovery)2002, ECWeb(International Conference on electronic Commerce and Web Technologies)2002, EGOV(International Conference on Electronic Government)2002 に出席し、研究動向調査を行った。

その報告を付属資料に示す。

### 1.3 人間主体の知的情報技術の研究調査の概要

#### 1.3.1 アウェアネス技術の動向の概要

國藤 進 委員

分散環境でも対面環境と同様なアウェアネス（気付き）を提供するアウェアネス技術が CSCW(Computer Supported Collaborative Work)、グループウェア研究者に注目されている。本報告では、この分野の代表的国際会議である CSCW2002 でのアウェアネス技術およびその周辺技術の研究発表を中心に、この分野の研究開発動向を述べる。アウェアネス技術がモバイル技術、センサー技術と統合し、「いつでも、どこでも、誰でも」対面環境にいるのと同様な雰囲気、熱気、視線等をアウェアできる環境の構築研究の裾野が広がっていることを紹介する。

#### 1.3.2 次世代Webとエージェントの概要

大須賀 昭彦 委員

WebサービスやセマンティックWebなど、機械がWebコンテンツの意味を理解した上で処理することを可能にする、新しいWebの世界を目指した研究が進められている。次世代Webと呼ばれるこの新しいWebの世界では、人間の代理人となってWeb上のさまざまな仕事を処理するソフトウェアエージェントが重要な役割を果たす。そこで、この章では、エージェント技術研究の最新動向をサーベイした後で、WebサービスやセマンティックWebなどの次世代Web技術とエージェント技術との関わりについて触れる。また、エージェント技術をWebサービス分野に応用した具体事例として、Webサービス向けマッチメーカーを紹介する。

#### 1.3.3 IT技術の法への応用の概要

新田 克己 委員

IT技術の発達により、法律分野では教育方法、および、紛争解決方法に大きな変化がおきている。法学教育においては、メディアを使った法学教育の例として、遠隔地からも学習できる模擬法廷システムや、パソコンを利用し、学生と教官がインタラクティブに情報を交換しながら講義を進める実験的な教育システムを紹介する。紛争解決方法においては、オンラインの法律相談、調停、仲裁システムについての内外の研究動向や、実用化に向けての問題点を紹介する。

### 1.3.4 WWWサーチエンジンの最新技術の概要

山名 早人 委員

2002 年末時点で約 70 億ページと推定され、かつ、年率 100%の勢いで増加を続ける Web ページの中から自分が必要とする情報を見つけ出すためには、Google に代表される WWW サーチエンジンが必要不可欠となっている。しかし、Google が検索対象とする Web ページ数は、2002 年末時点で約 31 億ページであり、世界中の全 Web ページを対象に検索することはできない。また、WWW サーチエンジンに入力されるキーワードは平均 2 語であり、2 語という限られた情報を用いて、ユーザが欲する Web ページを見つけ出すことは、一般的に困難である。このような問題に対して、WWW サーチエンジンは、①膨大な Web ページの効率的な収集とインデックス化、②検索結果のランキング、という 2 つの重要な技術を利用している。本報告では、これらの技術について最新動向を報告する。

### 1.3.5 デジタルライブラリとメタデータ - 概要と課題

杉本 重雄 委員

本報告では、はじめに簡単にデジタルライブラリの研究開発を振り返った後、デジタルライブラリとメタデータに関する概観と将来への課題に関して述べる。最後に、コンテンツの保存や相互利用といったデジタルライブラリにおける幾つかの基本的課題について述べ、デジタルライブラリの位置付けを試みる。

デジタルライブラリの研究開発活動は、Digital Library Initiativesに代表される新しい情報技術指向のものと図書館を基礎にした活動に大別される。また、ネットワーク上での情報蓄積と流通の発展がデジタルライブラリに大きな影響を与えている。こうした活動を通じて、デジタルライブラリにおけるInteroperabilityの問題、デジタルコンテンツの長期保存の問題といった、これからのデジタルライブラリにとっての中心的な課題が浮き上がってきている。

メタデータはデジタルライブラリに限らずインターネット上でのコンテンツやサービスの流通に重要な役割を果たすことが認められている。また、デジタルライブラリサービスが、情報資源の検索やアクセス支援という付加価値であるという視点からは、メタデータは価値の付加を実現する上での中心的なものであるともいえる。インターネット上では目的に応じて多様なメタデータ規則が提案されている。実際にそれらを応用するには、それらを目的にあわせて適用することや、複数の規則を組み合わせることも求められる。こうした多様性を認める必要がある一方、メタデータ間のInteroperabilityを高め

## 1.3 人間主体の知的情報技術の研究調査の概要

ることが重要であることもみとめられている。草の根的に応用システムやサービスが作り上げられていくインターネットにおいて、異なる規則で作られたメタデータ間の相互利用性を高めることはこれからの重要な課題である。

インターネット上での情報発信や電子出版が一般化し、デジタルライブラリはすでに現実のサービスともなっている。一方、そこで実際に提供されている機能と、期待された機能との間にはまだギャップがあるといえる。また、コンテンツのデジタル化によって新たに出現してきた問題もある。デジタルコンテンツの収集と保存、コンテンツの細粒度化によるサービスの変化など、いろいろな問題が残されている。

### 1.3.6 セマンティック・トランスコーディング-より実用的な“Semantic Web”に向けて-の概要

長尾 確 委員

次世代の高度な知識共有のインフラであるSemantic Webに不可欠なものは、ユーザが自由にコンテンツを作成し、さらにその共有化を促進し、知識として再利用可能にするためにコンテンツを意味的に拡張するツールやプラットフォームである。われわれは、これまでセマンティック・トランスコーディングという枠組みにおいて、現在のWebの若干の延長線上に、Semantic Webと同等の機能を有する仕組みを研究してきた。それは、主に以下の3つのシステムから成っている。

1. 既存のWebコンテンツに対して、機械による内容理解を促進する補足情報(アノテーション)を付与するツール
2. アノテーションデータを管理・共有するためのサーバ
3. アノテーションデータを用いてコンテンツを動的にユーザに適合させるためのトランスコーディングプロキシ

これらのシステムを実現した経験に基づいて、Semantic Webの早期実現のために、近い将来に、われわれが何を、どのように行うべきか、に関する一つの提言を行う。また、動画コンテンツに対するアノテーションとそれを用いた応用についても述べる。

### 1.3.7 多感覚情報のデータベース化と情報サービスへの応用の概要

加藤 俊一 委員

多感覚を用いた情報通信・情報サービスの基盤技術として、多感覚を用いたヒューマンインタフェース、多感覚の情報のデータベース化、多感覚の情報の内容に基づく連想・検

索の技術の必要性が高まっている。

本稿の範囲で扱う「感性」とは、人間が実世界のさまざまな事物に接した時に、どのようにセンシングし、これを類型化し、どのような概念のイメージや主観的な評価と結びつけるかの過程をいうものとする。

感性をモデル化する場合、従来は、あるメディアのコンテンツを対象に、それぞれのコンテンツごとに、かなりの数の教示用データを利用者から得て、統計的学習を行っていた。従って、利用者（被験者）の心理的身体的負担は大きく、感性のモデル化技術を応用する上でのボトルネックとなっていた。また、人間は多感覚を利用して人とコミュニケーションを行っているが、1つの感覚を分析・モデル化し、ヒューマンインタフェースに応用している研究がほとんどであり、システムの開発は非常に個別的・限定的に行われている。このような現状をブレイクスルーして、感性情報技術（本来は応用分野に横断的な技術である）を汎用性・一般性のある技術として確立し、モノ作りやサービス、ヒューマンコミュニケーションなどの様々な場面で利用可能とする新しい枠組みが必要となってきたのである。

本稿では、現行のモデル化の枠組みを踏まえながら、感覚情報の幅を視覚・聴覚・触覚・味覚・嗅覚の五感に拡大するとともに、より精密な感性のモデル化を行う上で重要な視点を整理して検討する。

### 1.3.8 感性のモデル化技術とEコマースへの応用の概要

加藤 俊一 委員

検索技術の向上により、膨大な商品群の中から自分が購入したいと考えている商品を検索することは、ある程度可能になった。しかし、「何かいいものがないかな？」というようにウィンドウショッピングするような感覚で、自分が興味を持ちそうな商品を探し出すのは、非常に大変である。しかし、「自分の興味の対象になりそうな商品の情報が欲しい」という思いは、消費生活の上では非常に基本的な要求である。このような要求に対して、消費者の主観的な商品の選択基準をモデル化することができれば、消費者ごとに興味を持つ可能性のある情報を適切に Recommend することができる。そうすることで、消費者は自分の選択基準にあったより多くの情報に触れることができる。

本稿では、電子社会時代に、市民生活に最も近い BtoC (Business to Consumer) 型の電子商取引 (Eコマース) として、商品情報のレコメンデーションシステムを取り上げ、これを実現する上での技術的課題を考察する。

### 1.3.9 テレイマージョンとビジュアルデータマイニングの概要

宮田 一乗 委員

CG 技術の進歩とネットワーク速度の高速化にともない、地理的に離れた空間をコンピュータとネットワークで構成される仮想空間内で統合する、テレイマージョン (Tele-Immersion) の研究が現実味を帯びてきた。一方で、大量の画像データの中から意味のあるデータを発掘するビジュアルデータマイニングの研究は、バイオサイエンスや気象学などにおける多元解析のツールとして、今後重要度が増すと考えられる。本報告では、これらの技術動向に加えて、テレイマージョンの環境下で、協調しながらビジュアルデータマイニングする手法なども紹介する。

### 1.3.10 昨今の音楽情報処理における研究プロジェクトについての概要

平田 圭二 委員

本稿では、昨今の新しい音楽情報処理研究の流れを象徴するような国際的研究プロジェクトを 5 つ紹介する。

昨年の報告書では、新しい音楽情報処理研究の特徴は、(a) 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、(b) 応用システムをインターネット/Web 上に展開すること、(c) 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の 3 点であると述べた。これらを念頭に置いて、今年 3 つの音楽情報処理関連の国際会議に参加した。得た所感は以下の通り。

(a) の試みは地道に続けられている。論文はある程度の件数発表されているものの、決定的なブレイクスルーには至ってないようだ。現在、確立しつつあるサブ研究分野としては、メロディ分割、メロディ類似度、楽曲構造分析、声部分離、演奏表情分析/合成、音楽データベースのインデキシングなどがある。

(b) に関しては実際に多数の事例を見出すことができたが、アプリの動作環境としてモバイル端末を想定したものも発表されていた。例えば、ユーザの嗜好を反映した playlist (ラジオ局やレコード店の推薦盤/曲リストのこと) の自動作成が目をつけた。

(c) に関しては対象タスクが中粒度に変化しただけではないように感じた。音楽情報処理におけるタスクは大きく生成系 (制作系) と認識系に分類できるが、現在注目されているのは中粒度でかつ認識系のタスクである。信号レベルのパターン認識技術に限定し、(a) のような意味レベルまで踏みこまなくとも、有用な商用アプリケーションが組めるのではないかという実績と期待の反映である (これは、昨今の統計的自然言語処理の隆盛を連想

させる)。また、生成系（制作系）に取って代わらないことで、他分野の研究者の参入が容易になり、一般の人でも利用可能な音楽システムの研究開発であることがより明白にアピールできるため、研究リソースが集まってきている。

現在欧米では、上述した音楽情報処理の新しい研究の流れを象徴するような研究プロジェクトが、NSF、FP5（FP6）等から支援を受けて進行中である/あった。本報告では、その中から OMRAS、CUIDADO、WEDELMUSIC、MUSICNETWORKS、SALIERI/GUIDO の5つを取り上げ紹介する。

### 1.3.11 次世代音声認識の手法と実用化の動向の概要

児島 宏明 委員

現在の音声認識は、HMM（隠れマルコフモデル）と統計的言語モデルに基づく手法が主流となっており、その枠組み内では技術的にはほぼ確立されている。本報告では、これに代るような次世代の音声処理手法に関して、認識アルゴリズムを中心に概観し、位置づけと方向性をまとめる。また、このような音声処理技術の応用に関して、これまでに、認識対象語彙の拡大や認識精度の向上などの改良が進められてきたにもかかわらず、現在のところ、広く普及し利用されるには至っていない。ここでは、普及の鍵となる要因について、小型化と低価格化を中心に検討し、その事例を紹介する。

### 1.3.12 アクティブオーディションを利用したヒューマンロボットインタラクションの高度化の概要

中臺 一博 講師

近年、ヒューマノイドを中心にロボットが注目されている。この延長線上には、将来、ロボットが人間社会で人と共生し、ソーシャルインタラクションを通じて、パートナーとしての役割を果たすことへの期待があろう。しかし、現状のロボットは、そうしたソーシャルインタラクションを可能にするほどロバストな知覚処理は難しく、その中でも特にロボットの聴覚処理には多くの課題が残されている。

そこで、本報告では、ロボット聴覚機能を実現し、自然なヒューマンロボットインタラクションを実現するため、アクティブな動作を聴覚と統合し聴覚を向上するアクティブオーディションを提案する。また、視聴覚情報の統合による知覚向上、一般的な音（混合音）の理解も同時にロボット聴覚には本質的であると捉え、これらを考慮したロボット聴覚システムの提案および工学的な実装を行う。

### 1.3 人間主体の知的情報技術の研究調査の概要

構築システムは、ロボット動作時のノイズをキャンセルすることで、アクティブな動作の利用を可能とする。また、視聴覚を統合することにより、複数の話者定位・追跡を行う機能と注意を向けた方向の音源を実時間で抽出するアクティブ方向通過型フィルタによって音源分離・分離音の音声認識を行う機能を有している。これらを通じて、ヒューマン・ロボットインタフェースとしての有効性を示す。



## 2. 米国政府支援 SDP(Software Design and Productivity)研究開発の動向

牧村 信之 幹事

1999年2月に、米国のPITAC（ITに関する大統領諮問委員会）は、ITに関する米国政府支援の研究開発に対し、長期的基礎研究とソフトウェア研究の重要性を訴える提言を行った。この提言を受けて、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及び情報技術研究開発（NITRD：Networking and Information Technology Research and Productivity）に「ソフトウェアの設計と生産性」(SDP：Software Design and Productivity)という研究領域が追加された。本章では、このSDPが何を目指しているかについて、2001年4月アーリントンで開かれたワークショップの内容から紹介し、2003年度のSDPの活動計画として、米国はどのような研究に取り組んでいるかを紹介する。また、SDPが目指すものに関連した注目すべき最近の技術、システムについて紹介する。

## 2.1 SDP 発足の経緯と PITAC の提言

1999年2月に、PITAC（ITに関する大統領諮問委員会）は、ITに関する政府支援の研究開発に対し、次の示すように、長期的基礎研究とソフトウェア研究の重要性を訴える提言[1]を行った。

- 1992年以来、IT関連の生産高はGDPの1/3を占め、数百万ドルの雇用機会を提供してきた。
- しかし、現状の政府の投資は、長期的なハイリスクな研究開発よりも、各省庁のミッションを追求する短期的な課題に偏りすぎている。
- このままであると、過去20-30年に渡って続いてきたIT革命へのアイデア供給の流れは途切れてしまう。
- 従って、2000-2004年の5年間に下記の研究開発テーマに4,743Mドルを長期的基礎研究開発に追加投資すべきである。
  - ソフトウェア
  - 大規模な情報化のためのインフラストラクチャ
  - 最高レベルのコンピューティング
  - 社会経済的影響
  - 連邦政府の情報技術研究の管理と実現

このPITACの提言を受け、政府は、2000年には「21世紀に向けての情報技術」IT<sup>2</sup>計画として366Mドルの研究開発投資を追加し、2001年にはこれまでのCIC計画（2000年の予算は1,089Mドル）と統合して、ITR&D計画として、1,928Mドルと大幅に増額している。そして、現在は、ITR&D計画はNITR&D計画と名前を変更し、予算規模は1,830Mドル(2002年)となっている。

PITACは、追加投資すべきとしたソフトウェアについては、以下のような現状、課題、であるとの調査結果を発表し、

- ソフトウェアに対する要求が、国家的なソフトウェアの生産能力を上回っている。
- 脆弱なソフトウェアに国家が依存している。
- 信頼性と安全性に優れたソフトウェアを構築するためのテクノロジーが不十分である。
- ソフトウェア・システムは、急速な勢いで多様化と洗練化が進んでいる。
- 一般的な市民にとって、ソフトウェアへの依存度はしだいに高まりつつある。
- ソフトウェアの基礎研究に向けた国家の投資が不足している。

その対策として、次のことを提案している。

- ソフトウェアの基礎研究を最優先課題とする。

## 2.1 SDP 発足の経緯と PITAC の提言

- ソフトウェアの開発手法、およびコンポーネント・テクノロジーに関する基礎研究に向けた投資を強化する。
- ヒューマン/コンピュータ・インタフェースと相互作用に関する基礎研究を支援する。
- 情報管理の方法に関する基礎研究を強化する。
- 主な情報テクノロジー関連の研究構想の中に、ソフトウェアの研究を主要な課題として位置付ける。

## 2.2 NSF SDP ワークショップ

PITAC の提案により、SDP の研究開発予算はついたものの、PITAC の提案は、具体的にどのような研究を行えばよいか必ずしも明らかではない。そこで、NSF の Frank Anger が座長となって、後述する米国の著名な研究者（他に省庁の研究機関担当者も出席）が、2001 年 4 月 18~19 日にアーリントンに集まって開催されたのが、本ワークショップ[2]である。

### 2.2.1 SDP ワークショップの狙い

本ワークショップでは、Frank Anger がイントロとして、ソフトウェアの現状は、以下のようにあり、

- 全米でソフトウェア技術者 200 万人
- サポートスタッフ 35 万人
- まだまだ空きポスト多数

従って、以下のことを達成しなければならない。

- ソフトウェア生産性の劇的向上
- 高いソフトウェア品質

と説明し、前述した PITAC の提言を説明した後、本ワークショップでは、以下の点をあきらかにするのがねらいであるとした。

- 解決すべき基礎的な問題は何か
- 真の障壁と挑戦とは何か
- どの方向がもっとも有効か
- 何が理想的な成果か
- どのように研究を進めるべきか

この「どのようなソフトウェア研究をすればよいか」の議論を基にして、SDP (Software Design and Productivity) の研究計画を作ろうということで、以下の 4 つのグループに分かれて検討がなされた。

- ソフトウェア研究の将来
- 新ソフトウェア開発パラダイム
- 実世界ソフトウェア
- ネットワーク中心のシステム

これら四つのグループの参加メンバを表 1 に示す。

表 1 SDP ワークショップ参加メンバ

ソフトウェア研究の将来	新ソフトウェア開発パラダイム
Barry Boehm, USC Benjamin Pierce, Univ. of Pennsylvania Doris Carver, Louisiana State Univ. Shankar Sastry, UC Berkeley Bonnie John, Carnegie Mellon Univ. Kevin Sullivan, Univ. of Virginia William Mark, SRI International	Grady Booch, Rational Software Ralph Johnson, Univ. of Illinois Gregor Kiczales, Univ. of British Columbia Charles Simonyi, Microsoft John Vlissides, IBM
実世界ソフトウェア	ネットワーク中心のシステム
Don Winter, Boeing Phantom Works Martin Feather, Jet Propulsion lab. Gabor Karsai, Vanderbilt Univ./ISIS Patrick Lardieri, Lockheed Martin Cleve Moler, MathWorks, Inc. Edward Lee, UC Berkeley	Ian Foster, Argonne National lab. Doug Lea, State Univ. of New York at Oswego Adam Porter, Univ. of Maryland Rick Schantz, BBN Technologies Jim Waldo, Sun Microsystems, Inc

## 2.2.2 各検討グループからの提言

### 2.2.2.1 ソフトウェア研究の将来

現状は、

- 現在のインフラは、低い信頼性、安全性、可用性で、まったく不適切である。
- 大規模ソフトウェアプロジェクトの半数は失敗している。
- 応用ソフトウェア同士が情報をやりとりできない。
- 過去のソフトウェア資産が 1 兆ドルもある。

との仮説に立って、以下の目標を立てている。

- 各領域の専門家がソフトウェアを使いこなせるようにしよう。
- 既製品ソフト業界の共通基盤となるアーキテクチャを開発しよう。
- ソフトウェア業界に対し、鍵となる概念・技術を提供しよう。
- プログラムの意図を機械が理解できる形で記録する方法を開発しよう。
- ソフトウェア開発の現状をしっかりと認識しよう。

この目標実現に向けて、特に注目すべき領域は、以下であり、

- ネットワーク埋め込み型
- 対人インタラクション

- グループインタラクションの計算機支援
- 領域専門家のための環境：教員、エンジニア、金融、科学者

その領域で基礎となる科学は、以下であり、

- ソフトウェア物理：並行処理のモデルなど
- ソフトウェア経済学：柔軟性の価値を定量化など
- ソフトウェア行動科学：人間の認知・記憶に関する知見をソフトウェア設計に応用したり、ソフトウェアと人間の役割分担を明らかにしたりすることなど

そのキーポイントは、以下であると提言している。

- ソフトウェア開発のプレイグラウンドを変えよう。
  - 「プログラム」の意味の変更：
 

Ptolemy（同期データフロー、状態遷移マシン、ランデブー方式同期メッセージング、連続的な時間、離散イベントといったコンピュテーションモデルのビジュアル化イメージでシステムを作成するアーキテクチャデザイン言語）

Simulink（ブロックダイアグラムを作成することによって、システムのシミュレーションや性能の評価、デザインの改善が行える Mathworks 社のモデリングツール）
  - 「プログラマ」の意味の変更:practitioners（開業医などの専門家）
- 以下のキーとなる技術を進展させよう。
  - 軽い形式的枠組み
  - 並行性・耐故障性・安全性などのモデル
  - 「価値」に基礎を置いたデザイン
  - 分散オープンソーステスト方式
- ソフトウェア経済学を変えよう
- ソフトウェアでハードウェア を定義しよう。  
(ハードウェアがソフトウェアをではなく)

#### 2.2.2.2 新ソフトウェア開発パラダイム

新ソフトウェア開発パラダイムでは、現状は、以下のような要因によって、複雑さが指数関数的に増加している。

- 組込型、モバイル、浸透性、大域性
- 長寿命、複数バージョン混在
- 共同開発

- 利用者ごとのカスタマイズ

この難問題を以下のような革新的概念によって、一挙に解決することが必要である。

- 国立ソフトウェア考古学センター構想
- 多面的ソフトウェア
- 協調開発環境

### **[国立ソフトウェア考古学センター構想]**

ソフトウェアの古典を蒐集および保存し、それをリバースエンジニアリングによって、その古典の芸術的本質を抽出し、Web で公開する。これにより、参考となるソフトウェアアーキテクチャの宝庫として、将来の研究の基礎として、特許取得のための事前芸術研究として、役立て、芸術的表現を奨励することをねらう。

### **[多面的ソフトウェア]**

高級言語、オブジェクト指向プログラミング、デザイン・パターン、特定領域用言語といったこれまでの歩みは、ある側面をより理解しやすくした、あるいは、ある側面をうまく表現したというものにすぎない。本来、ソフトウェアは、様々な異なる意図を表現するものであり、新たなシステムは、共存する異なる側面が互いに影響しあって生まれるものである。このような多面的ソフトウェアについて、従来は認識不足であったが、デザイン・パターン、協調開発環境、アスペクト指向プログラミング、特定領域用言語といった研究分野では、一部、多面的ソフトウェアの考えが取り入れられ始めている。今後、多面的ソフトウェアとしてすべきことは以下である。

- どの側面が肝要か  
性能、価格、使い勝手、過去からの連続性...
- 各側面をどう表現するか
- 諸側面をどう変形・統合しシステムとするか
- 特定領域用言語のためのツール整備の加速

### **[協調開発環境]**

協調開発環境とは、ソフトウェア開発のプラットフォームとして Web を仮想的な議論の場にするものであり、研究すべきことは以下である。

- ソフトウェア開発の社会学
- ソフトウェアプロセスの利用
- 考古学センターや多面的ソフトウェアを利用するプラットフォームの創造

これらの革新的概念の研究計画例には、以下のようなものがある。

- 側面中間形式（特定領域用言語のための汎用ツール、Weavers?）
- 対応代数（Correspondence Calculus）

- デモプロジェクト (Demonstration Projects)
- 領域回復 (Domain Recovery)
- ソフトウェア保存 (Preservation)
- ソフトウェア開発民俗学 (Ethnography)

### 2.2.2.3 実世界ソフトウェア

今日のソフトウェアは、偶然的複雑性、予測不能性、合成不能性、脆弱性、実世界との相互作用に向かないといった問題点を抱えている。

また、ソフトウェアは「コードの集まり」という考え方にハマッており、大規模にするには人力を注ぎ込む、システム全体像の理解が喪失している、仕様とその実現が分離している、といった事態を招いている。

10年後には、どこでもコンピュータだらけになって、これまでの世界とは全く変わった世界になるというのに、今日のようなソフトウェアの作り方をしていたら、世界は危険極まわりの場所となってしまう。それを避けたければ、技術の実施を遅らせねばならず、その場合には、自動車は、コンピュータ制御とはならず、交通渋滞にはまり込んでしまうだろう。

これを解決するには、モデリング言語、システムの合成、変換、過去のソフトウェア遺産の扱い、といった課題に取り組まねばならない。

#### [モデリング言語]

ソフトウェアの問題点は、「プログラムコード」、あるいは「仕様」にあるのではなく、「モデル」あるいは「設計」にある。すなわち、複雑な機能の実現、設計の理解、質問の定式化、振舞いの予測、といったことを行う際に、それを表現するモデリング言語を我々が持っていないことである。

これを解決するため、以下のような研究を行うべきである。

- システムのモデリング「言語」
- 有用な抽象化を見つけること
- 計算的システム理論
- 組合わせて使える抽象化
- 時間・並行・能力などの表現

#### [システムの合成]

システムを組合せて作る組織的方法、例えば、要素の枠組み、組合せの意味論、その場での (on-the-fly) 合成、既成のソフトウェア要素との統合、といったことが欠如している。

これを解決するため、以下のような研究を行うべきである。

- 意味の枠組みと理論
- 方法論 (methods) とツール
- 実験台と挑戦的問題
- 参照実装 (reference implementation)
- アーキテクチャ枠組みの定義
- 分散・分割の戦略
- 粒度とモジュラー性の制御戦略

### [変換]

異なる抽象間の変換についての理論、例えば、異なる抽象の間関係、生成器(変換器)、複数の視点を許す抽象化(multi view abstraction)、物理的環境の抽象化、高信頼性ソフトウェアとシステム(HCSS: High Confidence Software and Systems)との関係: 証明つきの変換、といったことが欠如している。

これを解決するため、以下のような研究を行うべきである。

- オープンな生成器インフラ(方法論、ライブラリ)
- 生成器の理論
- Methods for correct by construction transformers (?意味不明?)
- 変換器のパラメタ化

### [過去のソフトウェア遺産の扱い]

過去の遺産を扱う方法論、例えば、システムを徐々に現代化する方法、新しいものを古いものと統合する方法論、といったことが欠如している。

これを解決するため、以下のような研究を行うべきである。

- 遺産コードの部品化
- 遺産からの抽象化の抽出
- 漸次現代化、リバースエンジニアリング
- 再設計を安価に ⇔ ソフトウェア遺産からの進化

#### 2.2.2.4 ネットワーク中心のシステム

今後のシステムは、以下のような状況に進展すると予想される。

- すべてがコンピュータ
- すべてがネットワークコンピュータ
- すべてが相互依存する可能性
- 実世界と接続
- ますますヘテロに

- 接続はますます多様で不安定に

その状況下で、以下のような応用システムが登場してくるであろう。

- 交通制御: 自動車数千台からセンサーデータ
- 戦域管理: 敵対環境で多様粒度の調整・任務
- サプライチェーン管理
- 科学データの共同解析

数千ユーザ、数千資源に対し、ソフトリアルタイムで問合せ最適化の要請

- 家庭内電力管理

このような状況、応用システムにおいては、以下のような技術的課題が生じる。

- 資源 (QoS、時間・相互依存・エネルギー消費・大きさ) を意識したプログラミング
  - 工学的トレードオフを考慮する計算モデルの欠如
  - 複雑なシステムでは端点間の特性が不明
- 動的な資源の振舞い: 故障・不均一性 等
  - 計算モデルや端点間特性の欠如
- ソフトウェア遺産
  - 設計時に想定しなかった利用形態への対応
- 危機・信用・制御の管理
  - ポリシー・保障・管理のドメイン
  - 極度に動的なシステムの安全性確保と確認手法
  - プライバシ
- スケールの問題
  - 構成要素数、構成要素の大きさ
  - 単一の計算に関する要素数
  - ネットワーク中心システムは長時間無停止動作

これらの問題を解決するには、以下のような研究方向とアプローチが必要である。

- 「問題」を扱うプログラミングモデル
  - 工学的トレードオフ
  - 大規模システムの運用
  - 動的な資源の振舞い
  - 危機・信頼・安全の管理
- 「問題」の基本機構
  - 資源のトレードオフ: QoS、実時間性、等
  - 適応的振舞い

- さまざまな次元への拡大
- 分散管理
- 「問題」の全体を扱える、多レベルの資源管理
- ネットワーク中心の(小規模な)要素を、動的に大規模システムに合成する技術、相互運用性

この他、次のような研究も必要である。

- 新たな計算モデルを使うためのツール
- 高度に複雑化したシステムの検証・シミュレーション・理解
- 大規模分散応用システムの(自動)構成・管理

これらの研究によって、以下のような効果が期待される。

- 社会的には
  - 現状では作れないものを作れるように
  - ネットワーク化システムの信頼性・制御性向上  
工学的設計によって実現 ⇔ デバッグ
  - ソフトウェア開発コストの抑制
- 教育
  - 新たなソフトウェアの作成・システムの設計にあたる人材の供給

### 2.2.3 SDP ワークショップ提案のまとめ

4つのグループからのそれぞれの提案は、別々の方向を向いているのではなく、ほぼ同一方向を向いており、概ね、以下の提案にまとめられる。

- 有用な抽象化を含んだモデリング
  - モデリングは、対象とする特定領域のモデルを記述できるものとし、その特定領域の専門家が使いこなせる。
  - モデリングは、資源のトレードオフ・動的管理、危機・信頼・安全の管理といった、設計の意図を表現できる。
  - ある特定の側面からのモデリングを組み合わせ・合成・変換・統合して、様々な側面からのモデリングができ、システムの複雑さを軽減できる。
  - モデリングは、科学的、工学的設計に基づいた軽い形式的枠組みを持ち、これからコードの自動生成ができると共に適合性の検証も行える。
- オープンな協調開発環境
  - Web をベースとしたオープンで協調して開発できるプラットフォーム。
  - この環境には、芸術的ソフトウェアが蒐集され、ソフトウェアの宝庫とな

る。

- 柔軟性の価値なども定量化され、ソフトウェアの経済的社会的評価ができる。

- 既存ソフトウェア資産を生かす仕組み

- システムを徐々に現代化する仕組み
- 新しいものと古いものを統合する仕組み

## 2.3 2003 年度 SDP の活動計画

ワークショップからの提案を受けて、米国政府は、通称Blue Book 2003と呼ばれる「ネットワーク及び情報技術研究開発 — 2003年度大統領予算教書補足資料」[3]において、SDPに対する2003年度の活動計画を以下のように示している。

### 2.3.1 取り組み方針

2002年に、半導体業界は地球上にいる人間の数より多いコンピュータ用チップを生産すると予測されている。これらチップの約1%だけが認識可能なコンピュータに使われ、ほとんどは他のタイプの機器に組み込まれる。すなわち、小さなものでは携帯電話、ポケットベル、パーソナルデジタルアシスタント(PDA)やリモコンであり、大きなものとしては車、飛行機、製造機械、兵器システム、緊急警報ネットワークのような複雑な物理的システムなどである。組み込み型コンピューティング技術にとって最も有望な新しい方向は、ほとんどリアルタイムで重要なデータを収集し、転送することのできるネットワーク化されたマイクロセンサシステムである。このようなセンサ網は、環境と危険モニタリング、科学的農業、輸送管理、戦闘場面における偵察と早期警戒システムを含む、広範囲のアプリケーションに使われるようになっている。

IT研究者が、国家のコンピューティング及び情報通信インフラストラクチャ上で動いている現世代のソフトウェアの重大な技術的脆弱さに取り組もうとしたとき、IT研究者は、組み込み型技術における初期の変革において、独特の新しい難問に直面した。今日のソフトウェアは、その開発とメンテナンスに巨大な金額が費やされているにもかかわらず、必要とされているほど相互運用性がなく、拡張性もなく、また、コスト効率に優れているわけでもない。

この問題を解くかぎは、これまではプログラミング“ウイザード”という寂しい技術によって行われてきたソフトウェア設計と開発を、質の高い結果を得るために広く受け入れられる原理、方法、最良の経験で支配された形式的枠組みの科学と工学の学問分野に変換することである。技術者は、技術的要求を満たした詳細な設計図、関連する基本原理や関連材料に対する最先端の訓練で養われた知識、長持ちする品質保証された構成部品を作るためのコスト効率の良い開発プロセスを持たずに、つり橋を架けるような夢を見ることはない。わが国のインフラストラクチャに根を下ろす重要な基盤として、ソフトウェアは少なくとも科学的に設計されたものであるべきである。

NITRDの研究のひとつは、ソフトウェア開発の新しい科学のために、モデルを作り出し、試験し、評価することに重点を置いている。目標は、今日の特異的で少数の人のみが

理解できるコードから脱却し、合理的でモジュラリティがよく、実証的で再使用できる、安定した工学的設計に移行することである。科学的な原理と手法は、開発者やユーザが、自動化された技術を使って、ソフトウェア製品を設計し、構築し、試験し、厳しい使用テストをすることを可能にする。ここで自動化された技術は、ソフトウェア製品が利用される前に、それらの動作を確認し、大規模ソフトウェアの何百万のもの手作りのコード中に潜む、今すぐ簡単に見つけられない弱点を正確に指摘することができるようにする。設計と実装の局面を自動化し、実践的なテストベッドを使うことは、プログラミング段階を合理化するだけでなく、高価なデバッグ工程の効率を上げ、改良することによって、現在の極めて高い開発コストを抑制することになる。

最終的には、より高品質のより多くの製品を作ることを可能にするだけでなく、実質上開発費を下げ、維持を容易にすることにより、ソフトウェアの「生産性」を劇的に向上させる開発方法論を得ることが目的である。

ソフトウェアのための科学的基盤は、組み込み型システムにとって極めて重大なものであり、大規模な物理的システムの内側奥深くにある小さなコンピュータ・コンポーネントは、極めて多くのコンピューティング及びノン・コンピューティングタスクをリアルタイムにサポートするための機能を持っている。ちょうど、自分がどこにいるか、そして、または何をしようとも、我々の心臓と肺は命令されなくとも機能するように、組み込み型プロセッサはその仕事を自動的に絶え間なく行うことができなければいけない。ちょうど我々の命は、新しい細胞が生まれて古い細胞が死ぬことにほとんど影響されないように、組み込み型システムは、いくつかの個々の組み込み型プロセッサが損傷を受け、破壊され、付け加えられ、あるいはアップグレードされても、全体としてのシステムが自動的に機能し続けている間、動作し続けることを期待されている。

2003 年度に、NITRD 参加機関は、ソフトウェアの品質を向上させ、その費用を全面的に削減するための基礎研究を支援する。この先進的技術は不可欠の駆動力となり、我々は、これにますます依存度を深めることになって、大規模及び小規模アプリケーションの両方において、社会により、大きな利益をもたらすことになる。

### 2.3.2 2003 年度の SDP の主要な研究課題

- ・ 言語とコンパイラを含むソフトウェアとシステム科学；  
構築方法論；  
分散型で拡張性のあるシステムとネットワークの設計基盤；
- ・ 効率的で、信頼性の高いソフトウェア・コンポーネントとその統合プロセスを含む自動化されたエンジニアリング；

- 統合ソフトウェアとシステム開発のための(仕様、分析、検証などの) 方法論；
- 並列実行するネットワーク化されたアプリケーションの相互運用性；
- ドメイン特有の開発環境の速やかな構築とカスタマイゼーションが可能なツール環境；
- ・ 組み込み型アプリケーションとシステム開発プロジェクトのための技術に関するパイロットアプリケーションと実践的研究；

### 2.3.3 2003 年度における各機関の代表的活動

#### NSF：

- ・ 実践的なソフトウェア工学研究
- ・ コンポーネントベースのソフトウェアに於ける絶えず生じる変化に対する管理
- ・ ソフトウェアを発展させるプロファイルとパターン
- ・ リスクドリブンからバリュードリブンの開発モデルへと変更するための戦略的ソフトウェア設計

#### DARPA：

- ・ 組み込み型システムのためのプログラム作成(Program Composition for Embedded Systems：PCES)

PCES は、組み込み型システムのプログラミングのための新しい戦略を開発しており、技術的労力を抑えてより堅牢なコードを作ることを目的としている。この手法にはコアソフトウェアの設計が絡んでおり、並行操作、センサ、時間的制約のあるアクチュエータの同期、安全で高効率のキャッシュ、レジスタ、メモリ管理などの組み込み型システム全体に必要とされる横断的機能を含んでいる。目標は、ダイナミックな構成とテストのための自動化されたツールを用いた再利用可能なソフトウェアスイートである。

- ・ 組み込み型ソフトウェアのモデルベースの統合
- ・ ソフトウェアによる制御、センサ網あるいはシステム網で構築される高度に複雑なシステムでのダイナミックな自己構成と保証

#### NIH：

- ・ バイオメディカル・コンピューティング・アプリケーションを支援するソフトウェア研究

#### NSF/NASA：

- ・ 高信頼コンピューティングとコミュニケーションシステムの研究に関する共同研究プログラム

これは、現実世界のハードウェアとソフトウェア人工物に関する研究資金援助を評価

するための NASA の新しいテストベッド施設を使って、信頼性があり、コスト効率の良いソフトウェアをベースにしたシステムを設計し、試験し、実装し、発展させ、保証することをねらいとするプロジェクトである。

**NASA :**

- ・組み込み型あるいはロボットデバイスのための技術とツールを含む自動化ソフトウェア工学手法
- ・ベイジアン法を用いた仕様、ソフトウェアの実践評価

**DOE/NNSA :**

- ・全ての高度シミュレーションとコンピューティング (Advanced Simulation and Computing : ASCI) ハイエンドプラットフォームのための、共通のソフトウェア開発/実行環境の構築

これは、I/O、ストレージ、視覚化に対するニーズだけではなく、堅牢さと拡張性のためのエンドツーエンド ASCI アプリケーションのニーズを支援するものである。

**NIST :**

- ・消費者団体と使い勝手に関するデータを共有するための共通の報告書形式の開発
- ・自動化された、知識ベースの手法を用いたソフトウェア品質の決定
- ・産業界のパートナーと製造 B2B の相互運用性テストベッドを共有する計画

**NOAA :**

- ・コンポーネントを基本にした地球物理学環境のモジュラー研究モデルの開発

**ODDR&E :**

- ・組み込み型システムをチェックするソフトウェアモデルに関する大学における基礎研究
- ・リアルタイムの無停止ネットワークプロトコル

### 2.3.4 2003 年度の予算

表 2 に示すように、SDP に対する予算は 1.821 億ドルであり、NITRD 予算の約 1 割を占めている。

表 2 NITRD 関係機関の個別研究分野(Program Component Area)別の予算

2002年度予算 (単位: 100万ドル)

Agency	HECI&A	HECR&D	HCI&IM	LSN	SDP	HCSS	SEW	合計
NSF	225.1	66.4	132.6	102.8	44.4	45.8	58.8	676
NIH	79.5	7.7	82.0	105.6	5.6	3.6	11.4	295
DARPA	17.4	63.9	35.8	35.0	65.6			218
NASA	36.1	26.0	27.8	14.4	22.4	47.1	7.0	181
DOE	97.2	29.5	16.4	27.8			3.5	174
NSA		41.6		1.3		32.7		76
NIST	3.5		3.2	6.2	7.5	2.0		22
NOAA	13.8	1.8	0.5	2.8	1.8			21
ODDR&E		2.0	1.8	5.7	1.6	1.0		12
AHRQ			8.0	6.0				14
EPA	1.8							2
小計*	474.4	238.9	308.1	307.6	148.9	132.2	80.7	1,691
DOE/NNSA	42.1	33.5		25.9	33.2		4.2	139
合計*	516.5	272.4	308.1	333.5	182.1	132.2	84.9	1,830

HECI&amp;A: ハイエンド・コンピューティング基盤とアプリケーション

HECR&amp;D: ハイエンド・コンピューティング研究開発

HCI&amp;IM: ヒューマン・コンピュータ・インターフェースと情報管理

LSN: 大規模ネットワーク

SDP: ソフトウェアの設計と生産性

HCSS: 高信頼ソフトとシステム

SEW: 社会、経済、および労働力の面から見た IT 労働力開発

\*小計と合計に 2003 年度大統領予算教書と若干の相違があるのは、集計における個別研究分野の多少の変更と端数処理の違いによる

## 2.4 SDP の目指すものに関連した技術、システムの動向

SDP の将来の方向性として、前述のように、モデリング、形式的枠組み、オープンな開発環境、既存資産の活用を提言し、キーワードを並べているが、まだ、具体的なイメージがつかめない読者も多いことであろう。SDP に参加したメンバも将来構想であるから、具体的イメージを持っているわけでもなく、各メンバの将来イメージもそれぞれ異なっているので無理からぬ話である。しかし、参加メンバは、予算獲得の思惑があり、各メンバが現在研究していること、あるいは関心があることの延長線上に将来イメージがあるであろう。そこで、SDP が目指す方向に関する、以下に述べる現在の研究動向を把握し、それから、その延長線上の将来イメージがどんなものであるかを把握してもらいたい。

### 2.4.1 モデリング、特定領域記述言語(DSL: Domain Specific Language)

IT 革命が進行し、IT と他分野との融合が加速していけば、融合分野の専門知識が希薄な IT プログラマは対応が難しくなってくる。従って、IT プログラマがプログラミングするのではなく、各特定領域の専門家がプログラムできるものでなくてはならない。特定領域には、多くの場合、建築設計図、電気回路図、損益対照表といった特定領域固有のモデリングドキュメントが既に存在する。このようなモデリングドキュメントをベースに IT システムを記述しようとするアプローチが特定領域記述言語である。

これは、従来、プログラミングとは、インプリメントすることを設計してきたが、モデリングを設計することによって行おうとするアプローチであり、このモデリングが特定領域固有のモデリングドキュメントに他ならないのである。しかし、ただ単にモデリングドキュメントで終わらずに、実行できるものではなくてはならない。従って、モデリングドキュメントを「軽い形式的枠組み」にしたり、「メタモデリング言語」でモデリングドキュメントの意味を規定したりして、あるいは、暗黙的な知識を使ってシステムが自動的に、実行コードの生成をサポートにする。

特定領域記述言語の代表的なのは、スプレッドシートで表に関する処理はほぼ何でもこなすことができる。表以外で注目されるシステムに、Simulink、Ptolemy、OMG で検討が始まった MDA、Web サービスをベースにビジネスフローを記述する BPEL4WS がある。

#### 2.4.1.1 Simulink

Simulink[4]は、MathWorks 社の製品で、DSP の設計、コミュニケーションシステムの設計、その他シミュレーションシステム向けのダイアグラムブロックの開発ツールである。その特徴は、150 以上のブロックを含むライブラリやモデル（出来合いブロックダイアグ

## 2.4 SDP の目指すものに関連した技術、システムの動向

ラム) が準備されており、これらを組み合わせてブロックダイアグラムを作成、編集できる。

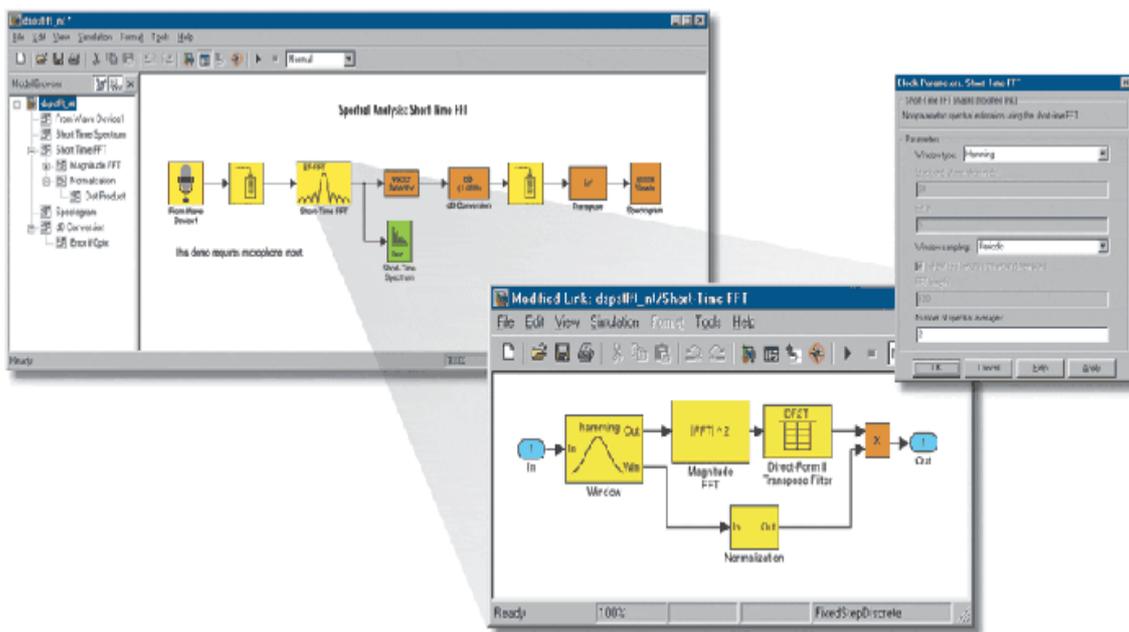


図 1 Simulink FFT の例

図 1 (<http://www.mathworks.com/products/simulink/description1.jsp>) は、短時間の高速フーリエ変換ブロック (FFT) の例であり、これは DSP ブロックを用いて作成されている。FFT ブロックに対するパラメタは右のダイアログボックスで入力される。ユーザがサブシステムの詳細を見たいければ、右下に示すその詳細ブロックダイアグラムを表示することができる。

Simulink では、ユーザ作成のカスタムブロックをシステムに組み込む、あるいは C 言語などのコードをブロックダイアグラムに組み込むことによって、システムを拡張することができる。また、作成したブロックダイアグラムを、インタラクティブに実行することができ、結果を即座に表示したり、グラフィカルデバッガでデバッグしたりすることができる。

### 2.4.1.2 Ptolemy

Ptolemy[5]は、カルフォルニア大バークレー校のプロジェクトであり、組み込みシステムのような複雑なシステムのモデリングと設計を支援するシステムの構築を目指している。その狙いは、以下の実現にある。

- ① 組み込みシステムの開発におけるモデリング、設計の方法論の確立
- ② 適切なコンピュテーションモデルの選択によるシステム設計品質の向上
- ③ 多数のコンピュテーションモデルを実行するモデルの生成と相互運用
- ④ コンポーネントベースの設計

Ptolemy プロジェクトでは、ビジュアルエディタである **Vergil** を用いてモデリング、設計を行う。この **Vergil** の特徴は、単一のコンピュテーションモデルではなく、複数のコンピュテーションモデルの中から選択できることである。**Vergil** では、継承構造、手続きインターフェースを主体とするオブジェクト指向ベースのモデリングではなく、並列性やコミュニケーションモデルを主体としたアクタベースとした、以下に示す9つのコンピュテーションモデルが用意されている。

Communicating Sequential Processes - CSP

Continuous Time - CT

Discrete-Events - DE

Distributed Discrete Events - DDE

Discrete Time - DT

Finite-State Machines - FSM

Process Networks - PN

Synchronous Dataflow – SDF

Synchronous/Reactive - SR

2つ目の特徴は、上記のモデルがグラフィカルに表現されることである。図 2 (<http://ptolemy.eecs.berkeley.edu/publications/papers/01/overview/overview.pdf>) は、Discrete Event モデルの例で、バス停に乗客がポアソン過程で到着し、バスが一定間隔で到着したときと、ポアソン過程でランダムに到着したときのシミュレーションの比較を示す。

図 3 (<http://ptolemy.eecs.berkeley.edu/publications/papers/01/overview/overview.pdf>) は、Finite-State Machine の例を示す。

3つ目の特徴は、上記のグラフィカルなモデルは、コンピュテーションモデルに基づいてため、意味の形式化がなされ、実行可能で、モデルが意図するものであれば、信頼性におけるプログラムとなる。

## 2.4 SDP の目指すものに関連した技術、システムの動向

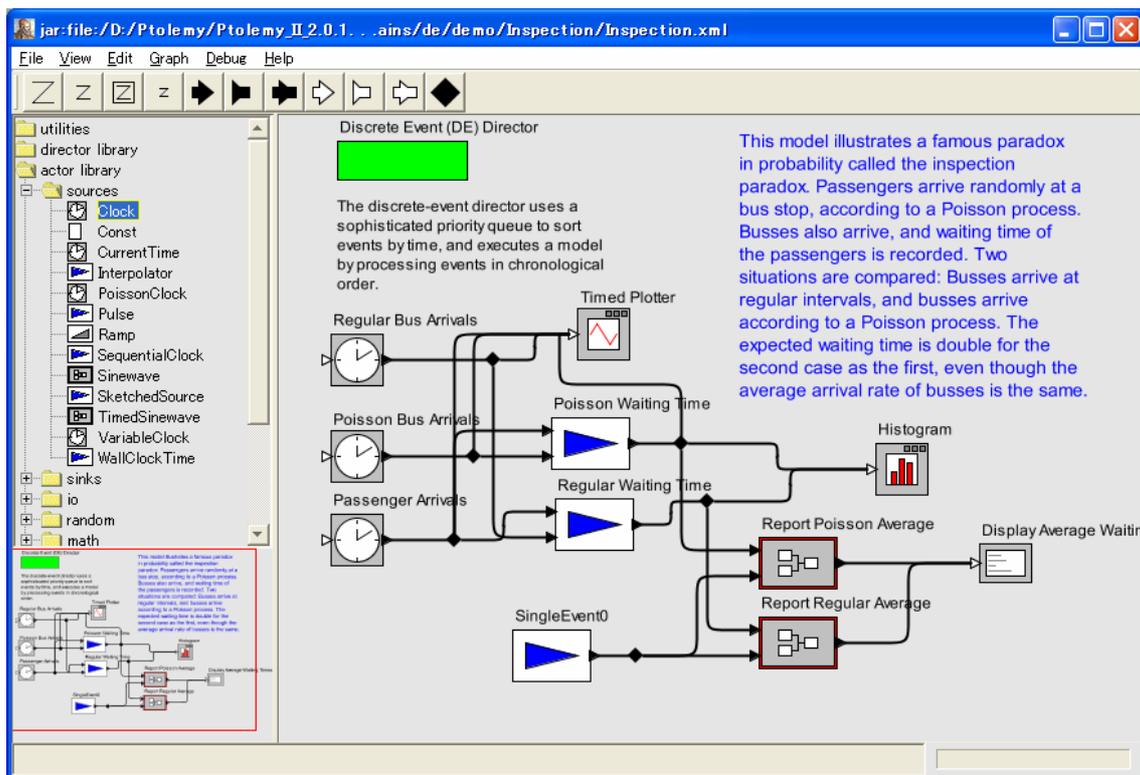


図2 Vergil のバス停における待ち行列のシミュレーション例

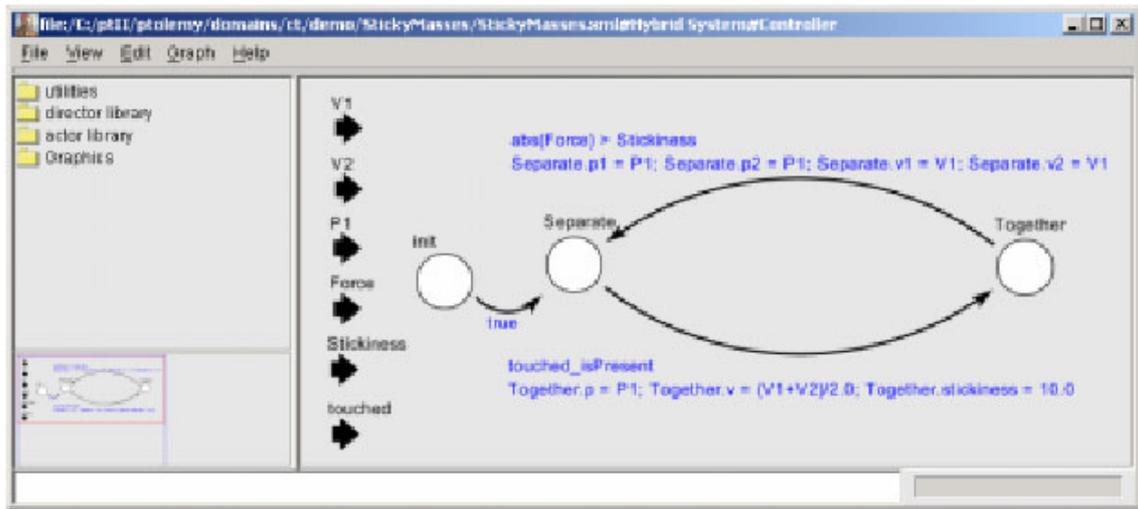


図3 Vergil の状態遷移図の例

### 2.4.1.3 UML,MDA

UML(Unified Modeling Language)は、オブジェクト指向に基づく開発方法論で使われるドキュメント(図)を統一したものであって、以下の9種類の図を規定している。

ユースケース図……………システム要件となる機能を表す図

クラス図……………クラスの階層構造、振る舞いを表す図

- オブジェクト図……………個々のオブジェクトの振る舞いを表す図
- シーケンス図……………オブジェクト間の協調動作（メッセージングの順序）を表す図
- コラボレーション図…オブジェクト間の協調動作（メッセージングとコンテキスト）を表す図
- 状態遷移図……………オブジェクトの状態の変化の様子を表す図
- アクティビティ図………操作の中で実行されるアクティビティの流れを表す図
- 配置図……………システムのハードウェアとソフトウェアの物理的なアーキテクチャを表す図
- コンポーネント図………コードのコンポーネントの物理構造を表す図

しかし、このような UML には大きな二つ課題を抱えている。一つ目は UML の図は、モデリング言語ではあるが、オブジェクト指向言語によるインプリメントを前提としたものであって、IT システム設計者、プログラマ向けのものである。業務で使うモデリングではなく、業務の専門家が使えるものではない。二つ目は、実行コードの生成までを可能とする情報は含んでいないため、設計からインプリメント、テストへとシームレスな開発が行えず、また、ツールを使えば、或る程度は補えるが、実行コードと UML との不一致が生じたりする。

この課題を解決する動きが、MDA(Model Driven Architecture)[6]、UML2 である。

MDA は OMG(Object management Group)が提案しているものであり、業務に視点を置いたモデリング(PIM: Platform Independent Model)とインプリメントに視点を置いたモデリング(PSM: Platform Specific Model)とに分けてモデリングを行い、この二つのモデリングをマッピングルールで関係付けを行おうとするものである。

PIM の一つとして、企業でのビジネスプロセスを対象としたエンタープライズ分散オブジェクトコンピューティング(EDOC: Enterprise Distributed Object Computing)向けに UML に対して、ビジネスプロセス、イベントドリブン、コンポーネント、パターンなど様々な拡張が検討されている。PSM も分散オブジェクトをサポートする CORBA、EJB、NET、XML が検討されている。図 4 (<http://cgi.omg.org/docs/ptc/02-02-05.pdf>) に EDOC の CCA(Component Collaboration Architecture)による販売の例を示す。

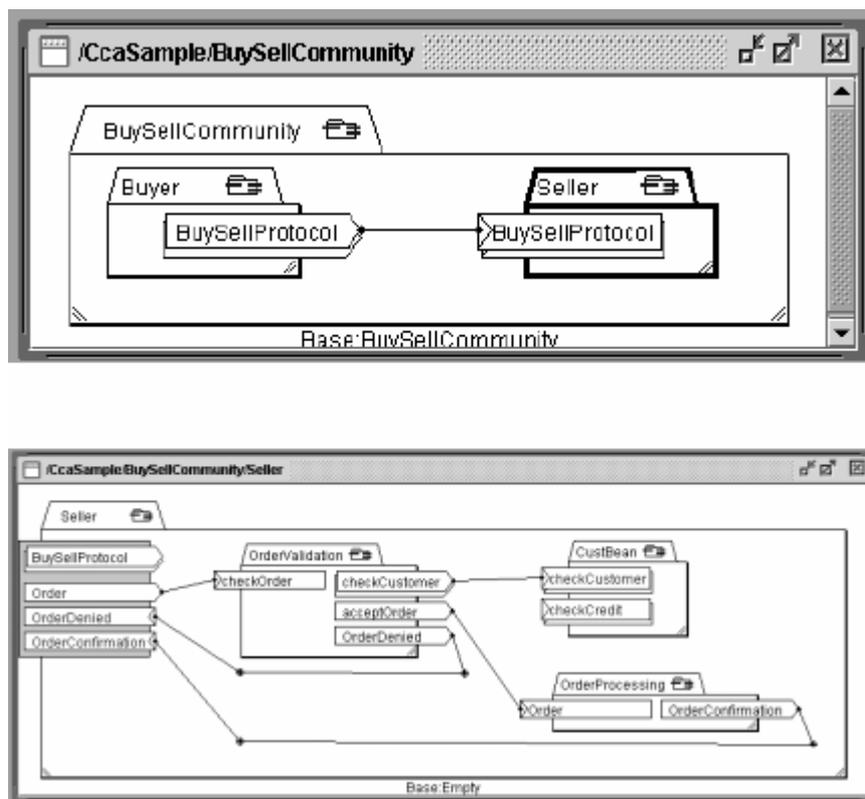


図4 EDOCのCCA(Component Collaboration Architecture)による販売の例

#### 2.4.1.4 BPEL4WS

BPEL4WS(Business Process Execution Language for Web Services)[7]は、ウェブサービスをベースとしたワークフローによるビジネスプロセスの振る舞いを記述するXMLベースのドキュメントであり、IBM、Microsoft、BEAが共同で提案しているものである。BPEL4WSは、XLANGとWSFLの概念を融合しており、それらに取って代わるものである。BPEL4WSは、ビジネスプロセスの二つの側面、ビジネス活動における参加者の振る舞いのモデルを進行させる実行プロセスと各業者間のメッセージ交換の振る舞いを記述するビジネスプロトコルとを記述する。

BPEL4WSでは、ルートのProcess要素の下位要素として、Partners要素、Containers要素、CorrelationSets要素、FaultHandlers要素、CompensationHandler要素、及びアクティビティを記述する。

Partners要素では、ビジネスプロセスの過程で登場する各業者について、Partner要素でサービスリンクタイプと役割を記述する。Containers要素では、ビジネスプロセスで使われる各データ内容について、Container要素で記述する。CorrelationSets要素では、各相互関係について、CorrelationSet要素で記述する。エラー処理については、FaultHandlers

要素、CompensationHandler 要素で記述する。

アクティビティは、ビジネスプロセスの活動内容を以下の要素で記述する。

- Receive 要素 メッセージの受け取り
- Reply 要素 メッセージに対する返答
- Invoke 要素 パートナーの PortType で提供されるサービスの起動
- Assign 要素 新しいデータで Container の内容の変更
- Throw 要素 例外処理の生成
- Terminate 要素 ビジネスプロセスの終了
- Wait 要素 或る時間待つ
- Empty 要素 無操作
- Sequence 要素 順次に各操作を行う
- Switch 要素 分岐操作を行う
- While 要素 繰り返し操作を行う
- Pick 要素 メッセージの受け取りまたはタイムアウトでアクティビティを実行
- Flow 要素 並列操作を行う
- Scope 要素 エラー処理の範囲を定義
- Compensate 要素 補償処理の呼び出し

## 2.4.2 多面的ソフトウェア

### 2.4.2.1 アスペクト指向

オブジェクト指向は、システムを「もの」、すなわちオブジェクトという側面から分割し、対象システム中の「もの」をモデル化し、その性質（手続きとデータ）を記述する。しかし、対象モデルの性質は、一つの側面ではなく、いくつかの側面を含んでいる場合が多い。このことに着目し、「もの」という側面からではなく、他の側面からもシステムを分割しようとするのがアスペクト指向[8]である。まず、オブジェクトの側面からシステムを分割し、分割したオブジェクトを横断するような関心事があれば（複数のクラスに同じ関心事があれば）、その関心事に対する機能を一まとめにして分割の単位とする。このように、アスペクト指向とオブジェクト指向とを組み合わせるシステムを構築する。

オブジェクトに跨る関心事には、すべてのオブジェクトに共通な関心事の例として、セキュリティチェック、課金、トレース、ロギング等があり、幾つかのクラス間の取り決めの例として、Observer—Subject のようなデザイン・パターンに見られるクラス、図形描画における図形要素と画面の再描画の関係といった構成要素と全体の管理の処理、クラス間の役割を明確化するメソッド呼び出しにおけるプレコンディション・ポストコンディシ

ョン等がある。また、システム統合を行う場合にもアスペクト指向は有効である。例えば、販売管理システムと配送システムとを統合する場合には、商品の重量、大きさ、配送スケジュール等といった配送の側面から機能を一まとめにして、販売管理システムに後付けで（販売管理システムに大きな変更を加えずに）追加することもできる。

このようなアスペクト指向をサポートするシステムには、Hyper/J、Aspect/J 等がある。Aspect/J を用いた図形描画における図形要素と画面の再描画の例で、具体的にどのようにコーディングするかを図 5 に示す。

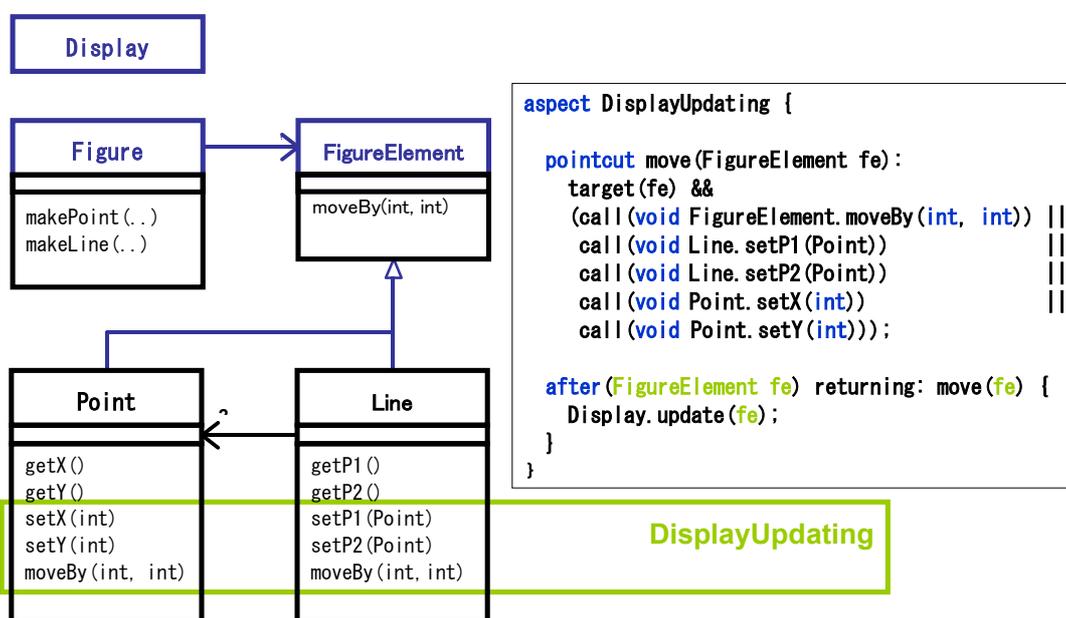


図 5 Aspect/J による図形要素と画面の再描画の例

図 5 の左に示したのが、オブジェクト指向による図形描画のクラスの階層構造である。ここで、図形要素の **Point**、**Line** が変更されたときには、必ず画面の再描画をしなければならないが、その処理は、各図形要素すべてで記述しなければならない、モジュール性がよくない。アスペクト指向では、その部分を別モジュール **aspect DisplayUpdating** で記述する。そこでは、**pointcut** 文で **Point**、**Line** が変更されたという関心事 **move** を (`call(FigureElement.moveBy|Point.setX|Point.setY|Line.setP1|Line.setP2)`) で指定し、そのときのオブジェクトを **fe** と指定する。そして、関心事 **move** が呼び出された後処理として行うべき処理（画面の再描画）を **after** 文で指定している。

### 2.4.3 科学、工学的な設計

現在のソフトウェアが抱える低い信頼性、生産性、相互運用性、柔軟性は、技術的な脆弱性に起因している。それを解決するには、ソフトウェアは、科学的原理、方法、経験に基づいて設計すべきであり、このような科学、工学的な設計によって、実行コードを自動生成したり、事前にエラー、矛盾性をチェックしたり、テストを自動化したりしようとする試みである。このような試みに、軽い形式的枠組み、**Semantic Web** がある。

#### 2.4.3.1 軽い形式的枠組み

軽い形式的枠組みとは何かを説明する前に、先ず、形式的枠組み、擬似形式的枠組み、及び、非形式的枠組みについて説明しよう。

形式的枠組みは、**Hoare** のロジックやリレーショナルデータベースの計算のように、数学的理論で裏づけされた概念で記述しようとする枠組みである。これは、類似、プリ／ポストコンデション、信頼／保証、連続性といった強力な概念の宝庫ではあるが、多くのアプリケーションにおいて、実用的ではない。

一方、非形式的枠組みは、デザイン・パターン、エクトリーム・プログラミングがその例で、優秀なプログラマ、マネジャーの経験を発掘して、その経験を実践しようとするもので、非常に効果的な方法である。しかし、人の経験的能力に依存しており、経験的能力を身に付けるまでの時間、教育コストが課題である。

また、擬似形式的枠組みは、**UML** がその代表例であり、意外に有効ではあるが、その有効性は、ソフトウェア設計を形式化、標準化することによって得られるものであり、改善の策と言えよう。

軽い形式的枠組み[9]は、型付けシステム、モデルチェッキング、ランタイムモニタリングといった実用的で強力な概念を採り入れる試みである。これらの概念を拡張してより強力にしたり、新たなこのような概念を見つけていけば、現在のソフトウェアが抱える問題を飛躍的に解決する可能性がある方法である。

例えば、型付けシステムは、コンポーネントとその動作環境との間の約束事を決めて置き、コンパイル時に、その約束事に従っているかを自動的にチェックするものであるが、実行時に動作環境が動的に変わるような場合においても、この型づけの概念、機構を適用されるようにすれば、証明付きコードを実現できる。また、モデルチェッキングは、ハードウェア設計において、流れを示すのに非常に有効であるが、これをソフトウェアに対しても適用できるようにする。ランタイムモデリングは、不正なふるまいを事前に推測するといった労力のいる仕事をしなくても、実行時に実際に不正な振る舞いをするコンポーネントを容易に検出できるようにする。

## 2.4.3.2 メタデータ、Semantic Web

メタデータ、Semantic Web[10]は、W3Cなどの団体が中心となって、その標準化、推進活動を行っている。メタデータは、データについてのデータであり、データを或る目的のために扱いやすくするデータのことである。例えば、以下のようなデータがある。

- ・データを構造化したデータ
- ・データリソースについてのデータ
- ・カタログデータ
- ・データを共通認識させる意味データ

Semantic Webは、データについての意味情報であり、W3Cのディレクターであるティム・バーナーズ・リーらによって提唱された。その考え方は、サイエンティフィック・アメリカン誌で、以下のように紹介され、計算機が情報の意味を理解できるようにし、各意味情報を用いて、自動的な処理、知的な推論処理を行うことによって、人々が知識共有をグローバルに効率よく効果的に行うことを目指している。

「The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.」

Semantic Webは、メタデータの拡張として記述される。

図6に示すように、Semantic Webは、XML Schema、RDF、RDF Schema、DAML+OILなどのメタデータの記述を階層的に定義することによって記述される。

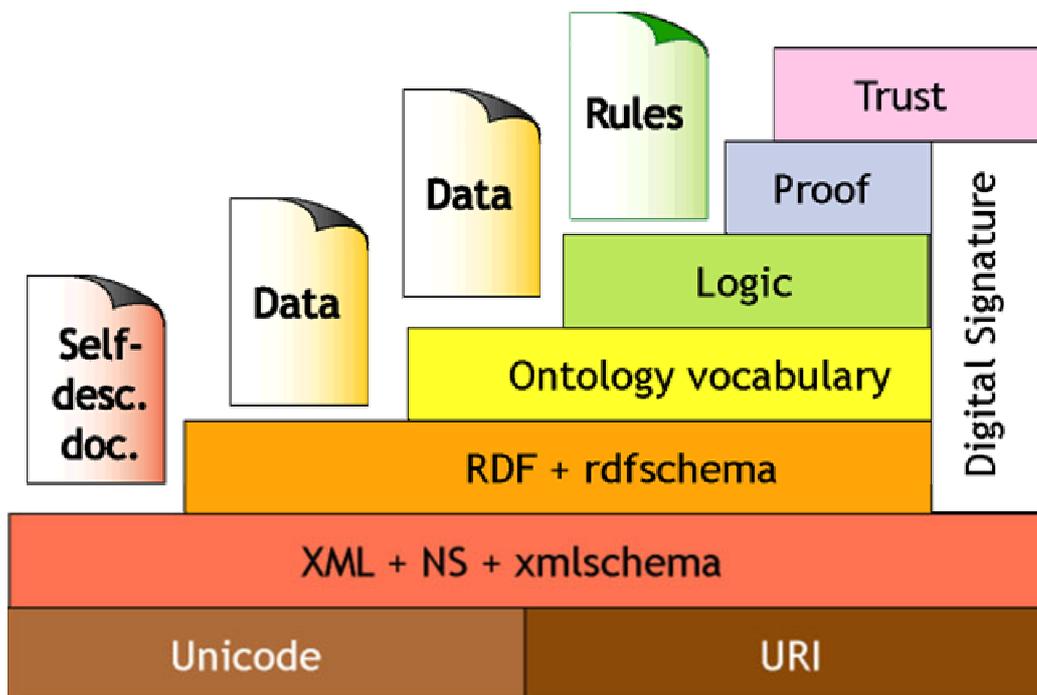


図6 Semantic Webのアーキテクチャ

Semantic Web プロジェクトの特徴は、セマンテックスの標準化に向けて、非集中型のアプローチを採用していることである。データに対し意味を付与し、計算機が理解可能とするには、XML のタグ、プロパティの用語を標準化し、その意味付けを明確化するメタデータの定義が必要となるが、そのメタデータの標準化にあたっては、一つに統一するのではなく、複数の標準が存在することを許容している。それぞれの目的のために、数多くのメタデータが定義され、標準化されてくると、それらの複数のメタデータには、同義語、同音異義語などが生じ、用語の相違が生まれる。このような場合には、同義語辞書（オントロジー）を作ることによって、その用語の相違を解消していこうとしている。この同義語辞書もメタデータで、図 6 の **Ontology vocabulary** レイヤで定義される。同義語辞書を作成してもセマンティクスギャップ（用語の相違）は完全には解消されないが、現実的なアプローチであると言える。データベースがデータの整合性、一貫性を死守するあまり、分散化、共有化が進まなかったのに対し、Web が、データの整合性、一貫性を放棄して分散化を進め、ボトムアップの草の根型の発展を遂げ、Web をグローバルな巨大なデータベース化するのに成功したように、セマンティクスギャップを許容することによって、ボトムアップの草の根型に意味、知識が蓄積されて、巨大な知識ベースが構築されることを目論んでいる。次に、図 6 に示した Semantic Web アーキテクチャの各レイヤの内容を以下に示す。なお、Logic、Proof、Trust レイヤは、構想段階では何も決まっていないので、省略する。

### (1) XML Schema

XML Schema[11]は、XML ドキュメントの以下に示すクラス情報を記述する。

- ・ 構造情報：どのようなエレメント、属性から構成されているか
- ・ データタイプ：そのエレメント、属性がどのようなデータタイプであるか

```
<xsd:element name="Price" type=xsd:decimal>
```

- ・ データの基本情報：エレメントの頻出回数の制約などの情報

```
<element name="item" minOccurs="1" maxOccurs="unbounded">
```

これには、意味情報はなく、DTD と同様の機能を提供するが、DTD は XML でないのに対し、XML Schema は XML であるので、XML で使われているツールが XML Schema に対しても使えるメリットがある。

### (2) RDF(Resource Description Framework)

RDF[12]は、メタデータを表すための枠組みであり、計算機が理解可能とするための情報記述を与える。RDF では、リソースとプロパティと値の三つ組で構造を表す極めてシン

フルな構造化モデルを採用している。

例えば、リソース : [http:// www.w3.org/Home/Lassila](http://www.w3.org/Home/Lassila) のプロパティ : 著者の値 : Ora Lassila は、次のように表される。

```
<rdf:Description about="http://www.w3.org/Home/Lassila">
  <Name>Ora Lassila</Name>
</rdf:Description>
```

値は、リソースを指してよいので、ネットワークとして定義していくことができる。

「[http:// www.w3.org/Home/Lassila](http://www.w3.org/Home/Lassila) (リソース) の著者 (プロパティ) は、個人 <http://www.w3.org/staffId/85740> (リソース) であり、その個人の名前 (プロパティ) は Lassila (値)、Email アドレス (値) は [lassila@w3.org](mailto:lassila@w3.org) である。」は、次のように記述される。

```
<rdf:RDF>
  <rdf:Description about="http://www.w3.org/Home/Lassila">
    <Creator>
      <rdf:Description about="http://www.w3.org/staffId/85740">
        <Name>Ora Lassila</Name>
        <Email>lassila@w3.org</Email>
      </rdf:Description>
    </Creator>
  </rdf:Description>
</rdf:RDF>
```

また、構造を表すモデルとしてコンテナモデルを採用し、<Bag>, <Sequence>, <Alternative>が用いられる。

### (3) RDF Schema

RDF Schema[13]は、RDF の以下のようなデータタイプを記述する。

- クラス、及び他のクラスとの関係を定義  
オブジェクト指向風に<Class>, <subClassOf>が用いられる。
- プロパティ、及び他のプロパティとの関係を定義  
<subPropertyOf>で他のプロパティの定義を継承することを示す。
- プロパティの値の範囲、属するクラスなどプロパティの制約を定義  
<range>でプロパティの値の範囲を、<domain>でプロパティが属するクラスを指定する。
- 意味 (ラベルとコメント) を定義  
<label>でラベルを、<comment>でコメントを指定する。

#### (4) メタデータの例—Dublin Core

以上の RDF、RDF Schema を用いてメタデータが定義され、各業界、団体で標準化作業が行われている。その主なものには、以下のものがある。

- Dublin Core ; 書誌情報の基本語彙集
- P3P : プライバシに関するメタデータ
- RSS : サイト情報の要約と公開
- WSDL : リモートプロシジャコール (SOAP) のメタ情報

Dublin Core[14]は、Web や文書の作者、タイトルなどの書誌情報の基本語彙集である。次の 15 の基本エレメントタイプが決められている。

Title Creator Subject Description Publisher Contributor Date Type  
Format Identifier Source Language Relation Coverage Right

さらに基本エレメントタイプよりもっと厳密なメタデータ記述を行うために、以下のよ  
うな修飾子が規定されている。

Title.alternative

Description.tableOfContents Description.abstract

Date.created Date.Valid Date.available Date.issued Date.modified

Format.extent Format.medium

Relation.isVersionOf Relation.hasVersion Relation.isReplaceBy Relation.replace

Relation.isRequiredBy Relation.requires Relation.isPartOf Relation.references

Relation.isFormatOf Relation.hasFormat

Coverage.special Coverage.temporal

#### (5) DAML+OIL Ontology

DAML+OIL Ontology[16][17]は、RDF Schema に加えて、より詳細なクラス、プロ  
パティの定義、制約、表記を記述する。

これらの記述のために以下のタグが規定されている。

cardinality cardinalityQ Class complementOf Datatype DatatypeProperty  
DatatypeRestriction Datatype value differentIndividualFrom disjointUnionOf  
disjointWith domain equivalentTo hasClass hasClassQ hasValue imports  
intersectionOf inverseOf maxCardinality maxCardinalityQ minCardinality  
minCardinalityQ ObjectClass ObjectProperty ObjectRestriction oneOf  
onProperty Ontology Property range Restriction sameClassAs  
sameIndividualAs samePropertyAs subclassOf subPropertyOf toClass  
TransitiveProperty UnambiguousProperty unionOf UniqueProperty versionInfo

### 2.4.4 オープンな協調開発環境

#### 2.4.4.1 オープンソースソフトプロセス

Linux は、オープンソースソフトプロセス (OSSP) [18]の成果である。OSSP を採用したが故に、Linux は、安定性・スケーラビリティ・カスタマイズ性・性能の向上、開発・バグ修正の素早さ、低コストをもたらし、マイクロソフトを脅かす存在となっている。OSSP は、単にソースを公開した方式にとどまらず、インターネットを最大限に活用した開発方式であり、市場原理に基づく開発方式とは異なるハッカー社会の贈与文化に基づく開発方式である。

#### [OSSP によるソフトウェア開発の特徴]

OSSP によるソフトウェア開発方法の特徴を以下に示す。

##### ① OSSP の開発組織

- ・地域的な広がり

国際化、標準化に優位性を持つ。

- ・万人が認めるリーダーの存在

Linux では Linus Torvalds の存在

- ・巨大な数の貢献者集団

Linux ではコード提供者 200 人以上、バグ修正者 1000 人以上

- ・金銭的動機ではなく、余暇に開発に貢献する個人集団

(ただし、商業版 Linux が登場して変わりつつある)

##### ② SP が成功する開発運営方法

- ・インターネットの共働技術を駆使

メーリングリスト、ニュースグループ、Web

- ・共通の目標

「UNIX を作り直せ」=>方向性を明確にし、開発チーム全員の意思決定

(「すばらしい OS を作ろう」といったあいまいな目標はダメ)

UNIX の経験=>何がうまくいって、何がうまくいかないかが分かる。

- ・共通の経験、スキル

UNIX/gnu=>参入障壁を低くし、参加できる開発者を増やす。

- ・並列競合開発とコンポーネント化の枠組み

コンポーネントの枠組みを確立し、複数の小チーム、個人が個別に開発。

競合した場合は、最もいい実装を選択。

- ・ 並列デバッグ

誰かが問題を見つけ、たいてい別の人がその問題を理解してそれを直す。

デバッグは、プロジェクトに参加する人の数に正比例して効率が上がる。

これが、ブルックスの法則（開発者の人数を増やしても、管理、調整作業が増大して効率が比例して上がらない）から逃れる鍵

- ・ 紛争解決

Linux では、「やさしい独裁者」モデル

Linus Torvalds がプロジェクトのリーダー。

大きなコンポーネントを信頼できる副官に移譲。

副官は、サブコンポーネント開発者に移譲。

Apache は、共同開発者による議決委員会方式。

- ・ 動機づけ

開発者の個人的な悩みを解決

自分のエゴの満足とハッカー社会での評判

ハッカーの社会は、贈与文化。

ものが豊富な社会（ソフトが自由に共有される社会）では、競争的な成功の尺度は仲間内の評判

ノウアスフィアの開墾

未開の土地を開拓し、自分のものにする

弱者の反発（反マイクロソフト）

- ・ コードの分裂の回避

万人が認めるリーダーの存在：Linus Torvalds

GPL ライセンス（派生したものはフリーにしなければならない）

ハッカー社会での評判が低下する恐れ

## [OSSP 開発を効率化する法則]

Linux 上の fetchmail を開発したレイモンドは、その OSSP を実践した経験から、OSSP 開発を効率化する法則を「伽藍とバザール」[19]に書いている。その法則を以下に列挙する。

- ①よいソフトはすべて、開発者の個人的な悩み解決から始まる。
- ②何を書けばいいかわかっているのがよいプログラマ。なにを書き直せば（そして使い回せば）いいかわかっているのが、すごいプログラマ。
- ③捨てることをあらかじめ予定しておけ。どうせいやでも捨てることになるのだから（フ

## 2.4 SDP の目指すものに関連した技術、システムの動向

レッド・ブルックス『人月の神話』第 11 章)

- ④まともな行動をとってれば、おもしろい問題のほうからこっちを見つけだしてくれる。
- ⑤あるソフトに興味をなくしたら、最後の仕事としてそれを有能な後継者に引き渡すこと。
- ⑥ユーザを共同開発者として扱うのは、コードの高速改良と効率よいデバッグの一番楽しい方法。
- ⑦早目のリリース、頻繁なリリース。そして顧客の話聞くこと。
- ⑧ベータテスタと共同開発者の基盤さえ十分大きければ、ほとんどすべての問題はすぐに見つけだされて、その直し方もだれかにはすぐわかるはず。
- ⑨賢いデータ構造と間抜けなコードのほうが、その逆よりずっとまし。
- ⑩ベータテスタをすごく大事な資源であるかのように扱えば、向こうも実際に大事な資源となることで報いてくれる。
- ⑪いいアイデアを思いつく次善の策は、ユーザからのいいアイデアを認識することである。時にはどっちが次善かわからなかったりする。
- ⑫もっとも衝撃的で革新的な解決策が、自分の問題のとらえかたがそもそも間違っていたという認識からくるということはよくある。
- ⑬「完成」(デザイン上の)とは、付け加えるものが何もなくなったときではなく、むしろなにも取り去るものがなくなったとき。
- ⑭ツールはすべて期待通りの役にたたなきやダメだが、すごいツールはまったく予想もしなかったような役にもたってしまう。
- ⑮ゲートウェイソフトを書くときはいかなる場合でも、データストリームへの干渉は最低限におさえるように必死で努力すること。そして受け手がわがどうしてもと言わない限り、絶対に情報を捨てないこと!
- ⑯自分の言語がチューリング的完成からほど遠い場合には、構文上の甘さを許すといろいろ楽になるかもね。
- ⑰セキュリティシステムのセキュリティは、そこで使われている秘密の安全性にかかっている。見かけだけの秘密は要注意。
- ⑱おもしろい問題を解決するには、まず自分にとっておもしろい問題を見つけることから始めよう。
- ⑲開発コーディネーターが、最低でもインターネットくらい使えるメディアを持っていて、圧力なしに先導するやりかたを知っている場合には、頭数は一つよりは多いほうが絶対にいい。

## [OPPS の強み]

指数関数的性質

- OSSP はインターネットと共に成長
- OSSP では「勝者がすべてをかつさらう」
- 開発者は最大の OSSP プラットフォームに貢献したがる。
- 大規模な OSSP プロジェクトのほうがもっとたくさんの「目先の問題」を解決する。

長期的な信用

- バイナリは死に絶えるがソースコードは永遠なり。
- コード分裂の不在が長期的信用につながる。

並列デバッグ

並列開発

完璧な API とドキュメンテーション

頻繁なリリース

## [OSSP の弱み]

OSSP プロジェクトを始めるのは難しい。

OSSP を信用させ、開発に貢献するには以下のことが基準を満たさないといけない。

大きな将来のノウアスフィア

大きな悩みを解決する

まずは問題のそこそこの部分を解決せよ

飽和点に達した後の開発

飽和点に達すると、OSSP が成功する要件である「共通の目標・スキル」、および「大きなノウアスフィア」が薄れてくる。これをどう克服するか。

非専門家からのフィードバック

OSSP は、「開発者の個人的な悩みを解決」が動機付け

技術水準の高いユーザ向けシステムになりがち

### 2.4.4.2 エクストリーム・プログラミング(XP)

エクストリーム・プログラミング[20]とは、Kent Beck らによって提唱されているソフトウェア開発方法であり、略して XP(eXtreme Programming)と呼ばれている。

XP は、コミュニケーション、シンプルさ、フィードバック、勇気に価値を置く開発手法である。その特徴は、変化を容認 (Embrace Change) するとの思想に立ち、その変化に対応できるように、初期設計はシンプルに行い、リファクタリングによる再設計を重視し、

また、変更コストが時間とともに増大しないように、テストを自動化するなどプログラミング、テストを重視している。また、チームに責任と権限が与えられ、メンバがプロセスを最適化していくセルフオーガナイズされたチームを目指し、人間であるプログラマに大切にしている思想に立っている。

### [12 のプラクティス]

XP は、これらの思想に立脚し、次の 12 のプラクティス（経験に基づいて有用性が立証された実践項目）を示している。

#### ① 計画ゲーム(The Planning Game)

ビジネス優先度と技術的見積により次回リリースの範囲を早急に決める。現実が計画と変わったら、計画を更新する。

#### ② 小規模リリース(Small Releases)

シンプルなシステムを早急に生産に投入する、それから新バージョンを非常に短いサイクルでリリースしていく。

#### ③ 比喩(Metaphor)

どの様に全体のシステムが機能するかを示すシンプルな喩え話(メタファー)をメンバが共有することで全ての開発を導く(ガイドする)。

#### ④ シンプルデザイン(Simple Design)

いつでもシステムは出来る限りシンプルに設計されるべきだ。余分な複雑さは見つけ次第取り除かれる。

#### ⑤ テスティング(Testing)

プログラマは継続的にユニットテストを書く、それは開発を続けるために完全に動かなければならない。顧客は、機能の開発が終わったことを示す機能テストを書く。

#### ⑥ リファクタリング(Refactoring)

二重コードを取り去り、コミュニケーションを改善し、単純化し、柔軟性を加えるために、プログラマは、システムの動作を変えずにシステムを再構成する。

#### ⑦ ペアプログラミング(Pair Programming)

全てのコードは 2 人のプログラマにより一台のマシンで書かれる。

#### ⑧ 共同所有権(Collective Ownership)

誰でも、どのコードでも、どこでも、いつでも、プログラマはコードを修正できる。

#### ⑨ 継続的インテグレーション(Continuous Integration)

システムを一日に何回もインテグレートしビルドし、テストを 100% パスさせる。

⑩ 週 40 時間(40-Hour Week)

週 40 時間以上仕事をしてはいけないのがルール。

⑪ オンサイト顧客(On-Site Customer)

現実のユーザをチームに加えて、フルタイムで質問に答えられるようにする。

⑫ コーディング標準(Coding Standards)

プログラマは、コーディング標準に従って全てのコードを書く。

## 2.4.5 コンポーネントウェア

### 2.4.5.1 デザイン・パターン

デザイン・パターン[21]は、ソフトウェアの拡張性・再利用性を高める構成法のカタログである。

ソフトウェア開発の生産性を高める有効な手段の一つが再利用である。繰り返し使われる機能を集めた関数ライブラリやオブジェクト指向に基づき機能とデータを一つにまとめて内部を隠蔽したクラスライブラリといったコンポーネントは、再利用形態の例である。

この他の再利用の形態として、フレームワークがある。フレームワークは、コンポーネントのように完成品ではなく、半完成品である。繰り返し使われる枠組み（半完成品）だけを提供し、開発者が目的に応じて、その枠組みに拡張を施し、最終的に完成品を効率よく作成することを目指している。通常、互いに関係を持つクラスを集めたクラスライブラリとして提供され、開発者はそれらを継承するなどして独自のクラスを作成し、フレームワークが用意した枠組みに沿ったアプリケーションを作成する。

しかし、フレームワークは、大きな問題を抱えている。それは、その設計アイデアをよく理解し、クラス間の約束事などの仕掛けを理解しなければ、フレームワークを使いこなすことはできないということである。その問題を解決するのがデザイン・パターンである。設計の中で繰り返し使われる設計アイデアをデザイン・パターンとして取り出して名前を付け、ポキャブラリを多くの人々の間で共有することによって、設計アイデアや、それによる構築法の理解を助け、再利用を可能にするのがデザイン・パターンである。

### [GOF のデザイン・パターン]

ソフトウェアのデザイン・パターンのアイデアは、1991 年の OOPSLA (Object-Oriented Programming System, Languages and Applications) で出され、Erich Gamma、Richard Helm、Ralph Johnson、John Vlissides の 4 名による『Design Patterns』が出版されて、世の中に広まった。これには、表 3 に示す 23 個のデザイン・パターンが記述されおり、4 人組になぞらえて、GOF(Gang Of Four)の 23 パターンと呼ばれている。

表 3 GOF の 23 デザイン・パターン

分類	生成	構造	振る舞い
クラス	Factory Method	Adapter (クラス)	Interpreter Template Method
オブジェクト	Abstract Factory Builder Prototype Singleton	Adapter (オブジェクト) Bridge Composite Decorator Facade Flyweight Proxy	Chain of Responsibility Command Iterator Mediator Memento Observer State Strategy Visitor

各デザイン・パターンについて、目的、別名、動機、適用可能性、構造、構成要素、強調関係、結果、実装、サンプルコード、使用例、関連するパターンの説明がなされている。例えば、**Observer** パターンの目的、適用可能性、構造は、以下のように示されている。

[目的]

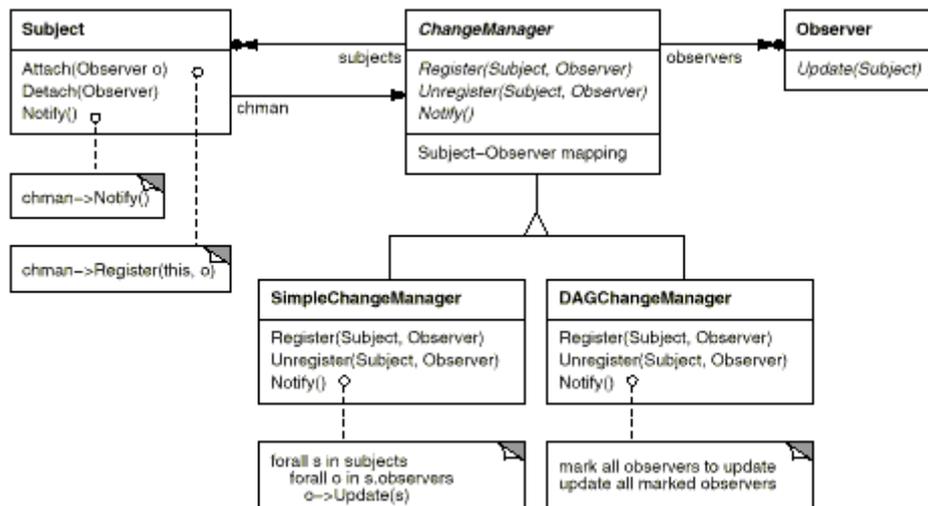
あるオブジェクトが状態を変えたときに、それに依存するすべてのオブジェクトに自動的にそのことが知らされ、また、それらが更新されるように、オブジェクト間に一对多の依存関係を定義する。

[適用可能性]

次のような状況で **Observer** パターンを使う。

- 抽象化により、2つの面が、一方が他方に依存しているという形で現れる場合。これらの面をそれぞれ別々のオブジェクトにカプセル化することにより、それらを独立に変更したり、再利用したりすることが可能になる。
- 1つのオブジェクトを変化させるときに、それに伴いその他のオブジェクトも変化させる必要があり、しかも変化させる必要があるオブジェクトを固定的に決められない場合。
- オブジェクトが、他のオブジェクトに対して、それがどのようなものなのかを仮定せずに通知できるようにする場合。別の言い方をすると、これらのオブジェクトを密に結合したくない場合。

[構造]



### 2.4.5.2 Web サービス

Web サービスとは、広義には、インターネットの標準的なプロトコルを使ってアクセスできるサービス全般を意味する。狭義には、インターネットの標準的なプロトコルとして、サービスの通信プロトコルを規定する SOAP(Simple Object Access Protocol)、サービスのインタフェース定義を規定する WSDL(Web Services Description Language)、サービスディレクトリを規定する UDDI(Universal Description, Discovery, and Integration)を使ってアクセスできるサービスを意味する。

そのコンセプトは、サービス指向アーキテクチャに基づいて、サービスとして表されるソフトウェア・コンポーネントを疎結合でつなぐものである。技術的には、オブジェクト指向、コンポーネントウェア、及び分散処理技術の延長線上にある漸進的な技術であるが、しかし、ソフトウェアアーキテクチャ的には、プラットフォーム独立の実現、クラサバモデルから P2P へ、ビジネス的には、ソフトウェアからサービスへ、ソフトウェアエコシステムの登場、無料 Web から有料 Web 時代へ、といった革新的な変革をもたらすものである。

以下に、Web サービスの基本技術である SOAP、WSDL、UDDI について説明する。

#### [SOAP]

SOAP[22]は、XML を用いてメッセージ交換するための通信規約であり、サービスを実行するときに用いられる。下位の通信層は通常 HTTP を用いるが、これに限定した仕様ではなく、例えば SMTP を用いることもできる。

SOAP によって送受信される XML は、ルートに Envelope 要素があり、その下位要素

に、Header と Body 要素が存在する構造を持つ。Header 部は、主にメッセージの制御のために扱われ、例えば、署名情報などが格納される。Body 要素に、通信したい内容が格納される。この Body の中身は、XML Schema で定義した任意の XML 文書が記述できる。

特に SOAP には、RPC(Remote Procedure Call)を実現するために用いる表現の規約があり、要求側では、Body 要素の中に、メソッド名、引数に関する XML 情報が記述され、一方、応答側に、結果を表す XML 情報が記述される。この RPC に用いる引数や結果の値に利用できる型は、XML Schema で定義された基本データ型(byte, short, integer, double など)と、その複合型(構造体と配列)である。構造体は、XML Schema の文書型定義を用いて表現し、配列は、SOAP encoding による特別な表現が用意されている。

この他、SOAP には、SOAP Messages with Attachments という仕様があり、MIME などでエンコーディングした複数のパートにまたがる文書を転送することができる。

### [WSDL]

WSDL[23]は、Web サービスのサービスインタフェースを XML で定義するための言語であり、サービスがどのようなものであるかを理解するために用いられる。また、この WSDL によって、Java の interface 定義等のプログラミング言語から WSDL を自動生成したり、SOAP のクライアントプログラム用プロキシクラスを生成したりする支援ツールが可能となり、ソフトウェア開発が容易となる。WSDL の XML 文書は、definitions 要素の中に次のような要素を記述する。

- **types**  
引数や返却値に現れる型を宣言する部分。XML Schema による文書型の規定を行う。
- **message**  
交換する一つのメッセージのフォーマットを規定する。message 要素の中に part 要素があり、これが引数のひとつずつに対応する。
- **portType**  
まとまった操作の集合を表す。portType の中には、一つの操作を表す operation 要素がある。さらに、operation 要素の中には、input 要素、output 要素が記述され、これが、それぞれ message 要素と結び付けられる。
- **binding**  
転送プロトコルへのバインディングを行う。例えば、SOAP binding では、WSDL で定義したインタフェースと SOAP メッセージの関係を定義する。
- **service**  
service 要素は port 要素を持ち、サービスのエンドポイントなどを指定する。

## [UDDI]

UDDI[24]は、Web サービスの公開と検索を行うレジストリの仕様を規定し、主にレジストリのデータ構造とレジストリへのアクセスインタフェースの定義を行っており、どのようなサービスがあるかを検索するために用いられる。

その内容は、businessEntity, businessService, bindingTemplate の3階層からなるXML文書と、tModel を表現するXML文書からなる。

- **BusinessEntity**  
サービスを提供する企業に関する情報(企業名、所在地、企業コードなど)を記述する。
- **businessService**  
サービスの内容や種類に関する情報を記述する。
- **bindingTemplate**  
サービスに接続するための情報(tModel へのリンク、エンドポイント)を記述する。
- **tModel**  
サービスにアクセスするためのインタフェース情報(WSDL 等)へのリンクを記述する。

そして、これらの情報を登録、検索するためのAPIとしては、次のものが用意されている。(XXX は検索対象を示す)

- **参照 API (Inquiry API)**
  - **find\_XXX** キーワードなどを用いた検索のためのAPI。検索の結果は対象となった要素のキーがリスト形式で返される。
  - **get\_XXX**: find\_XXX で検索されたキーから、その実体(内容)を取り出すときに使用するAPI。
- **発行 API (Publication API)**
  - **save\_XXX**: オブジェクトを登録するためのAPI。
  - **delete\_XXX**: レジストリのエントリを削除するためのAPI。

## 2.4.6 既存ソフトウェア資産を生かす技術

### 2.4.6.1 アプリケーションインテグレーション技術

現在の社会は、膨大な既存アプリケーションを抱えている。しかも、その多くの既存アプリケーションは、脆弱であり、スパゲティ状態である。SDPが目指すものが実現したとしても、これらの既存アプリケーションを即座に捨てるわけにはいかない。SDPが実現する新しいソフトウェアアーキテクチャを核としたインフラ上に既存アプリケーションを追加・統合し、段階的に消滅させていかねばならない。このために、アプリケーションイン

## 2.4 SDP の目指すものに関連した技術、システムの動向

テグレーション技術が必要となってくる。

アプリケーションインテグレーション技術は、図 7 に示す 5 つのレベルがある。

統合ミドルウェア	ビジネスプロセスマネージャ		
	インテグレーションブローカ		
	スーパーサービス		
	ゲートウェイ		
基本ミドルウェア	データ管理	通信管理	プラットフォーム管理

図 7 アプリケーションインテグレーション技術の階層

基本ミドルウェアのデータ管理は、リモートのファイル、DB を論理的にローカルにあるかのように扱う。たとえば、JDBC、ODBC、NFS、Microsoft の Windows などである。通信管理は、RPC、メッセージ指向ミドルウェアのように、リモートとの通信を同じシステム上の通信のようにする。プラットフォーム管理は、通信管理のスーパーセットで、サーバとサーバ間、サーバとクライアント間のプログラムローディング、メモリ管理等のリソース管理を同一のシステム上の管理のように行う。

統合ミドルウェアのゲートウェイは、異なった DB システム、通信システム、プロットホームを共通のインタフェースで扱えるようにする。スーパーサービスは、異なった複数の OS、サーバにまたがって、ディレクトリ、セキュリティ、トランザクションマネージャといった共通の機能を提供する。インテグレーションブローカは、二つのアプリケーション間のメッセージの内容を変換したり、コンテンツの内容に応じて宛先を決定したりして、アプリケーションの統合をサポートする。ビジネスプロセスマネージャは、アプリケーション統合をマルチステップまで広げて自動化し、ビジネスプロセスのワークフローをサポートする。

繋ぎ方として、メッセージバス方式とハブ方式がある。統合のパターンには、システムが物理的にも論理的にも独立で一方向のデータ同期の方式と、システムが物理的には独立であるが、論理的にはワークフロー的に複数のプロセスが一方向に繋がるマルチプロセス方式と、システムが物理的にも論理的にも繋がり、双方向のデータの同期を取り遅延なく整合性をとる複合アプリ方式とがある。今後は、複合アプリ方式が重要となっていく。

繋ぐ時の通信のモデルには、対話型、リクエスト・リプライ型、メッセージ・パッシング型、ストア&フォワード型、パブリッシュ・サブスクライブ型の 5 つがある。

ストア&フォワード型は、キューを持つことにより、メッセージの完全性、配信保証をし、パブリッシュ・サブスクライブ型は、送信者は宛先を指定しなく、受信者が受信した

い情報を論理的に指定する方式で、送受信者の相互作用を最も柔軟にできる。この2つの通信型のサポートによって、イベントベースのビジネスプロセス統合を実現できることに特徴がある。例えば、新しい注文を受けると、販売報告システム、製造システム、請求伝票システムへ即時に通知され、ビジネスプロセスの統合が遅延なくデータの同期を取って実現できる。

以上のようなアプリケーションインテグレーション技術で既存アプリケーションを統合していくには、脆弱でスパゲティ状態の既存アプリケーションを標準化、部品化、抽象化、再構築するのを自動化、支援するツールとして、コンバータ、アダプタ、ラッパー、リバーエンジニアリング等で構成されるアプリケーションインテグレーション環境が必要と becoming.

#### 2.4.6.2 CBS (COTS-Based Systems Initiative)

CBS[25]は、カネーギーメロン大のSEI(Software Engineering Institute)で研究開発中のプロジェクトで、商業的ソフトウェア・コンポーネント(COTS : Commercial Off The Shelf)、既存ソフトウェア・コンポーネントを用いて、システムを構築するための開発手法に取り組んでいる。CBSは、従来のシステムコンテンツ(要求、価格、スケジュール、OS・サポート環境等のシステムに求められる性格)の選択、アーキテクチャと設計、実装の工程を順次に行うウォーター・フォールモデルではなく、システムコンテンツ、アーキテクチャと設計、市場の於ける製品の能力を同時に検討するアプローチを目指している。

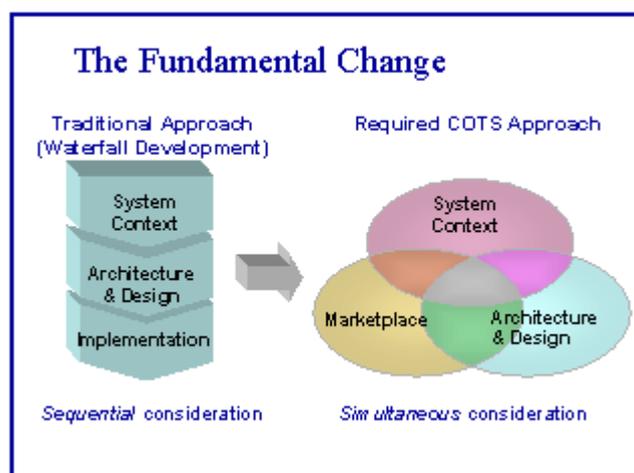


図8 CBSの開発手法のアプローチ

COTS、既存コンポーネントからシステムを構築するには、以下のような質問に答えて行かねばならない。

- どんな技術、製品が適切か？
- 我々のシステムに製品のミスマッチがどの程度あるか？
- 個々のコンポーネントを制御しづらくなるにも拘らず、信頼性、安全性、性能といったシステム特性をどのように設計するか？
- COTS 製品とカスタムコードをどのようにインテグレートしていくか？
- 長いライフタイムに亘ってシステムを出荷する間に COTS の良さをどのように生かすか？

このような難しい質問に答えるために CBS は次の 3 領域をキーポイントとしている。

### (1) 製品と技術の評価の実施

製品と技術の評価に際しては、直接的なシステム要求に応えると同時に、その技術、ソフトウェアの進化が途切れることがないかを考慮しなければならない。そのため、候補となる製品、技術は、関係する製品、技術との関連のなかで評価されなければならない。したがって、問題領域での実験に着手して、クリティカルな質問に答え、COTS の製品、技術の適用に対するガイドラインを作成していくようにしなければならない。

### (2) 入手とマネジメントの実施

CBS では、システムの開発者という立場からシステムの消費者、インテグレータという立場へと変化し、コンポーネントのライセンスング、知的財産権の交渉、開発・メンテナンスコストの見積もり、スケジュールの予測、人材の管理、リスクの同定と軽減といったことに対する戦略が必要となる。さらに、システムに柔軟性を持たせるための知識、いろいろな創造的なソリューションを可能とするドキュメントも要求される。このために、CBS は、CBS をうまく活用した組織の実践例の情報、あるいは、CBS における入手の失敗を避ける情報を提供し、CBS に対する要求の作成、リスクの同定等、CBS プロセスを支援する。

### (3) 設計とソフトウェア・エンジニアリングの実施

CBS システムでは、コンポーネントをインテグレートするために「つぎはぎのコード」が使われ、このため、可読性、発展性、信頼性に欠ける傾向がある。これを克服するため、CBS コンポーネントの人工的な見せかけを記述する参照モデルを開発、提供している。CBS のシステムエンジニアリングでは、図 9 に示すようなプロセスを経てシステム開発が行われる。最初はコンポーネントの多くのプロパティは未知であり、ブラックボックスとして扱われ、システムマティクな調査によって、プロパティが明らかにされる。プロパティが明らかにされると、他のコンポーネントと矛盾したり、対立したりするものが明らかに

なり、その矛盾、ミスマッチがコンポーネントの改造プロセスで除かれる。ミスマッチが除かれると、コンポーネントはシステムに組み込まれ、そして、他のコンポーネントと共に再構成されて、進化していく。

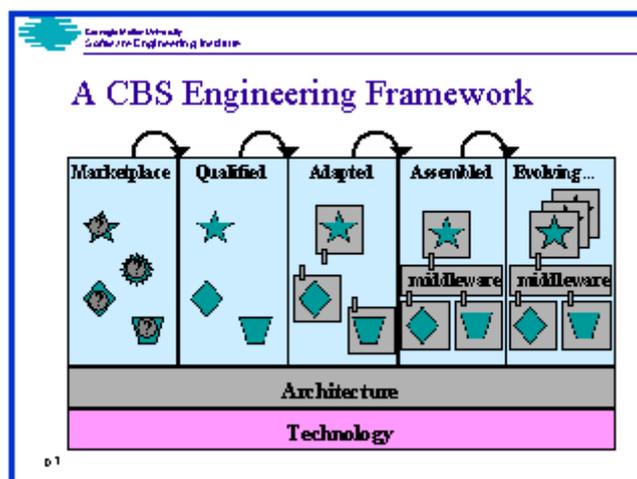


図9 CBSのエンジニアリング・フレームワーク

この参照モデルは、順次的な段階で行われるものでなく、コンポーネントのプロパティによってアーキテクチャ・パターンが決められるものである。

#### [参考文献]

- [1] “Information Technology Research: Investing in Our Future”,  
<http://www.ccic.gov/ac/report/>(日本語版は  
[http://www.icot.or.jp/FTS/REPORTS/H10-reports/AITEC9903Re1\\_Folder/AITEC9905R1-a8-fm.html](http://www.icot.or.jp/FTS/REPORTS/H10-reports/AITEC9903Re1_Folder/AITEC9905R1-a8-fm.html))
- [2] “Planning Workshop on New Visions for Software Design and Productivity”,  
<http://www.itrd.gov/iwg/sdp/planning/index.html>
- [3] “Networking and Information Technology Research and Development – Supplement to the President's Budget for FY2003”,  
<http://www.hpcc.gov/pubs/blue03/index.html>  
 (日本語版は <http://www.icot.or.jp/FTS/Ronbun/BlueBook2003-J.PDF>)
- [4] Simulink, <http://www.mathworks.com/products/simulink/>
- [5] Ptolemy, <http://ptolemy.eecs.berkeley.edu/>
- [6] MDA, <http://www.omg.org/mda/>
- [7] “Business Process Execution Language for Web Services, Version 1.0”,

<http://www-106.ibm.com/developerworks/webservices/library/ws-bpel/>

- [8] Gregor Kiczales, Eric Hilsdale, Jim Hugunin, Mik Kersten, Jeffrey Palm, and William G. Griswold, “Getting Started with AspectJ”, Communications of the ACM, Vol.44 No.10 pp59-65(2001)
- [9] 軽い形式的枠組み、 Benjamin C. Pierce, “Software Research: Where do we go from here?” , <http://www.itrd.gov/iwg/sdp/planning/presentations/UPenn-Pierce.pdf>
- [10] TIM BERNERS-LEE, JAMES HENDLER and ORA LASSILA, “The Semantic Web” <http://www.sciam.com/2001/0501issue/0501berners-lee.html>
- [11] David C. Fallside, “XML Schema Part 0: Primer”  
<http://www.w3.org/TR/xmlschema-0>
- [12] Ora Lassila, Ralph R Swick, “Resource Description Framework(RDF) Model and Syntax Specification”, <http://www.w3.org/TR/REC-rdf-syntax/>
- [13] Dan Brickley, R.V. Guha, ”Resource Description Framework(RDF) Schema Specification 1.0”、 <http://www.w3.org/TR/rdf-schema/>
- [14] “Dublin Core Metadata Element Set, Version 1.1: Reference Description”  
<http://dublincore.org/documents/dces/>
- [15] “Dublin Core Qualifiers”、 <http://dublincore.org/documents/dcmes-qualifiers/>
- [16] “Annotated DAML+OIL (March 2001) Ontology Markup”  
<http://www.daml.org/2001/03/daml+oil-walkthru.html>
- [17] “Reference description of the DAML+OIL (March 2001) ontology markup language” <http://www.daml.org/2001/03/reference.html>
- [18] オープンソースソフトプロセス、 <http://www.opensource.org/halloween.html>  
(日本語版は <http://www.post1.com/home/hiyori13/freeware/halloween.html>)
- [19] “伽藍とバザール”, <http://www.tuxedo.org/~esr/writings/cathedral-bazaar/>  
(日本語版は <http://www.post1.com/home/hiyori13/freeware/cathedral.html>)
- [20] ケント ベック, “XP エクストリーム・プログラミング入門—ソフトウェア開発の究極の手”, ピアソンエデュケーション
- [21] エリック ガンマ, ラルフ ジョンソン, リチャード ヘルム, ジョン ブリシディース, “オブジェクト指向における再利用のためのデザインパターン”, ソフトバンクパブリッシング
- [22] “Simple Object Access Protocol (SOAP) 1.1”, <http://www.w3.org/TR/SOAP/>
- [23] “Web Services Description Language (WSDL) 1.1”, <http://www.w3.org/TR/wsdl>
- [24] “UDDI Version 3.0”, <http://uddi.org/pubs/uddi-v3.00-published-20020719.pdf>
- [25] CBS, [http://www.sei.cmu.edu/cbs/cbs\\_description.html](http://www.sei.cmu.edu/cbs/cbs_description.html)



### 3. 人間主体の知的情報処理に関する研究調査

#### 3.1 アウェアネス技術の動向

國藤 進 委員

##### 3.1.1 あらまし

グループウェア研究が一段落した 1990 年代になって、同期・対面環境では当たり前の存在感、実在感や臨場感がグループウェア環境では欠落していることに気付き、存在感、実在感や臨場感等のアウェアネス（気付き）を補完する研究[1] が勃興してきた。その草分けは Dourish らのジェネラル・アウェアネス「誰が誰と話し、誰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」[2] といった日常の同期・対面作業では当たり前の情報がグループウェア環境では欠けているという認識から出発した。まず存在感や実在感のアウェアネスを伝達するグループウェア環境の構築研究が起こり、ついで対面環境以上の臨場感のアウェアネスを伝達するグループウェア環境を構築したいという研究開発が盛んに行われている。

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、石井は既にゲイズアウェアネスの提供できる環境[3] 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[4]、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付かせる存在のアウェアネス研究[5]、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[6] 等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[7]、情報取得アウェアネス[8]、Web アウェアネス[9]、コミュニティアウェアネス[10]等の様々なアウェアネス研究が行われた。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。

JAIST では臨場感、雰囲気、気配、熱気、凄み等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ており、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や遠隔教育が可能となる。日本学術振興会の未来開拓学術研究推進事業で、宮原は深い感激をも伝達できる電子的 AV システム[11]を試作し、深い感性をも共感できるアウェアネス基盤が構築できることを実証[12]した。これらの研究は癒しを促し、切れる子供達を救う新しいヒーリングテクノロジーとして開花する可能性がある。具体的応用研究として、アウェアネス技術とセンサ技術を融合し、一人所帯あるいは独居老人向きのアウェアホーム[13]を構築しようという動きもあり、興味深い。

知識科学の立場では、アウェアネスは形式知のみならず暗黙知をも支援できるグループ

ウェア環境を構築する研究開発と見做すことができる。サブリミナル・マインド[14]に関する認知心理学の実験によると、人間の記憶には再生、再認、再学習の三つの階層がある。アウェアネス研究はこのうち「言われる、見せられる、あるいは示される」と分かる再認レベルの記憶を、暗黙知でなく形式知に引き上げる技術ととらえることができる。なお脳科学の領域でも、意識の階層構造[15]として自己意識、アウェアネス、覚醒という三階層を考えており、自己意識の下位にアウェアネスと覚醒という広大な無意識世界があることを主張している。その意味でアウェアネス研究の奥行きは深い。

### 3.1.2 CSCW2002 にみるアウェアネス技術

CSCW(Computer Supported Cooperative Work) はアメリカの ACM が隔年(偶数年)に開催するこの分野で最高の国際会議である。類似の国際会議としてヨーロッパで開催される ECSCW があり、これは奇数年に開催されている。

シアトルで開催された CSCW98 に参加した際、研究発表およびデモセッションを中心に「アウェアネス(Awareness)」に関する研究発表が相次ぎ、富士通北陸システムズの中川と参加し、「これからはアウェアネス関連の研究開発を大学、企業関係者が真面目にやらないといけない」と確信した。その後、CSCW2000 では ATR の角が Web and Awareness というワークショップを設け、我々の指摘を実証してくれた。ここに、中川と國藤は既に申請済みの国有特許を受け、当時共同研究した成果を帰国後、直ぐに商品化(商品名 WebCoordinate)した。

今年度の CSCW2002[16,17] はアメリカのニューオーリンズの Hyatt Regency Hotel で開催され、参加者 500 名弱であった。招待講演 2 件、論文発表 39 件、パネル 4 件、ポスター紹介 1 件、併設ワークショップ 12 件、博士論文コロキウム 13 件、チュートリアル 12 件、デモンストレーション 16 件、インタラクティブポスター 36 件、ビデオ 8 件、特別イベント 1 件の多岐に渡る国際会議であった。このうちアウェアネスというキーワードを含む研究発表が論文発表で 3 件、博士論文コロキウム 2 件、デモンストレーション 2 件、インタラクティブポスター 3 件あった。すなわちキーワード・レベルでは 10 件であるが、内容的には 20 件以上のアウェアネス関連の研究発表があった。ここではそのうち我々が注目すべき、将来的にインパクトを持つと思われるアウェアネス研究を 12 ほど紹介する。

全体的にみた大きな研究開発動向としては、基礎研究、応用研究のフェーズを越え、徐々にではあるが具体的商品に繋がる開発研究のフェーズに突入したと言える。また個人活動や小集団活動のアウェアネス研究からコミュニティ活動や組織活動のアウェアネス研究に力点が移りつつある。そのために必要なデータや情報を AV 機器やセンサを用いて徹底的

に蓄積・保存し、そこからどのような知識や知見が得られるのかといった社会科学的分析と、計算機実験を行い、その結果を新たな CSCW ツールの設計イシューへ反映するといった研究が増えてきつつある。

### 3.1.2.1 アウェアネス技術に関する研究発表

#### [Papers ペーパー発表]

#### **Ambiguities, Awareness and Economy: A Study of Emergency Service Work**

M. Pettersson, D. Randall and B. Helgeson, Blekinge Institute of Technology, Sweden  
pp.286-295

本研究はスウェーデンにおける緊急サービスセンターで実施された学習経験から帰結される知見を纏めたものである。この学習は CSCW コミュニティーにとって馴染みの仕事の様相分析に焦点を当てて行なわれた。実際にそこで行なわれている仕事の実際の記録を取り分析すること、技術的に適切な質問が行なわれたかを理解すること、最も重要なのはいかに技術がその使用に際してうまく機能したかである。ここでは二つの事例に焦点をあてた。最初の例では、質問された論点は緊急事が同定され、取り扱われる方法に関する物だった。それはオペレータに取り扱われるべき典型的問題で、携帯電話社会では日常起こる問題である。ここでの対処法は注意深く聴くことであり、Computer Aided Dispatch System の機能を活用しようとするれば「しつこく聴取すること」で対処できる。第二の焦点はこれらのセンターではコンピュータ化された地図が存在するにも係わらず、地図の解釈の適切性に焦点を当てている。これら二つの事例はあいまいな状況で注意深い緊急サービスを行なうには、アウェアネスという概念がシステムの設計イシューに含まれねば行けないことを示している。

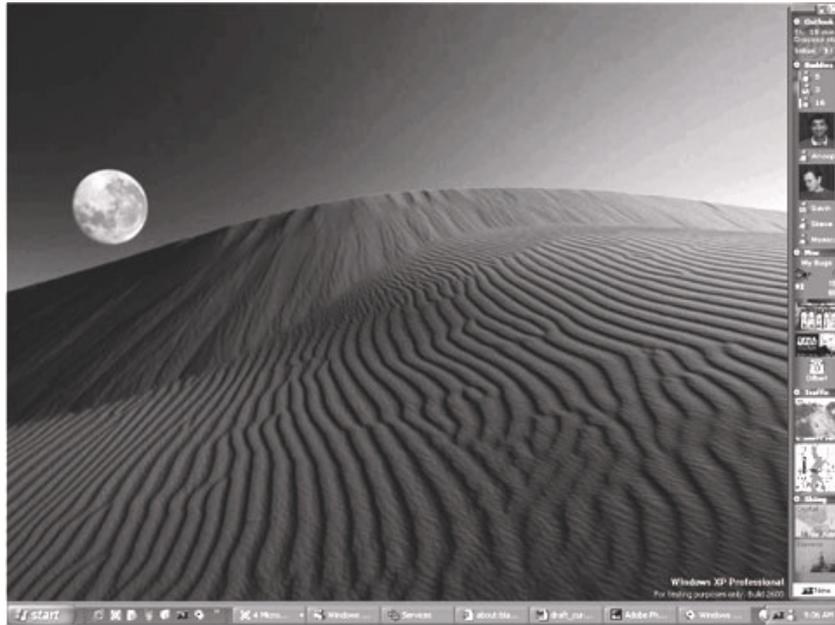
#### **Designing and Deploying an Information Awareness Interface**

Jonathan Cadiz, Gina Venolia, Gavin Jancke, and Anoop Gupta, Microsoft, USA

pp. 314-323

本稿は、情報アウェアネスを提供するインタフェースである Sideshow(図1参照)について述べている。

Sideshow は DHTML と C で記述された Windows アプリケーションである。天気情報やニュースヘッドラインなどのインターネット上にある情報源をユーザが選択すると、アイコン化された情報一覧をデスクトップに常駐する Sideshow から随時参照することができる。予定表やインスタントメッセージの機能も持ち、これらの情報やその変化に対するアウェアネスを促すことができる。



(a)



(b)

Figure 1: Sideshow is a sidebar on the Windows desktop (a). The bar is filled with items called tickets. Users can get more information about a ticket by hovering their mouse over it, which causes a tooltip to appear. The tooltip for a person ticket is shown in (b). Enlarged images are available in Figures 2 and 3.

図 1 Sideshow の画面イメージ

著者は、多くの既存アウェアネス研究が、研究室レベルでの使用に留まっている理由の一つは、ユーザがシステム利用に支払う対価としての手間が、その使用による利益に比べて大きいという問題によるとしている。本システムは、情報源の選択や利用方法に関するカスタマイズの簡易性と自由度を追求することで、様々なユーザが自分の求める利便性に沿って利用可能なシステムとし、これを克服した。

### Stimulating Social Engagement in a Community Network

David R. Millen and John F. Patterson, IBM Research, USA.

pp. 306-313

本稿はオンラインコミュニティでの社会的交流(Social Interaction)を促進または抑制する要因に関する調査研究である。1999年に開始されたオンラインコミュニティ(図2参照)を対象にフィールドスタディを行い、社会的交流の促進要因として、コミュニティシステムのデザイン、参加メンバの役割選択、対話内容といった3つの側面からの検討がなされた。

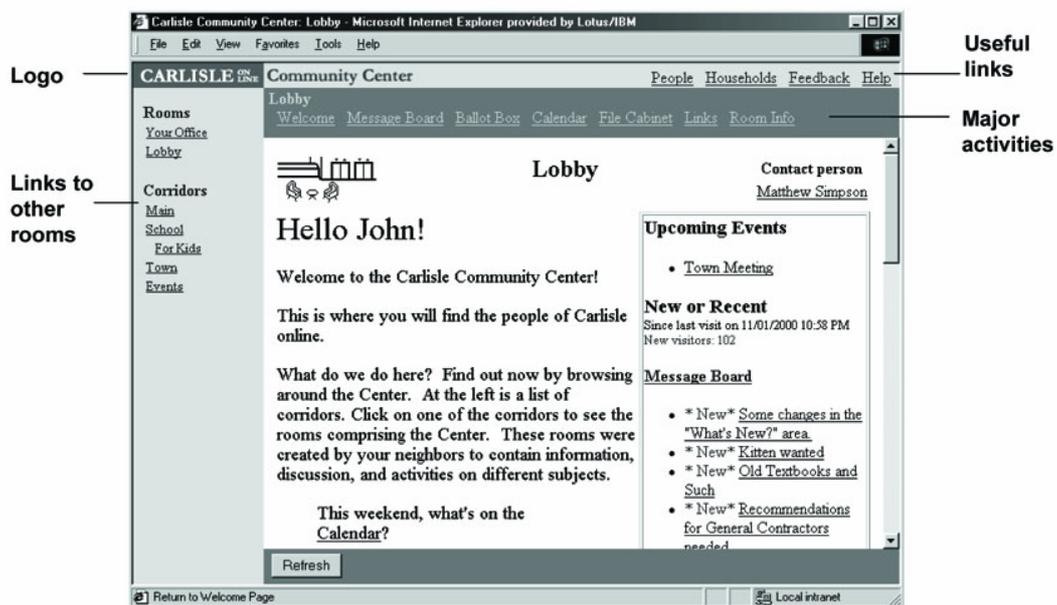


Figure 1 Screen shot from the Carlisle Community Center

図2 オンラインコミュニティページの画面イメージ

デザインの側面からは、システムを参加者全員が集える"ロビー"と複数の"チャンネル"からなるツリー構造にすると、参加者はロビーで全体の動向に関する社会的アウェアネスを得ることができることとされた。メンバの役割選択の側面からは、コミュニティ内での社会的交流に時間を割くことができるサポートメンバを募集することが、特にコミュニティの初期立ち上げ時期には有効であるとされた。また、対話内容の側面からは、サポートメンバによる、人々の興味関心が高いと思われる話題への誘導が、社会的交流の促進に有効であるとされた。

### Work Rhythms: Analyzing Visualizations of Awareness Histories of Distributed Groups

James "Bo" Begole, John Tang, Randall Smith, and Nicole Yankelovich, Sun Microsystems, Inc., USA

本稿では、分散ワークグループを対象とした作業履歴分析によるアウェアネス支援の提案を行っている。作業履歴とは、E-mailの送受信や、キーボードおよびマウス操作の有無を1分単位の時系列で記録したもの(図3参照)である。人々の生活や作業には「リズム」があるという考え方から、記録された作業リズムの規則性を分析的に利用することで、他者の作業状況に関わるアウェアネスを支援することを目的とする。

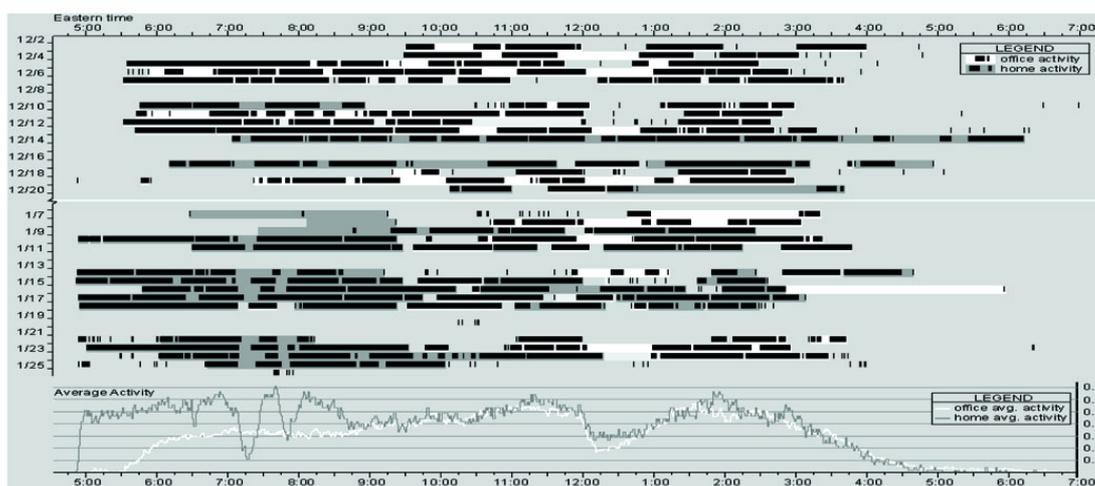


Figure 6. Activity and location information for an individual. This person switched from office to work-from-home after 1/7. His activity rhythm has changed significantly (especially in the mornings) as can be seen by the difference in the mean activity graph at the bottom (white is mean activity up to 1/7 and dark grey is mean activity after 1/7).

図3 可視化された作業アクティビティの画面イメージ

実験では、時差が3時間ある分散ワークグループを対象に、作業履歴の収集が行われた。長期的な収集結果から、日々の出勤退勤や昼食時間といった、作業場の特徴的リズムパターンが抽出された。曜日による作業アクティビティの差異といったパターンの抽出も行わ

れている。本研究では、このようなパターンから、例えば、時差があるユーザ間で連絡を取り合うときに、相手型の作業パターンから最初の連絡に適切な時間を知るといったアウェアネスキュー（手掛かり情報）を得ることが可能である、といった提案がなされている。

#### [Demonstrations デモ発表]

##### **DiamondTouch SDK: Support for Multi-User, Multi-Touch Applications**

Alan Esenther, Cliff Forlines, Kathy Ryall, and Sam Shipman, Mitsubishi Electric Research Labs, USA

pp. 127-130

本発表では、映像の投影が可能なタッチパネル型入力デバイス DiamondTouch (図 4 参照)の紹介と、同開発キット及びアプリケーション例の提案が行われた。

タッチパネルは、小型 PDA や大型プラズマディスプレイなどで採用されることが多い、画面上の任意の一点を押圧により指示することができる入力デバイスである。一般的なタッチパネルでは、複数点への同時押圧を検出することができない。

本装置の特徴は、特殊な RFID ペンなどを使わずに、複数点押圧の認識と押圧者の識別が可能なことである。これにより、テーブル型の共通画面上で、複数ユーザが同時に表示対象の操作や指示を行うことが可能になった。

DiamondTouch の表面は、それぞれがユニークな信号を発信する送信アンテナのアレイになっている。ユーザは座布団型などの受信アンテナ端子に触れておく。どのユーザがどこに触っているかは、受信された信号を解析することで判別可能である。

開発キットによるアプリケーション例としては、一面に表示された衛星地図上の指示したポイントに、観光情報や道路情報などそれぞれのユーザが選択した情報を表示するものなどが示された。

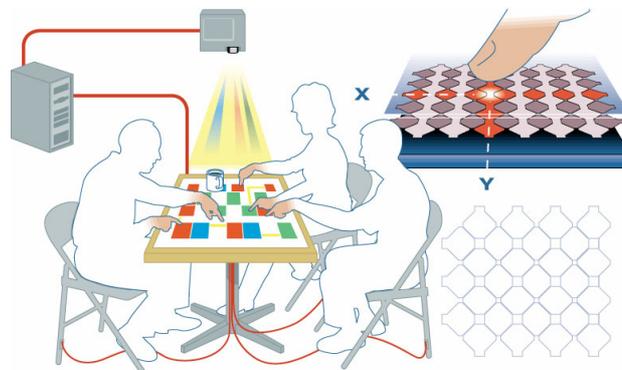


図 4 DiamondTouch 使用イメージ

## Encountering Awareness Information with GroupSense Displays and Tools

Andreas Dieberger, Stephen Farrell, Beverly Harrison, Eser Kandogan, Thomas P. Moran, and Barton A. Smith, IBM Almaden Research Center, USA and Bogdan Dorohonceanu, Rutgers University, USA

pp.95-98

本発表が提案する GroupSense システムは中規模ワークグループにおける出会いアウェアネスを支援するものである。協調的活動を行うグループ内メンバがその所在情報を共有することによる作業上の利便性向上を目的としている。

システムは、ユーザによる IN(出勤)/OUT(帰宅)の申告からアウェアネス情報を生成する。本システムは、インスタントメッセージなどの既存ソフトと異なり、情報のプッシュ型配信を行わない。環境に設置された専用ディスプレイ(図 5 参照)やウェブページにアウェアネス情報を表示し、訪問者がそこを参照することで気づきを促すという手法を採る。

150cm x 90cm の大型プラズマディスプレイでは、メンバの在勤状況を魚の泳ぐ様子で表現するといった手法がとられた。またウェブページでは、ユーザが記した連絡先や今日の予定といった情報も共有することで、より利便性を増すという方策がとられている。

本システムの評価は行われていないが、システム停止時に得たクレームなどから、ユーザにとって無くてはならないツールとして成功したであろうとされている。



Figure 2. GroupSense fish tank showing awareness information in a pleasant natural visualization.



Figure 3. GroupSense awareness information incorporated into a colleague's door panel.

図 5 設置型 GroupSense ディスプレイの例

#### [Interactive Posters インタラクティブポスタ発表]

##### **SitComm: Situation-Aware Interpersonal Communication**

Kristine Nagel, Georgia Institute of Technology, USA, and Gregory Abowd, GVVU - Georgia Institute of Technology, USA

pp. 191-192

本発表では、著者らが進める **Aware Home** プロジェクトの一部として、状況アウェアネスを備えたインターコム(内線電話機)の提案を行っている。

既存の内線電話の特徴は、同期分散型環境を対象に音声による軽快なインタラク션을実現しているところにある。装置の軽快さと利便性を追求する手法として、著者らはこれに状況アウェアネス機能の付加を行った。

具体的には、誰がどこに居るかといった状況の一覧表示が可能な **Interactive InOut Board** を、インターコムの対話相手先選択端末として備えた。相手先を選択すればすぐに繋がれるという軽快さを持ちつつ、相手がどこにいるかというアウェアネス情報を得ることができる。本装置は、**Aware Home** プロジェクトの中で使用され、実際の長期評価などは今後行われる計画である。

##### **Communicating through Handheld Wireless Tablets: Livenotes and Shared Group Awareness**

Matthew Kam, Orna Tarshish, Dan Glaser, Alastair Iles, and John Canny, University of California, Berkeley, USA

pp. 143-145

本発表において Kam らは、アウェアネス機能を備えた共有白板 **Livenotes**(図 6 参照)を提案している。これは小集団対象の同期分散型グループウェアであり、無線で繋がれた携帯型タブレット上に実装されている。共有白板としての機能では、複数の書き込み可能なページを持ち、各ユーザは自由にページを選んで書き込みを行うことで、個人のメモや複数人での共同作業に使うことが出来る。

共同作業環境では、他ユーザの存在や作業状況に対する気付きであるワークスペースアウェアネスが重要と言われている。**Livenotes** では、利用ページをページ数を表すナビゲーションバーから選択する方式を取っており、このバー上に他ユーザの書き込み状況をページ毎にアイコン化して表示することで、作業状況に対するアウェアネスを支援している。

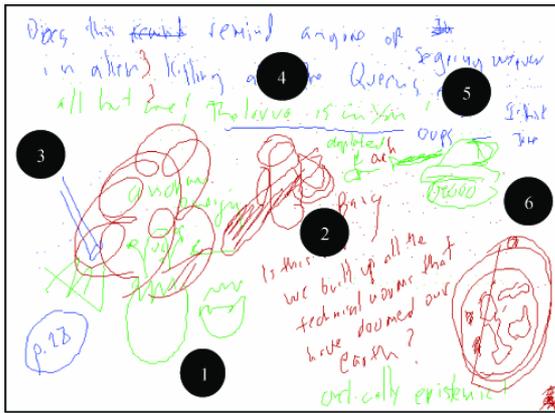


Figure 2: Screen-shot of visual dialogue between 3 users.

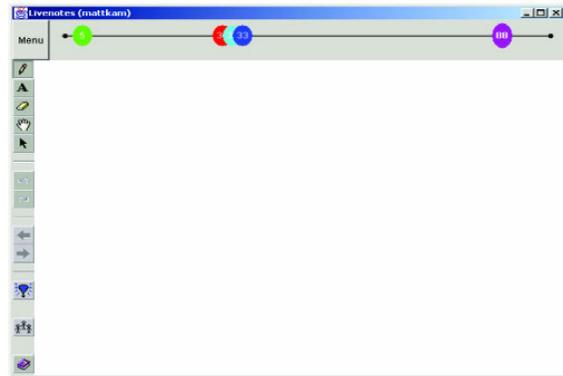


Figure 3: Livenotes with awareness widget.

図6 共有白板 Livenotes の画面イメージ

### 3.1.2.2 アウェアネス関連技術に関する研究発表

[Papers ペーパー発表]

#### Improving Interpretation of Remote Gestures with Telepointer Traces

Carl Gutwin and Reagan Penner, University of Saskatchewan, Canada

pp.49-57

ジェスチャーによるコミュニケーションは対面および分散環境の両者において重要な役割を果たしている。しかしながら、グループウェアにおけるジェスチャーはネットワークのジッターによって引き起こされる動きの崩壊の故に、しばしば見て、解釈するのが難しい。遠隔ジェスチャーに見やすさを改良する一つの方法はトレースを用いることである。即ち、遠隔ポインタの動きの最後の数モーメントの可視化である。本研究は人々がジェスチャーを解釈するのに役立つトレースの効果に関する評価実験を行なった。我々はジッターの遅れがだんだん大きくなる時、テレポインタのトレースが人々の意思決定に関する正確さと確信の度合いを動的に改良するのを発見した。このことは遠隔共同作業において、テレポインタ・トレースやインタラクション履歴の可視化がコミュニケーションを豊かにすることを示唆している。

#### Developing CSCW tools for Creativity -Empirical Findings and Implications for Design

C. Magerkurth, T. Prante, and N.A. Streit, Fraunhofer IPSI, Germany

pp.106-115

CSCW 研究者では世界で最も発想支援に近い研究を行なっているのがドイツの GMD グループである。彼らが分散グループでのアイデア発見支援ツールに関する経験的学習を行い、CSCW ツールの設計に活かすための知見をまとめ報告した。学生の集団を異なる仕

事の様式で、空間配置と白板ツールを用いる創造的問題解決に従事させた。学習結果を分析することで、多くの要求が演繹された。これら要求によって知らされるツールの組は典型的な利用のシナリオにそって提供されるべきである。すなわち、それはマインドマッピングシステム BeachMap、アイデアを継続的ボトムアップ構造化する革新的なインタラクション技法 MagNets、および旅行中も使える非同期アイデア生成のための PDA ツール PalmBeach という三つのソフトウェアコンポーネントからなる。

#### Creating Assemblies: Aboard the Ghost Ship

J. Hindmarsh, C. Heath, D. vom Lehn, King's College London, U.K. and J. Cleverly, Staffordshire University, U.K.

pp.156-165

シカゴの SOFA 展覧会に展示され、大好評を博したインタラクティブアート幽霊船(図7参照)のデモ展示結果を報告した論文である。伝統的なオブジェクトとビデオ技術の組合せに遭遇した人々が、それらをどのように理解するかについて、作品とデモ見学者とのインタラクションの膨大な視聴覚記録を取った。

人々の行動とインタラクションの分析結果は美術館やギャラリーにおいてインフォーマルなインタラクションや社会性を生じさせる技術的組合せ・研究開発を示唆する設計方針を得るのに利用できる。デモ展示会場における見学者の視線の記録から、インタラクションと共参画に関する幾多の設計方針を得たことが興味深い。

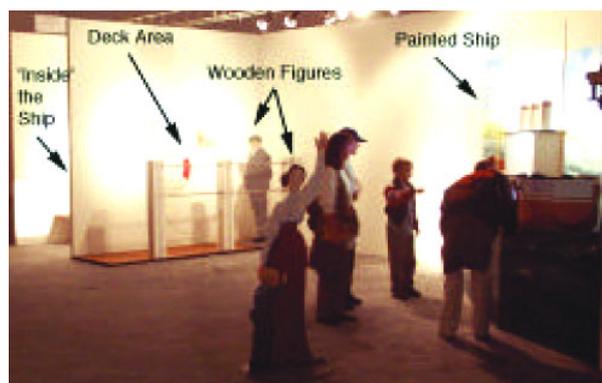


Figure 1: Ghost Ship

図7 Ghost Ship デモの様子

## Sharing and Building Digital Group Histories

C. Shen, N. B. Lesh, F. Vernier, C. Forlines, and J. Frost, Mitsubishi Electric Research Laboratories, USA

pp.324-333

組織、家族や研究機関では文化や歴史を共有・発展させることが多い。本研究では、こういった過ぎ去った集積的出来事についての会話や物語を促進するシステムについて述べる。ユーザはテーブルトップ・インタフェース上の写真、ビデオ、テキスト文書といった分散したマテリアルのデジタルアーカイブを活用する。ソフトウェアとインタフェースの双方が自然な会話とリフレクションを助長する。この研究は多数の共存するユーザがデジタルコレクションを開発するための進行中のプロジェクトの一応用である。プロジェクト名はPDH(Personal Digital Histirian)で、専用のテーブルとインタフェース(図8参照)が開発されている。各種の応用を意図したソフトウェアシリーズが事例研究として提案されていたが、それらは新しいテーブルウェアとでもいうべき研究分野を示唆しており、極めて興味深かった。



Figure 1: The PDH table.

図8 PDH テーブル使用の様子

### 3.1.3 まとめ

本章では、最近のグループウェアや CSCW 研究における主要研究課題の一つであるウェアネス技術について、CSCW2002 の調査結果を中心に、その研究開発動向を述べた。昨年の報告書[18]に記したように、我々は既に Web 探索ウェアネス[19,20]、およびカンバセーションウェアネス[21]を実現する環境を構築した。位置情報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・ウェアネス環境の構築事例[22,23]を試作した。更に位置情報ウェアネスサーバを利用した情報環境を、JAIST 知識科学研究科棟内に知識創造支援システム[24]として構

築し、各種応用システムを研究開発中[25,26,27]である。また宮原のように視覚・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の臭覚アウェアネスの研究[28]や宗森らの効果音・振動アウェアネス研究[29]が興味深い。

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。アイコンタクトの提供という素朴な研究から出発し、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時、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、新世代のマルチメディア・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。

#### [参考文献]

- [1] 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、2001年7月。
- [2] P. Dourish and S. Bly: Supporting Awareness in a distributed Work Group, in Proc. of CHI'92, pp.541-547, ACM, 1992.
- [3] 石井 裕：グループウェアのデザイン、共立出版、1994。
- [4] 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.
- [5] 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。
- [6] 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.
- [7] 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.
- [8] 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.111-121,1999年1月号。
- [9] 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.2820-2827,1998年10月。

- [10] T. Ishida(ed.): Community Computing and Support Systems, Springer, Lecture Notes in Computer Science 1519, 1998.
- [11] 宮原 誠: 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研究成果報告書、理工領域-6 マルチメディア高度情報通信システム「未来映像音響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」、2002年3月.
- [12] 林 秀彦、國藤 進、宮原 誠: 高品位映像の評価ー脳波を指標とする客観評価法ー、映像情報メディア学会誌、Vol.56、No.6、pp.954-962, 2002年6月号.
- [13] 椎尾一郎: 基礎講座 ユビキタスコンピューティング 第1回 ユビキタスコンピューティング@ホーム、ヒューマンインタフェース学会誌、Vol.4 No.3, pp.123-130, 2002年3月号.
- [14] 下條信輔: サプリミナル・マインド、中公新書、1996.
- [15] 荳阪直行編: 脳と意識、朝倉書店、1997.
- [16] ACM2002 Conference Proceedings of CSCW2002, New Orleans, Nov. 16-20, 2002.
- [17] ACM2002 Conference Supplement of CSCW2002, New Orleans, Nov. 16-20, 2002.
- [18] 日本情報処理開発協会先端情報技術研究所: 人間主体の知的情報技術に関する調査研究V、平成14年3月.
- [19] R. Sakamoto and S. Kunifuji: Collaborative World Wide Web Browsing System through Supplement of Awarenesses, Proceedings of KES'2000 Vol.1, University of Brighton, pp. 233-236, 31 August, 2000.
- [20] 坂本竜基、國藤 進: 創造的な Web の利用を支援する協調的 Web ブラウジングシステム、日本創造学会論文誌第6号、2002年12月27日.
- [21] S. Ito and S. Kunifuji: Supporting Conversational Awareness in Text-based Conference System, Proceedings of KES'2000 Vol.1, University of Brighton, pp. 221-224, 31 August, 2000 .
- [22] 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、第17巻3号、pp.301-312、2002年5月.
- [23] 坂本竜基、角 康之、中尾恵子、間瀬健二、國藤 進: コミックダイアリー-漫画表現を利用した経験や興味の級友支援、情報処理学会論文誌、第43巻第12号、pp.3582-3595、2002年12月.
- [24] 山下邦弘、國藤 進、西本一志、伊藤孝行: 知識創造キャンパスの実現、サイエンティフィック・システム研究会編、SS研究会ニューズレター選集2002、pp.61-72、2002年5月.
- [25] 若江智秀、小林 薫、藤波 努、國藤 進: 公開型コミュニティ指向メッセージャー

による実世界コミュニティの活性化、情報処理学会第 64 回全国大会、特別トラック「グループウェアとネットワークサービス」優秀発表賞、東京電機大学、平成 14 年 3 月 12-14 日。

- [26] 山下邦弘、國藤 進、西本一志、伊藤孝行、宮田一乗：知識創造ビル内位置情報ウェアネスサーバの設置とその応用 -追跡型情報掲示板システムの構築-、情報処理学会第 46 回グループウェアとネットワークサービス研究会、機械振興会館会議室、2003 年 1 月 15 日。
- [27] 森田篤志、山下邦弘、國藤 進：インタレスト・コンシェルジュ：“待ち状況”に共通興味を案愛する情報サービス提供サービスシステム、情報処理学会インタラクシオン 2003 ポスター発表、2003 年 2 月 27-28 日。
- [28] 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、1999 年 3 月。
- [29] 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.25-30,2001 年 3 月。

## 3.2 次世代 Web とエージェント

大須賀 昭彦 委員

### 3.2.1 はじめに

Web サービスやセマンティック Web など、機械が Web コンテンツの意味を理解した上で処理することを可能にする、新しい Web の世界を目指した研究が進められている。次世代 Web と呼ばれるこの新しい Web の世界では、人間の代理人となって Web 上のさまざまな仕事を処理するソフトウェアエージェントが重要な役割を果たす。そこで、本章では、エージェント技術研究の最新動向をサーベイした後で、Web サービスやセマンティック Web などの次世代 Web 技術とエージェント技術との関わりについて触れる。次に、エージェント技術を Web サービス分野に応用した具体事例として、Web サービス向けマッチメーカーを紹介する。Web サービスにおける現状の UDDI では、検索の際にサービスの意味までは考慮しないため、今後、多数の Web サービスが登録された際に、その中から適切なサービスを探し出すのが難しくなる。Web サービス向けマッチメーカーは、この問題を解決するために開発された、Web サービスのための検索エンジンである。Web サービスの提供者が WSDL や OWL などの言語を併用してサービスを定義すると、Web サービス向けマッチメーカーはオントロジーや制約条件などの意味情報を活用し、利用者の要求に合った適切なサービスを探し出す。

### 3.2.2 エージェント技術に関わる最近の動向

#### (1) エージェント技術の標準化

FIPA (Foundation for Intelligent Physical Agents)[1]は、エージェント技術の標準化を進める非営利団体である。1996 年にスイスのジュネーブで登録され、50 前後の企業・大学が参加して活動を続けている。活動の主な目的は、異種エージェントシステム間の相互運用性を確保する標準仕様の策定である。2002 年 12 月に、それまでにまとめられたエージェント通信言語(ACL; Agent Communication Language) やインタラクションプロトコル、エージェント管理などに関する 23 の仕様が FIPA 全メンバーの投票を経て標準仕様(Standard)として承認された[2]。これを受けて、今後、産業・学術分野において FIPA 仕様の活用が盛んになっていくものと予想される。

#### (2) FIPA 仕様に基づくエージェントプラットフォーム

FIPA 仕様を先取りしたエージェントプラットフォームも次々と登場している。FIPA-OS(FIPA Open Source)[3]、JADE(Java Agent DEvelopment Framework)[4]、

JADE を携帯機器向けに軽量化した LEAP(Lightweight Extensive Agent Platform)[5]が有名であるが、この他にも Zeus[6]、AAP(April Agent Platform) [7]、Bee-gent[8]、Grasshopper[9]、microFIPA-OS[10]などが知られている。多くのプラットフォームは Web ページからダウンロード可能である。

### (3) AgentCities プロジェクト

FIPA 仕様に準拠したプラットフォームを用いて、インターネット上でさまざまなエージェントサービスを提供する国際プロジェクト AgentCities[11]が盛り上がりを見せている。エージェントサービスとして各都市(city)の情報を提供することから、この名が付けられたものである。グローバルにプラットフォームを管理する Platform Directory[12]を見ると、現時点で 100 前後のプラットフォームが登録されており、国内からも AgentCity Tokyo と Sendai がこれに参加している。このプロジェクトでは、当面はエージェントサービスの相互運用技術の開発を進めるが、将来的にはセマンティック Web や Web サービス、グリッドコンピューティングなどとの連携を図り、ダイナミックで知的かつ自律的なエージェントの実現を目指していく。

### (4) セマンティック Web とエージェント

セマンティック Web とは、Web 上のコンテンツの意味を取り扱う技術である。これにより Web コンテンツの意味を機械的に扱うことが可能となるため、エージェントが文字通り人間の代理人となって、Web 上のさまざまな仕事を処理することが期待される[13,14,15]。このような知的エージェントが Web コンテンツの意味を理解する上で、オントロジーが重要となる。オントロジーを記述する言語として、W3C(World Wide Web Consortium)では当初 DARPA で開発された DAML (DARPA Agent Markup Language)[16] と EU で開発された OIL (Ontology Inference Layer)[17] を統合した DAML+OIL[18] の採用が検討されたが、その後、これをベースとする新たな言語 OWL(Web Ontology Language<sup>1</sup>)[19] が提案され、今後は OWL 中心に検討が進むものと思われる。

### (5) Web サービスとエージェント

XML や SOAP といった技術をベースに、Web サービスは急速に広がりを見せている。Web サービス技術に関しては、W3C[20]や WS-I (Web Services Interoperability

---

<sup>1</sup>略称が WOL ではなく OWL となったことにより、文献によってはこれを Ontology Web Language と呼んでいる。

Organization)[21]などで仕様の標準化が進められている。このような Web サービス分野にセマンティック Web やエージェント技術を応用する試みがある。Web サービスのためのオントロジー DAML-S(DAML Services)[22]の活動などがその例である。現状の DAML-S は DAML+OIL に基づいているが、OWL 版の登場も遠くないと思われる。また、サービスの意味を扱う Web サービスのための検索エンジンも開発されている。次節では、Web サービス向けマッチメーカーと呼ばれる Web サービスのための検索エンジンを紹介する。

### 3.2.3 Web サービス向けマッチメーカー

#### (1) Web サービスの課題

人がブラウザを介してインターネット上のコンテンツを閲覧する際、コンテンツの検索には Yahoo!や Google などの検索サイトが利用される。これらの検索サイトで使われている技術は、基本的には入力されたキーワードを含む Web 上のコンテンツを探し出すというものである。しかし、こうした技術は Web サービスの検索にそのまま使うことはできない。

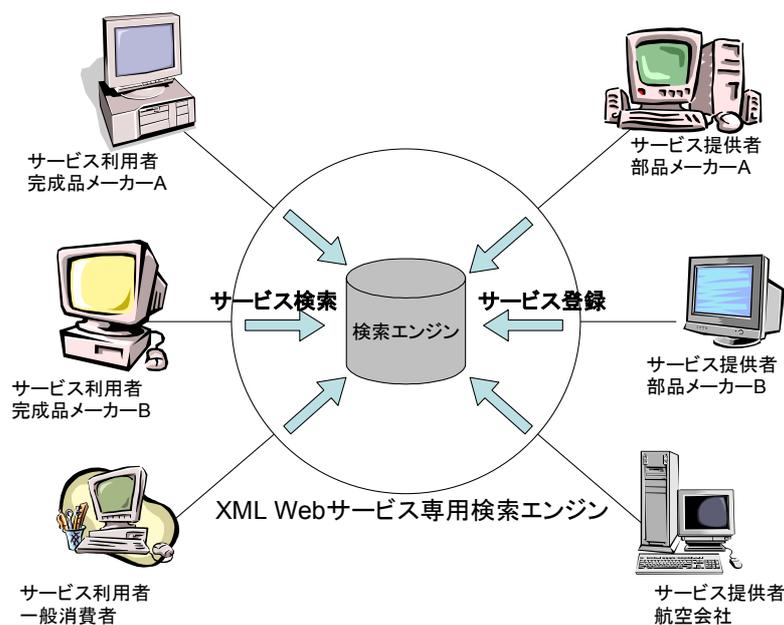


図1 Web サービスにおける検索エンジンの位置づけ

例えば、完成品メーカーが「ネジ」を 1,000 個発注しようと部品メーカーの電子商取引サイトを探すような場合を考える。必要なネジの個数に相手先が対応可能かどうか、納期や配送地などの条件はこちらの希望に合うかどうかなど、単純なキーワード列では指定できないことが多い。また、部品メーカーのサイトで「ネジ」ではなく「Screw」という言葉を使っていた場合は、キーワード検索ではこのサイトを見つけることすらできなくなる。

そこで、Web サービスには、従来の検索エンジンとは異なる専用の検索エンジンが必要となる(図1)。

## (2) Web サービス向けマッチメーカーの概要

Web サービスとは、簡単に言えば、インターネット上で SOAP (Simple Object Access Protocol)[20]と呼ばれる XML 形式のメッセージを使って行なわれるリモートプログラム呼出しである。実際に呼び出されるのは Java や C#で記述されたプログラムであり、そうしたプログラムの呼出しインターフェースを定義するのが WSDL (Web Services Description Language)[20]と呼ばれるサービスインターフェース記述言語である。また、この WSDL をサービス提供者名やサービス名、アクセスポイントに関する情報と一緒に登録するための専用のレジストリが UDDI (Universal Description, Discovery, and Integration)[23]である。SOAP、WSDL、UDDI の3つは Web サービスの基本的な枠組みとすることができ、国際標準仕様の策定も進められている。

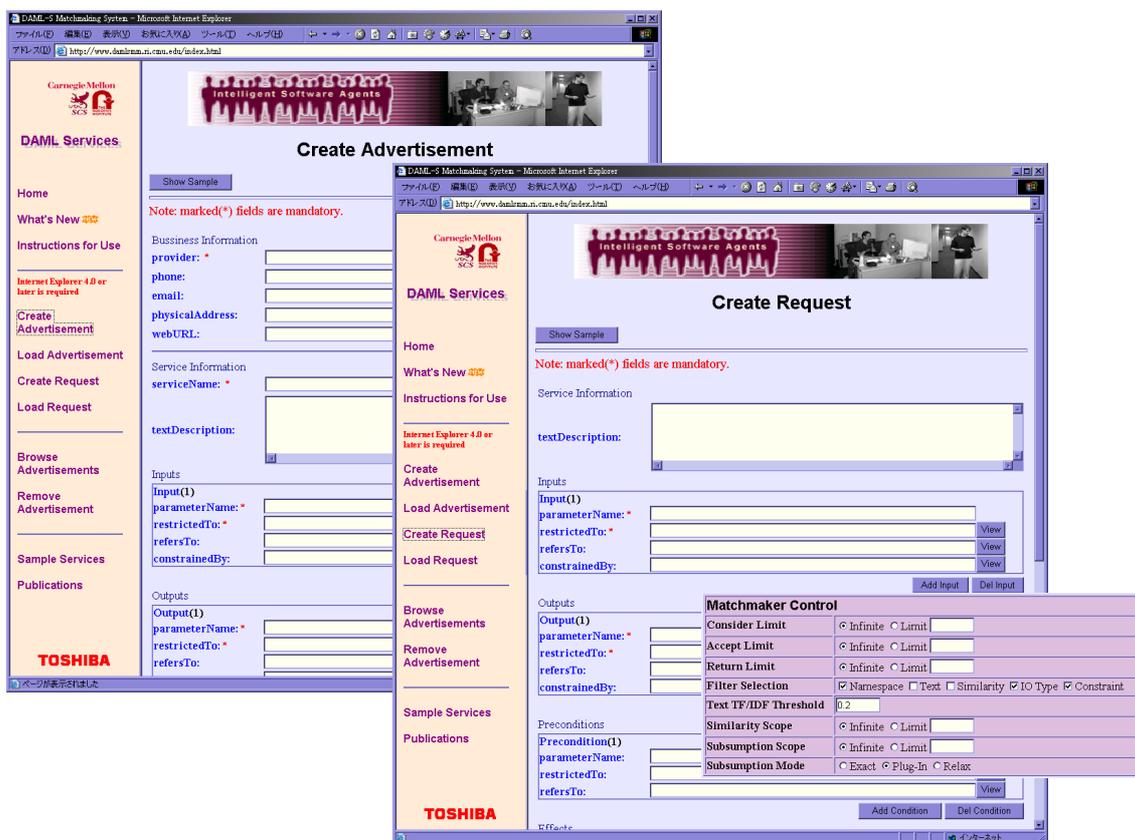


図2 CMUサイトのWeb サービス向けマッチメーカー画面例

しかし、この枠組みでは従来の検索エンジンと同様に、サービス提供者の名前やサービスの名前、サービスインターフェースの名前をキーワードで検索することしかできない。そこで、Web サービスの基本的な枠組みを変えることなく、サービスが提供するもの（「ネジ」や「Screw」など）の意味や、サービスを呼び出す際の条件（注文個数や納期、配送地など）を対象にサービスを検索する技術が開発された[24,25,26,27]。この技術は、東芝が米国カーネギ・メロン大学(Carnegie Mellon University; CMU)のエージェント技術をベースに Web サービス向けに拡張したものである。この技術に基づいて作られたシステムが Web サービス向けマッチメーカーであり、既に図 2 に示すような試用版が CMU のサイトから公開されている[28]。

### (3) Web サービス向けマッチメーカーによる高度なサービス検索

Web サービス向けマッチメーカーでは、サービスのインターフェースにオントロジーや制約条件といったメタ・データを新たに付加している。そして、それらに対し「意味検索」や「制約条件間の包摂関係を判定するアルゴリズム」をフィルターとして適用し、候補となるサービスを絞り込んでいく（図 3）。

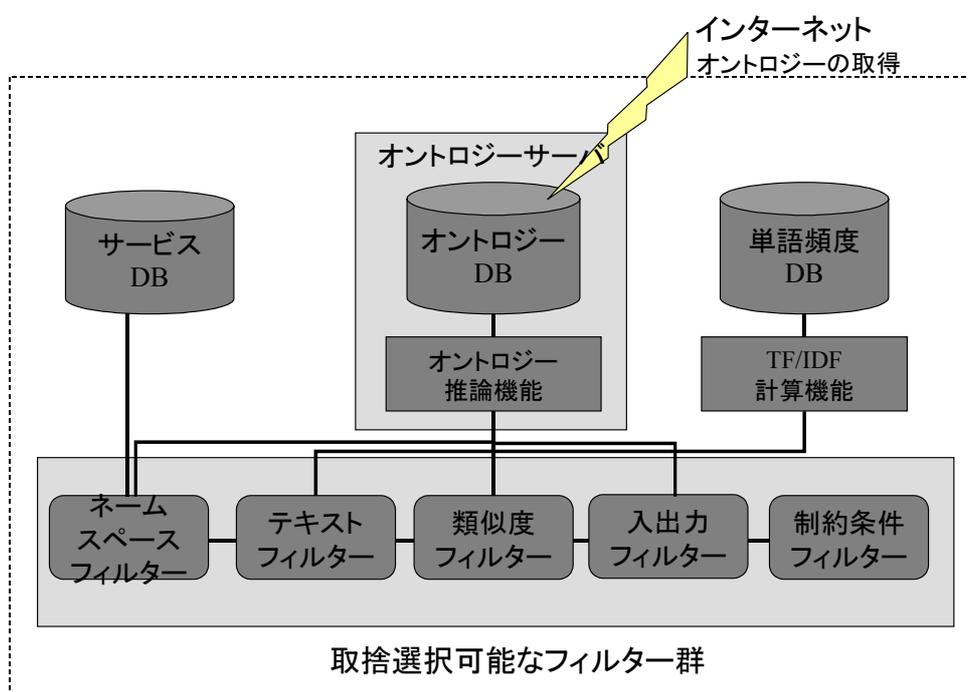


図 3 Web サービス向けマッチメーカーの構成

ここで、メタ・データとは情報の意味や内容を判断するために付加するデータのことである。オントロジーとは一般に語彙体系と呼ばれるもので、意味に基づいて言葉の関係を定義したグラフによって表現される。また、制約条件とはサービスを呼び出す際に満足すべき条件を論理的に記述したものである。以下、簡単な例を用いて説明する。

例えば、従来の UDDI レジストリにおける検索では、「ネジ」を提供している部品メーカーを検索するには、キーワードとして「ネジ」を与え、サービス提供者の名前やサービスの名前に「ネジ」を含むサービスを探すしか方法はなかった。しかし、Web サービス向けマッチメーカーを使えば、図 4 にあるようなオントロジーを参照して「ネジ」と「Screw」が意味的に同一であることを発見し（意味検索）、単語として「ネジ」が使われていなくても対象となる部品メーカーを見つけ出すことができる。また、「部品」や「小ネジ」など「ネジ」から見て上位や下位の概念も探し出し、検索時に指定するオプションによって検索結果に含めることができる。オントロジーを検索する深さや検索の時間も指定できる。

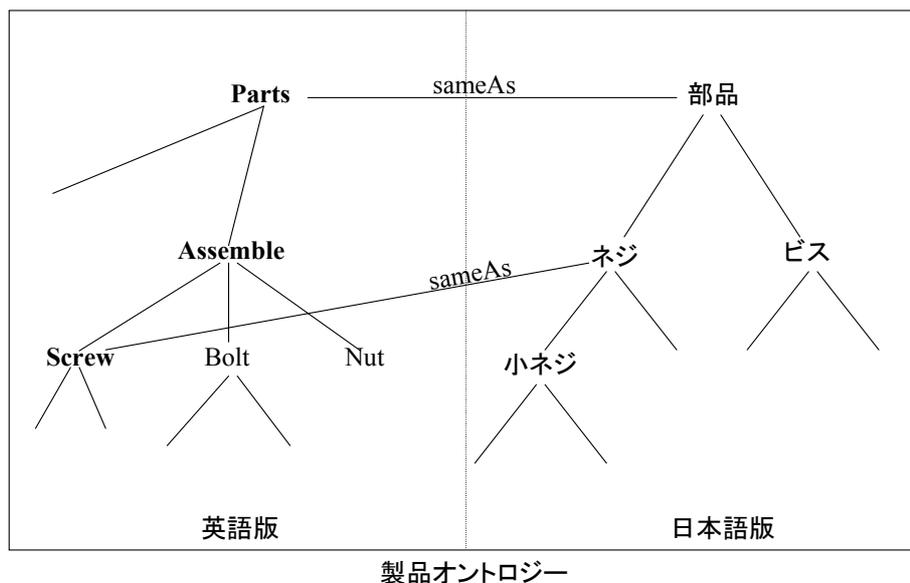


図 4 オントロジーの例

また、従来の UDDI レジストリでは、納期や配送地に関する制約条件を対象にサービスを検索することはできなかったが、Web サービス向けマッチメーカーでは以下のような検索が可能である。利用者が次の制約条件を持つサービスを検索したいとする。

- (1) 注文個数 > 1,000 個
- (2) 配送地 ⊂ 東京
- (3) 納期 ≤ 7 日

仮に次のような制約条件を持つ部品メーカーのサービスが登録されていたとする。

- (a) 注文個数 > 100 個
- (b) 配送地 □ 関東
- (c) 納期 ≤ 5 日

すると、Web サービス向けマッチメーカーは、要求された制約条件(1)~(3)をサービスの制約条件(a)~(c)が満たしているということを論理的に判断し(制約条件間の包摂関係を判定するアルゴリズム)、このサービスを検索結果として返す。

その他にも、利用者がサービス呼出し時に与える情報の種別(品名、注文個数、配送地など)を指定し、その種の情報で呼び出すことができるサービスを検索することも可能である。同様に、利用者がサービス呼出しの結果として得たい情報の種別(納期など)を指定し、その種の情報を返すサービスを検索することもできる。

つまり、WSDL では引数のデータ型や個数は指定できても、引数の情報が持つ意味や出力結果が持つ意味までは定義できなかつた。Web サービス向けマッチメーカーでは、サービスを呼び出す際に与える引数や、サービスの出力結果にオントロジーを付加することで、サービス自体が持つ意味を規定できる。また、WSDL では記述することのできない制約条件を論理的に与えることで、提供されているサービスが自分の要求に合うかどうかを判断することができる。こうした技術を現在の Web サービスの枠組みを壊さずに、UDDI へのアドオンとして組み込んだものが Web サービス向けマッチメーカーである(図5)。

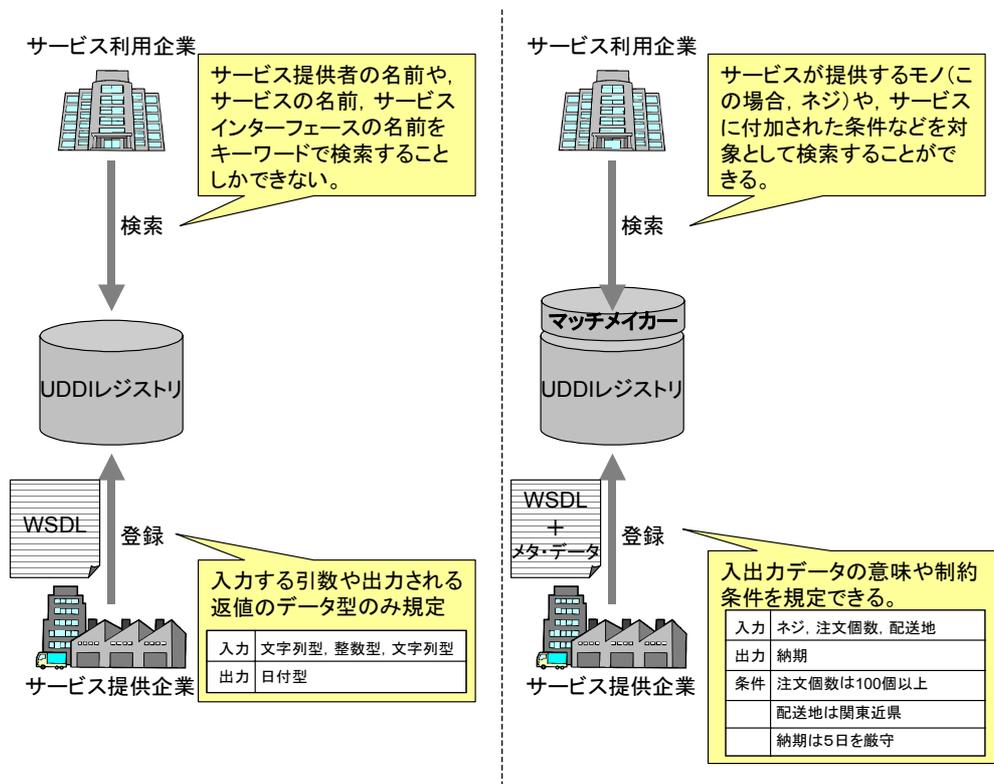


図5 UDDI と Web サービス向けマッチメーカーの比較

#### (4) エージェントによるダイナミックな Web サービス統合

現在の UDDI 利用では、人間が UDDI から返されたサービスのリストを見て、ひとつひとつのサービスインターフェースをチェックし、Web サイトを訪れて要求に合うサービスかどうかを判断することになる。Web サービス向けマッチメーカーを用いることで、このような手間を大きく軽減することが期待できる。また、Web サービス向けマッチメーカーから返されるサービスは、利用者が要求したサービスのインターフェースに整合するものであることが保証されるため、返されたサービスを利用者側のプログラムから直接呼び出すことも可能となる。このため、Bee-gent[8]のようなサービス連携エージェントフレームワークを適用することで、Web サービス向けマッチメーカーを使って探し出した複数のサービスを動的に連携させることも可能になる。

今後、Web サービスはダイナミックにサービスを統合して活用するものに徐々に変わっていくと考えられるが、そこではこの種の技術が重要な役割を果たす。

#### [参考文献]

- [1] <http://www.fipa.org/>
- [2] <http://www.fipa.org/repository/standardspecs.html>
- [3] <http://www.nortelnetworks.com/products/announcements/fipa/>
- [4] <http://jade.cselt.it/>
- [5] <http://leap.crm-paris.com/>
- [6] <http://more.btexact.com/projects/agents/zeus/>
- [7] <http://www.nar.fujitsulabs.com/aap/>
- [8] <http://www.toshiba.co.jp/beegent/>
- [9] <http://www.grasshopper.de/>
- [10] <http://www.cs.helsinki.fi/>
- [11] <http://www.agentcities.org/>
- [12] <http://www.agentcities.net/>
- [13] <http://www.semanticweb.org/>
- [14] 萩野達也 編：特集「セマンティック Web」、情報処理学会誌、Vol.43、No.7 (2002)、pp.708-750.
- [15] 溝口理一郎 編：特集「Semantic Web とその周辺」、人工知能学会誌、Vol.17、No.4 (2002)、pp.383-416.
- [16] <http://www.daml.org/>
- [17] <http://www.ontoknowledge.org/oil/>

- [18] <http://www.daml.org/2001/03/daml+oil-index.html>
- [19] <http://www.w3.org/2001/sw/WebOnt/>
- [20] <http://www.w3.org/2002/ws/>
- [21] <http://www.ws-i.org/>
- [22] <http://www.daml.org/services/>
- [23] <http://www.uddi.org/>
- [24] Kawamura,T., Hasegawa,T., Ohsuga,A., and Yamamoto,J.: Proposal of Semantics-based Web Service Matchmaking, in Proc. International Conference on Computational Intelligence and Multimedia Applications 2001 (ICCIMA01), IEEE (2001) , pp. 87–92.
- [25] 川村隆浩、長谷川哲夫、大須賀昭彦: XML Web サービスのマッチメーカー、東芝レビュー、Vol.58、No.2 (2003).
- [26] Paolucci,M., Kawamura,T., Sycara,K. et al.: Semantic Matching of Web Services Capabilities, in Proc. First International Semantic Web Conference (ISWC2002), IEEE (2002), pp. 333-347.
- [27] Paolucci,M., Kawamura,T., Sycara,K. et al.: Importing the Semantic Web in UDDI, in Proc. E-Services and the Semantic Web Workshop (ESSW 2002) (2002).
- [28] <http://www.damlsmm.ri.cmu.edu/>

### 3.3 IT 技術の法への応用

新田 克己 委員

#### 3.3.1 はじめに

IT 技術の発展が法律分野に大きな変革をもたらしている。1つは、法学教育での変革である。教材を Web で配信することにより、いつでもどこでも教育できる環境ができてきただけでなく、実務経験を通じてしてしか得ることが出来なかったような法律の実技技術が習得できるような高度な教育技術が提言されるようになってきた。特にわが国では来年から法科大学院の制度がスタートする。法科大学院では、法律の未経験者を 3 年で司法試験レベルまで立ち上げる必要があることから、メディアを使った効率の良い教育システムを開発することが急務になっている。

もう 1 つの変革は、紛争の仲裁・調停（いわゆる ADR）における変革である。各国では ADR の増加に対応するため、オンラインによる ADR のシステムが試作されている。わが国でも一部の ADR 機関により、オンラインシステムが作られているが、相談自体は人手によるものであり、今後、相談件数の増大とともに、相談員の不足が予想される。IT 技術により、ADR を支援することが必要になる。

本章では内外における法学教育、および、ADR の研究動向を報告する。

#### 3.3.2 教育支援システム

欧米では、Web 配信や CD-ROM による法律の講義は普及しつつあり、今回調査した法律の知識ベースの国際会議 Jurix2002[10]に併設された Law and Education ワークショップにおいても、ほとんどの発表がそのような法律教材に関するものであった。教材としてのコンテンツがかなり多く開発され、非常に充実したものになっていることは推察できたが、その反面、教育システムとしての興味ある報告はイタリアの DidaLex 以外はあまりされていなかった。

一方、わが国では、e-Learning としての法学教育システムはまだ普及していないが、法科大学院による法学教育の方法を探究するプロジェクトとして、文部科学省科研費の学術創成研究「コンピュータネットワークを用いた法学教育の実践・評価システムの創成」（代表：松浦好治名古屋大学法学部教授）と、特別推進研究「法創造教育方法の開発研究—法創造科学に向けて」（代表：吉野一明治学院大学法学部教授）が教育システムとしては新しい試みを行っている。法科大学院では、単に法律の知識を習得させるだけでなく、法律の実務教育の導入部分を担当することが期待されている。これには、コミュニケーション能力や説明能力を持ち、事実認定などの実務的な体験を持ち、創造性を持った論理展開能力

を持つ法律家を育てるための教育を行うことが必要になる。そのため、単にビデオ教材を Web で配信するような単純な e-Learning システムと異なる高度な教育システムの開発が必要となる。

### (1) DidaLex

イタリアの Istituto di Teria e Tecniche dell'Informazione Giuridica – ITTG)は法学教育のオンライン化について進歩的な試みを行い、注目されている。ここで開発されている DidaLex システムは、法令、判例のみならず、ビデオ教材や音声や法的手続きのチャート(図1)や関連の Web ページを検索する機能を備えた総合的な法学教育のシステムである。このシステムの特徴は、学生はいろいろなページを結合して、独自のハイパーテキストを作り出せることと、テストによる自己評価システムがあることと、システムと学生の対話によって、理解度の低い点についてはフィードバックを行い、そのたびに必要なインストラクションのページを提供されることである。必要なインストラクションを出すためには、何を伝えるかだけでなく、どのように伝えるか、さらには、学生がどのような学ぶかを分析する必要があり、その課題に取り組んでいる。これはオンラインの教育システムに関する普遍的な問題でもある。

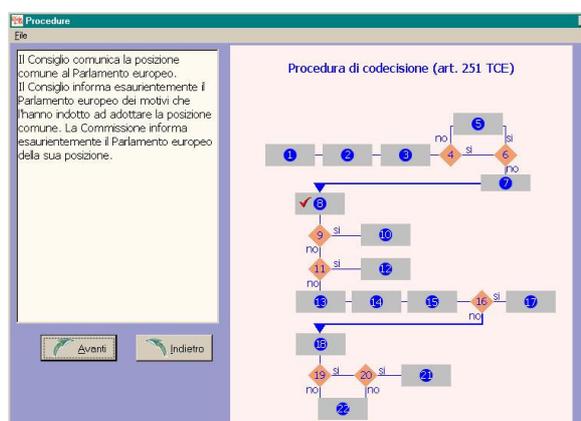


図1 DidaLex システムの画面

### (2) 「コンピュータネットワークを用いた法学教育の実践・評価システムの創成」

名古屋大学大学院法学研究科を中心とした、このプロジェクトでは IT 技術を使った実技指導の研究を、以下のサブプロジェクトに分けて進めている。

1. Web を介した専門家のアドバイス提供、仲間による評価システム
2. Web による理解度確認システム
3. デジタル基本六法プロジェクト

## 4. 映像収録・分析システム

## 5. WebCT カスタマイズプロジェクト

このうちで4番目の収録・分析システムが模擬裁判システムとして実現されつつある。

模擬裁判システムにおいては、現実の法廷に模して作られた演習室に4台のTVカメラがセットされている(図2)。それぞれは、裁判官、原告席、被告席、証言台を撮影している。裁判官席、原告席、被告席、証言台にはマイクがあり、音声感知スイッチコントローラは、現在、音声を発している者を感知して、その発信源のTVカメラの信号を選択して、傍聴者のモニタに表示すると同時にビデオディスクに記録することになっている。

このように模擬裁判の記録を音声と画像で記録することにより、従来にない教育効果を期待することができる。1つは、映像にインデックスを付加することにより、場面に応じた原告や被告の映像を検索することができ、しぐさや表情を分析することによって、法廷論争の技術を身につけることができることである。2つ目は、ネットワークを利用することにより、専門家のコメントを容易に各場面に付加することができ、単なる記録だけでなく、分析結果を含めた教材として利用できることである。また、このような模擬裁判記録は専門家にとっての分析記録としても利用することができる。3つ目は、学生相互のコメントを出すことにより、実践的な授業が期待できることである。

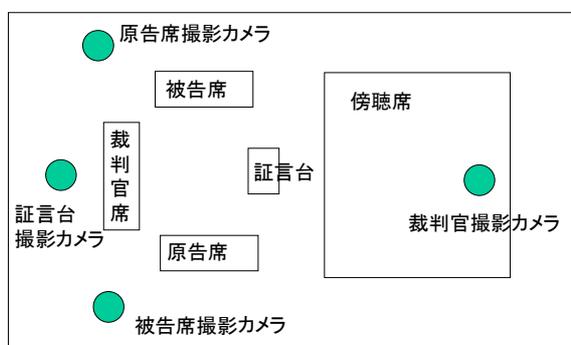


図2 テレビカメラを配置した模擬法廷

IT技術を利用した2つ目の教育システムは、学生の投票を利用した実習である。たとえば、和解シミュレーションにおいては、あらかじめ学生を原告側、被告側、裁判所側に役割分担をさせて、Webページを介して、和解の演習問題の提示と、分担ごとに異なる教材提示を行う。学生は、それぞれの担当に応じて和解計画書を作成し、期限までにWebページに書き込みを行う。書き込まれた和解計画書はどの学生からもWebページからアクセスすることができ、他人の計画書と自分の計画書を比較することができる。Webページであるため、外部の専門家からのコメントを和解計画書に付加することで教育効果を高めるこ

とが可能となる。次に学生同士で投票を行い、多くの票を獲得した学生が代表者となってロールプレイングを行う。ロールプレイングについては、前述の模擬法廷システムを利用する。

### (3) 「法創造教育方法の開発研究－法創造科学に向けて」

明治学院大学を中心とするこのプロジェクトにおいては、法科大学院において、創造力をもつ法律家を育てるため、以下の点で研究を進めている。

1. 法創造基礎の理論的解明
2. 実務と教育における法創造の実際の解明
3. 法創造教育方法の開発
4. 法創造教育支援システム

このうち、1 番目の法創造基礎においては、ゲーム理論、交渉学、経済学などの知見を利用した法制定の効果の分析などを行っている。たとえば、法改正が期待したとおりの効果を生み出すかどうかを、アンケートの分析によって予測するなど、単なる理論ではなく、検証方法まで関連させた法律の設定効果の方法論についても検討されている。

2 番目の法創造の実際の解明においては、アメリカのロースクールの教育の分析を行い、ソクラティックメソッド、ケースメソッド、プロブレムメソッド、ディスカッションメソッドなどを用いて、知識を伝えるのではなく、知識を発見する教育方法の分析を行っている。たとえば、民法の専門書においても難解とされていて、十分な解説がなされていないような最高裁の事例について、学生の反応を電子メールで確認しながら、関連知識を少しずつ与えながら、問題点を徐々に把握させる講義の実験を行い、学生に知識を教え込むのではなく、知識を発見させる教材の開拓を行っている。

4 番目のテーマとして開発されている模擬裁判支援システムは、前述の名古屋大学のシステムと異なり、原告と被告が対面するのではなく、オンラインのシステムを用いて、遠隔地での裁判を想定している(図3)。原告と被告はクライアントとして模擬法廷のサーバに接続し、裁判官エージェントの指揮に従って模擬裁判を行う。したがって、原告と被告は相手の顔を直接見ないで、文字または映像を介してのみの裁判となる。裁判官エージェントは同一課題に関する過去の事例ベースを有しており、類似場面の訴訟指揮事例を用いて、指揮の自動化を試みている。過去に類似事例がない場合には、教師が介入して、裁判官エージェントを制御する。従来の模擬裁判は、時間も場所も限定され、教師の負担も大きいため、学生全員にロールプレイングを体験させることができない。最近では、場所の制約を軽減するため、ビデオ会議システムを応用した遠隔地の模擬裁判も実験されているが、それでもなお、教師の負担は大きい。それに対し、エージェントを用いた模擬裁判支

援システムは、教師の負担を軽くし、できるだけ多くの学生に模擬裁判を予備的に体験させるためことを目的としている。したがって、このシステムである程度の経験を積んだ学生は、次段階として、本格的な模擬裁判へ移行することを想定している。

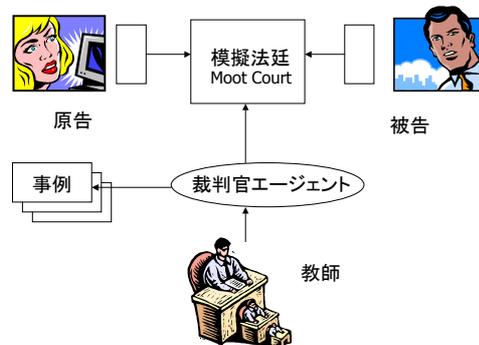


図3 模擬裁判支援システム

### 3.3.3 紛争解決支援システム

ADR(Alternative Disputation Resolution)は裁判によらない紛争解決手段である。たとえば、国民生活センターなどの代表される各地の消費者センターでは商品に対する苦情相談を行っており、助言や専門家の斡旋だけでなく、必要に応じて仲裁や調停を行うところもある。[21]には、わが国におけるADRの実情と課題が詳細に説明されている。

ADRをオンラインで行う試みが内外でもなされている。たとえば、[19]には、アメリカにおけるオンラインADRのサービスが24掲載されている。その多くは、メールを使った相談システムであり、実際は専門家による調停を前提としているものが多い。その中ではCybersettle[14]が和解のための金額の調整のためのコミュニケーションツールを有しているようである。

わが国でも、たとえば、電子商取引協議会(ECOM)ではオンラインによる苦情相談を受け付けている。オンラインであるから、相談者とのやりとりは図4のようにすべて文書の形で記録されることになる。

苦情相談がなされると相談員は事実の確認を行うために、何回かメールのやりとりを行うのが普通である。その際、必要に応じて、関連のWebページにアクセスしたり、相手先に確認のメールを出すこともある。結果として、多くの場合は仲裁や調停にいかずに、助言だけで終わることになるが、その際、かなり多くのケースが類似した対応をとることになる。たとえば、知らない間に海外のWebページに接続され、非常に高額の国際電話料金を請求されたケースは、非常に相談件数が多いが、このようなケースの対処方法は、ほとんどは、

自分の PC を調べ、接続先が変更になっていないかを確認する、

「電話料金の請求を免れることは困難であるが、減額を依頼してみる」と助言する、  
ことすんでしまうものである。

このような典型事例は、相談員を呼び出すまでもないので、多くの消費者センターでは、  
典型事例集を Web ページに記載して、ユーザに無駄な相談を避けてもらうようにしている。

ECOM ではこのような典型事例を構造化し、質問に答えるだけで助言ができる Web ペ  
ージを試作している（図 5）。現在は、このような助言ページの評価と、事例からの助言ペ  
ージの自動生成の検討を行っている。

相談者タイプ： 個人
相手方タイプ： 個人
取引目的物： ノートパソコン
取引目的価格： 150,000 円
取引カテゴリー： ネットオークション
取引目的授受区分： 受領
代金決算区分： 未払い
原因区分： 広告相違
相談内容： オークションでノートパソコンを 15 万円で落札した。送られてきたノートパソコンを見ると、写真には写っていない部分に大きな目立つ傷がついていた。私はこのパソコンを外出先でも使うつもりなので、このような大きな傷は非常に気になります。そこで、出品者に値段の引き下げを求めました。しかし出品者はパソコンの機能自体には何ら問題はないので、値下げ要求には応じられないとのこと、話し合いに一切応じる様子はありません。値下げ要求をしているので、パソコンの代金はまだ払っていません。オークションの主催者にも連絡したが、なんにもならず、本当に困っています。調停に持ち込んで欲しい。
対処： 本件は、当該オークションにおける「出品物の説明と現物の状態」の相違に原因があると判断いたします。調停による解決を希望ということですので、出品者の B 氏の参加意思が確認され次第、調停を行います。

図 4 苦情相談の例

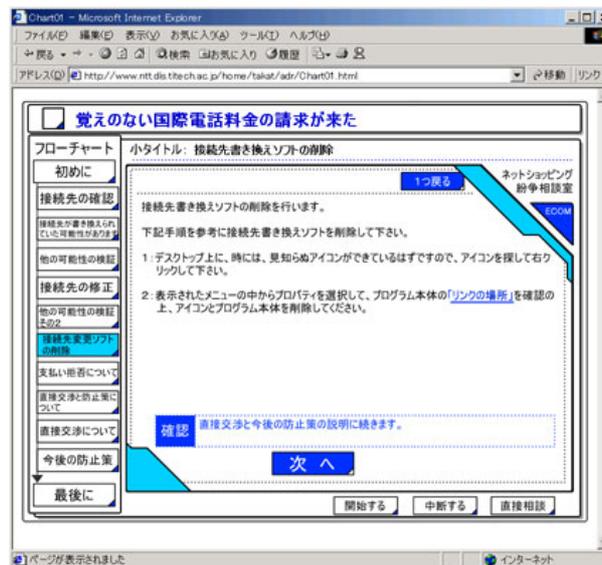


図5 助言の Web ページ

### 3.3.4 関連会議など

#### (1) メディアを使った法学教育に関する学術団体

以下に学術団体の主なものをあげる。ここにあげた学術団体には、メディアに関連しない純粋に法律だけの学術団体は省いた。

#### Law, Education and Technology

BILETA	British & Irish Law, Education and Technology Association
CALI	Center for Computer-Assisted Legal Instruction
LETA	Law & Educational Technology Association
LCC	Law Courseware Consortium
	Law and Technology
IAAIL	International Association for Artificial Intelligence and Law
ISLAT	International Society for Law and Technology
JURIX	Foundation for Legal Knowledge Based Systems Education and Technology
AACE	Association for the Advancement of Computing in Education
ALT	Association for Learning Technology

EDUCAUSE	Association for Transforming Education Through Information Technology
IATED	International Association of Science and Technology for Development
IAIED	International Artificial Intelligence in Education Society
IEEE/LTTF	Learning Technology Task Force
IMS	Global Learning Consortium Inc.
OTEC	Educational Technology Expertise Centre of the Open University of the Netherlands
SURF	Dutch Cooperative Organisation Network Services and ICT for Institutions for Higher Education and Research
PROMETEUS	European Partnership for a Common Approach to the Production of e-learning Technologies and Content

上記のうちで、IAAIL の主催する隔年の国際会議 ICAIL(International Conference on AI and Law)、JURIX が開催する国際会議、BILETA や LCC らが the Law Courseware Consortium(LCC)と共催する隔年の国際会議 SubTech は注目する必要がある。ただし、SubTech は招待者のみしか発表することはできない。

## (2) 情報ネットワーク学会

日本での関連学会としては、2002年に発足した「情報ネットワーク学会」が活発な活動を行っている。この中で、「サイバーコート研究会」は、

1. 司法の IT 化の可能性を探り、問題点を研究する。
2. 国内外の機関とインターネット経由の遠隔司法の実証実験を進めるとともに、その問題点を研究する。特に、従来のテレビ会議システムのみでなく、IPNet を使ったシステムの実証実験を進める。

の2つを目的とし、「ODR 研究会」は、

オンライン技術を用いた裁判外紛争処理手続 (Online Dispute Resolution) について、技術面と手続面の双方からのアプローチにより、実現可能性とその問題点を明らかにしていく。法的理論面での検討ももちろんだが、オンライン上のセキュリティ、自動処理の可能性、実証実験などの技術面の研究も重視する。

を目的とするので、本報告で述べた法学教育や ADR への IT 技術の応用と重なるところがある。今後の成果が期待される。

### (3) 法とコンピュータ学会

「法とコンピュータ学会」においても、「インターネットと ADR」が 2001 年の研究会の報告テーマとなっている。

#### [参考文献]

- [1] <http://www.Iri.jur.uva.nl/LETA>
- [2] <http://www.vuw.ac.nz>
- [3] <http://www.cooley.edu>
- [4] <http://rocket.vub.ac.be/>
- [5] <http://w3.idg.fi.cnr.it/didalex/>
- [6] <http://www.uninettuno.it/nettuno/index.htm>
- [7] <http://www.ijcai-03.org/1024/index.html>
- [8] <http://www.ecom.or.jp/adr/program.htm>
- [9] <http://homepage2.nifty.com/civilpro/adr/koza/>
- [10] <http://www.dcs.kcl.ac.uk/staff/aspasia/jurix02/>
- [11] <http://www.iadb.org/mif/eng/conferences/presentations/gelinas/>
- [12] <http://www.wm.edu/law/courtroom21/legalsystem/>
- [13] <http://personal.law.miami.edu/~froomkin/articles/adr-slides/>
- [14] <http://www.cybersettle.com/flash.htm>
- [15] <http://www.courttv.com/>
- [16] <http://www.law.washington.edu/ABA-eADR/home.html>
- [17] <http://www.adr.gr.jp/>
- [18] [http://www.kokusen.go.jp/jcic\\_index.html](http://www.kokusen.go.jp/jcic_index.html)
- [19] [http://www.cibertribunalperuano.org/enlaces\\_todo.htm](http://www.cibertribunalperuano.org/enlaces_todo.htm)
- [20] <http://www.clicknsettle.com/>
- [21] 「ADR の現状と理論」ジュリスト No.1207,2001.



## 3.4 WWWサーチエンジンの最新技術

山名 早人 委員

### 3.4.1 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 Google(<http://www.google.com/>)や FAST(<http://www.alltheweb.com/>)など数々の WWW サーチエンジンが存在する。2002 年末時点で世界中に存在する総 Web ページ数は、約 70 億ページと推定<sup>1</sup>され、年率 100% の勢いで増加を続けている。

Web ページの中から自分が必要とする情報を見つけ出すためには、Google に代表される WWW サーチエンジンが必要不可欠となっている。しかし、Google が検索対象とする Web ページ数は、2002 年末時点で約 31 億ページ、世界第 2 位の FAST は約 21 億ページであり、これらの現存する WWW サーチエンジンを用いても世界中の全 Web ページを対象にした検索はできない。また、WWW サーチエンジンに入力されるキーワードは平均 2 語であり、2 語という限られた情報を用いて、ユーザが欲する Web ページを見つけ出すことは、一般的に困難である<sup>2</sup>。このような問題に対して、WWW サーチエンジンは、①膨大な Web ページの効率的な収集とインデックス化、②検索結果のランキング、という 2 つの重要な技術を利用している。

以下では、まず、インターネットの発展に関わる統計データを紹介した後、これらの技術についての最新技術を紹介する。

### 3.4.2 インターネットの発展

図 1、図 2 に示すようにインターネットに接続するコンピュータ台数は毎年増加を続けている。1998 年から 3 年間の平均でみるとコンピュータ台数、Web サーバ数は、それぞれ年率約 1.5 倍、2 倍で増加を続けている。ただし、2002 年 1 月までと 2002 年 1 月以降での傾向に変化が見られ、図 3 に示すようにインターネットに接続するコンピュータ台数の増加率と Web サーバ台数の増加率が逆転し、2002 年 1 月以降、Web サーバ台数の増加

<sup>1</sup> NEC 北米研究所の Lawrence らが 1999 年 2 月時点で約 8 億ページが存在すると推定している[1]。この時点での各サーバが持つ Web ページ数は平均 186 ページであった。一方、2002 年末時点の Web サーバ数は、Netcraft 社 (<http://www.netcraft.co.uk/>) の発表によれば 3812 万台であり、この台数と 186 ページ / Web サーバとの積をとることにより、世界中の Web ページ総数を約 70 億と推定した(テキストページのみ)。なお、Web ページの平均容量は 10KB / ページであり、テキストのみでも 70TB の容量となる。

<sup>2</sup> スペースを語の区切りとした場合。したがって、日本語の場合は入力されたクエリーが、形態素解析等によりさらに複数の語に分割されるため、サーチエンジン側から見ると入力される語の平均は、2 語以上となる。

率が鈍化してきていることが分かる。

### 3.4.2.1 Web ページの特徴

2000年のIBM, SUN, Compaqの研究者らの調査[2]によれば、Web ページは、図4に示すような構造（蝶ネクタイ構造）をしている。相互にリンクを張り合うWeb ページ（図4中のSCC）は、全体の28%であり、outlink<sup>3</sup>のみを持つページ（図中IN）と inlink<sup>4</sup>のみを持つWeb ページ（図中OUT）がそれぞれ22%、22%であり、その他のリンク的には中心部分（図中のSCC, IN, OUT）と関連がないページが全体の28%である。

したがって、単純にリンクのみを辿ることで、全てのWeb ページを収集することはできない。これに対して、一つの解決手段は、Netcraft社(<http://www.netcraft.co.uk/>)がWebサーバ数を調査する手法として採用しているように、「IPアドレスを順番に直接指定しhttpプロトコルの標準ポートである80番ポートをアクセスする手法」であるが、Webサーバが80番ポート以外で立ち上がっている場合や、Webサーバのトップディレクトリからのリンクが存在しない場合等は、やはり、収集できないページが発生する。

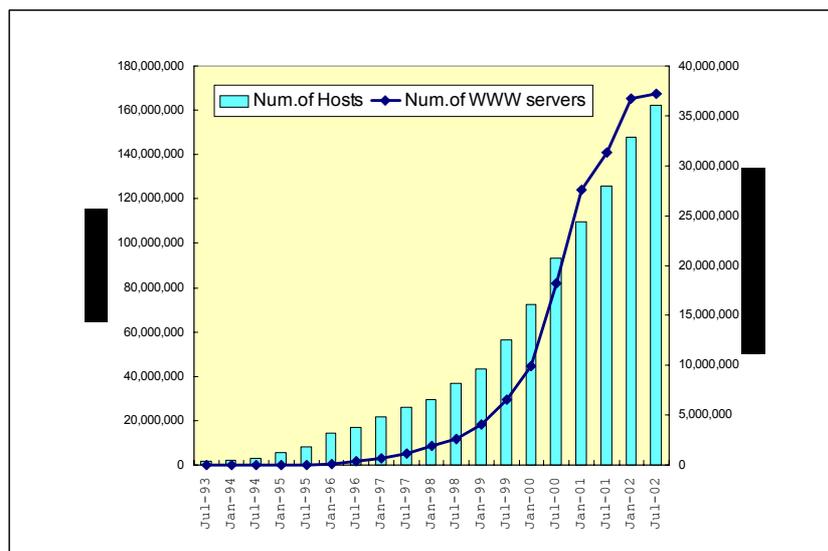


図1 インターネットに接続するコンピュータ台数とWebサーバ数の推移<sup>5</sup>

<sup>3</sup> 当該ページから他のページへのリンク

<sup>4</sup> 他のページから当該ページへのリンク

<sup>5</sup> Internet Software Consortium(<http://www.isc.org/>)のInternet Domain Survey及びNetcraft社(<http://www.netcraft.co.uk/>)のWWW Server Surveyの公開データから作成した。

### 3.4 WWWサーチエンジンの最新技術

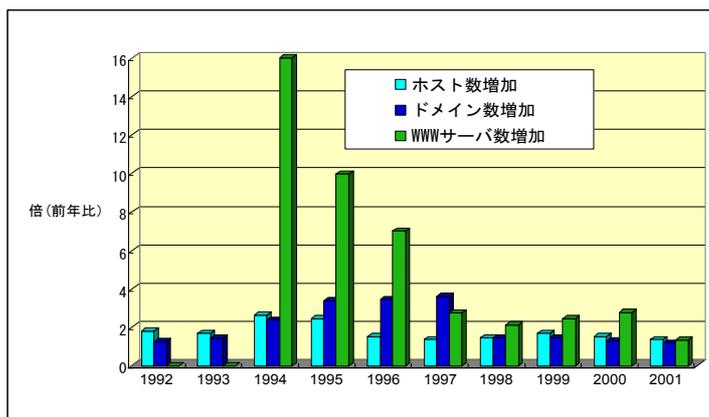


図2 インターネットに接続するコンピュータ台数、ドメイン数、Webサーバ数の増加率<sup>5</sup>

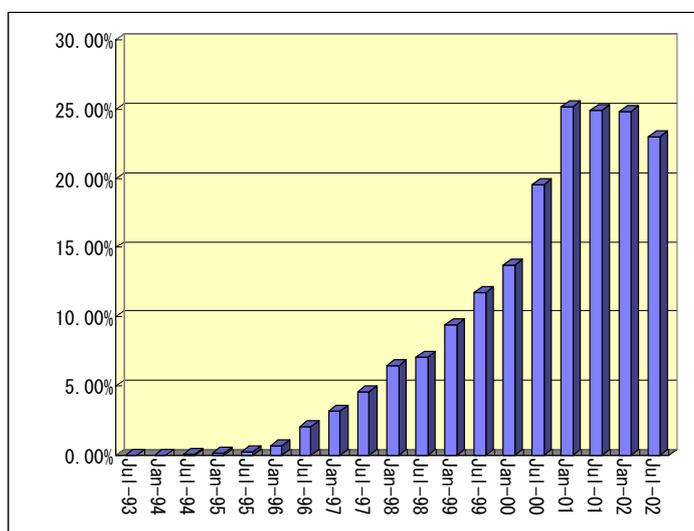


図3 インターネットに接続するコンピュータ台数に占めるWWWサーバ台数<sup>5</sup>

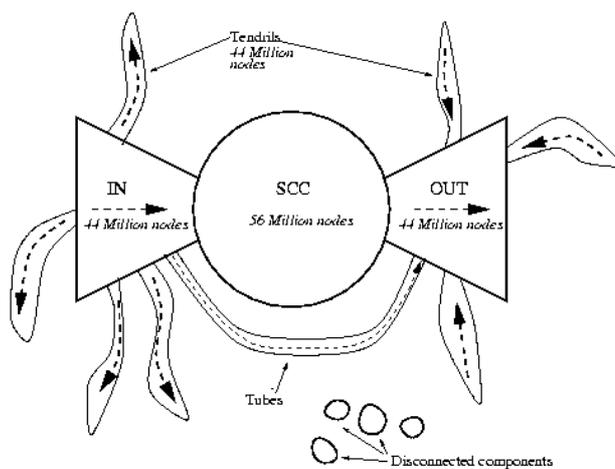


図4 Web ページの特徴[2]

### 3.4.2.2 世界と日本の Web ページ数推計

2002年7月時点でのドメイン毎の Web ページ数(URL数)の推計を表1に示す。各ドメインの割合は Network Wizards 社(<http://www.nw.com/>)のデータに基づいている。また、Lawrence 氏が提案した手法[1]に基づき、2002年7月現在での全世界の Web ページ数を約 70 億 URL と仮定し、各ドメイン数の比率から各ドメインの URL 数を算出した。また、1 ページのバイト数を 10KB と仮定し、各ドメインにおける総バイト数を算出した。

表 1 全世界でのドメイン別 Web ページ数の推計(2002年7月)

	ドメイン名	全ドメイン中の割合	総ページ数(億 URL)	サイズ(GB)
net	Networks	32.4%	22.68	22,680
com	Commercial	30.0%	21.00	21,000
edu	Educational	5.0%	3.50	3,500
jp	日本	5.0%	3.50	3,500
au	オーストラリア	2.0%	1.40	1,400
ca	カナダ	2.0%	1.40	1,400
de	ドイツ	2.0%	1.40	1,400
it	イタリア	2.0%	1.40	1,400
uk	イギリス	2.0%	1.40	1,400
br	ブラジル	1.0%	0.70	700
es	スペイン	1.0%	0.70	700
fi	フィンランド	1.0%	0.70	700
fr	フランス	1.0%	0.70	700
nl	スイス	1.0%	0.70	700
org	Organizations	1.0%	0.70	700
se	スウェーデン	1.0%	0.70	700
tw	台湾	1.0%	0.70	700
us	アメリカ	1.0%	0.70	700

さらに、表 2 に 2001 年末時点での jp ドメインの第二レベルドメイン毎の Web ページ数の推計を示す。これは、第二レベルドメイン毎に約 50 の Web サーバにランダムにアクセスし、各サーバが持つ Web ページ数を求め、JPNIC(<http://www.nic.ad.jp/>)が公開しているドメイン数との積をとることで求めた。なお、表 1 では jp ドメイン内の Web ページ

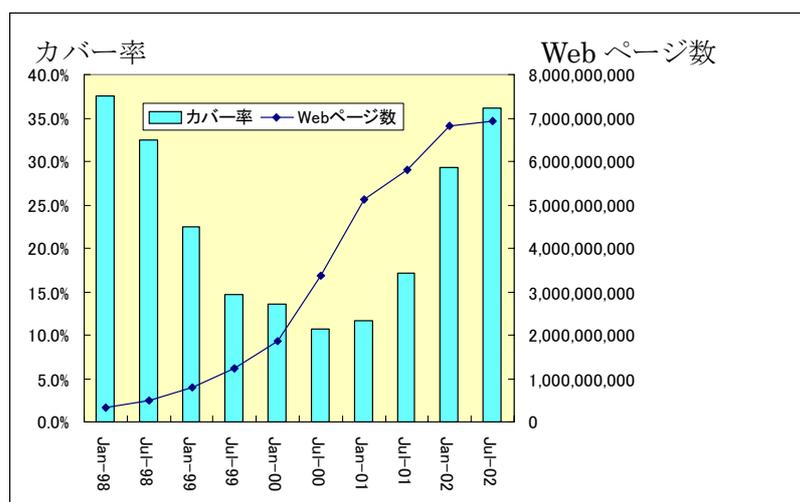
数は約 3.5 億と算出されている。これに対し、西村、山名らが 2001 年に行った jp ドメイン内の Web ページ数は推計[3]<sup>6</sup>でも、表 2 に示すように、約 3.1 億ページであり、表 1 の推計がほぼ実態に即していると判断できる。

表 2 日本の Web ページ数(2001 年 12 月現在)[3]

ドメイン名	第二レベルドメイン数	1ドメインあたりの平均ページ数	結果
AD	299	900	269,100
AC	2,679	4,491	12,031,389
CO	214,152	1,200	256,982,400
GO	669	4,111	2,750,259
OR	13,706	356	4,879,33
NE	16,931	1,533	26,955,223
GR	10,416	10,416	3,853,920
ED	3,139	3,139	822,418
地域	3,854	1,511	5,823,394
汎用	1	3,348	3,348
合計			313,370,787

### 3.4.2.3 推定 Web ページ数に対する商用検索システムのカバー率の推移

推定 Web ページ数を元に、1998 年 1 月以降の商用検索システムが全 Web ページ数 (推定) に対してカバーする率を求め、図 5 に示した。2000 年からカバー率が上昇に転じているのは、Google が検索対象 Web ページ数を増加させてきたことに起因する。

図 5 推定 Web ページ数に対するサーチエンジンのデータカバー率の推移<sup>7</sup>

<sup>6</sup> 当時は汎用ドメインが 1 つしか存在しなかった。現在は 20 万を超える汎用ドメイン (ドメインリストは入手不可) があり、これらの汎用ドメインをどのように扱うかを考える必要がある。

<sup>7</sup> 各 WWW サーチエンジンが持つデータ量の推移は、SearchengineWatch(<http://www.searchenginewatch.com/reports/sizes.html>)のデータを用いた。

### 3.4.3 効率的な収集とインデックス化に関する研究動向

本節では、膨大な Web ページの効率的な収集手法とインデックス化の高速化手法について紹介する。

#### 3.4.3.1 効率的な Web ページ収集手法

Web ページ収集には、収集ロボット（あるいはクローラー(Crawler)）と呼ばれるソフトウェアを用いる<sup>8</sup>。しかし、収集ロボットの収集能力に物理的な限界がある以上、何らかの方法で「効率的」な収集を実現する必要がある。

効率的な収集を実現する手法としては、全ての Web ページを集めることは元々不可能だという前提に立ち、①決められた時間内で**重要と判断される Web ページ**<sup>9</sup>から優先的に収集する方法と、②収集ロボット自体を分散化し高速に収集をする方法、さらに、③Web ページの更新差分のみを収集できるように WWW サーバ自体を専用化する方法、の3つに分類することができる。

##### 3.4.3.1.1 Web ページの重要度に基づいて効率的な情報収集を行う方法

Web ページの重要度として、①Web ページの更新頻度を用いる方法[4][5]、②Web ページのスコア<sup>10</sup>（**3.4.4** で詳述）を用いる方法[6][7]がある。

###### (1) Web ページの更新頻度を用いる手法

Web ページの更新頻度を用いてデータの鮮度を保つ手法に関する研究を示す。

① Stanford 大学の Cho らは、2000 年に「Web ページの実際の更新頻度を調べると共に、それを基にした収集方法」を提案した[5]。まず、PageRank<sup>11</sup>を元選ばれた重要なサイトを対象として、更新頻度の調査を行い、得られた結果からポアソン分布に基づいて更新頻度のモデル化を行う。更新頻度の分布を図 6 に示す。X 軸は更新の間隔、y 軸は該当する Web ページの全対象 Web ページに対する割合を示す。ポアソン分布は、ランダムに起こる事象の確率分布を求めるのに適しており、Web ページの更新頻度の確率分布を求める場合にもよく用いられる。Cho らは収集した Web ページの更新頻度の分布が、ポアソン分布に従うことを確認し、その前提に基づいて、収集された Web ページの

<sup>8</sup> よく用いられるものとして、GNU wget(<http://www.gnu.org/software/wget/wget.html>)がある。

<sup>9</sup> **3.4.3.1.1(1)**を参照。

<sup>10</sup> 検索結果を並べる際に用いるランキングで用いるスコア。

<sup>11</sup> **3.4**を参照。

新鮮さを最大に保つ収集方法を考案している。新鮮さとは、収集された Web ページが最新のものと同じである割合のことを指す。

通常の収集ロボットは全てのページを一定の周期で収集する。しかし、更新頻度が高いページは、索引付けがなされて検索システムで検索可能になったときには、既に最新の情報ではなくなっているという問題を持つ。しかし、Cho の提案する収集方法では、ページ毎に収集頻度を変えることにより、検索システム内のデータの新鮮さを常に一定に保つことができる。

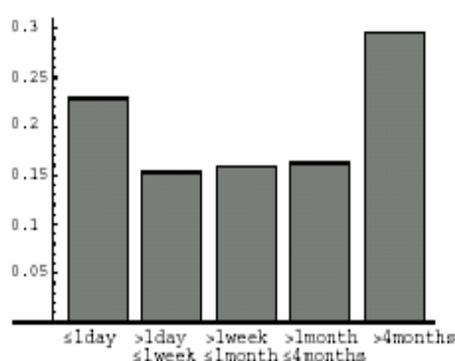


図6 更新頻度の分布 ([5]より引用)

また、新鮮度について、Cho らは2つの指標を定義している：

- freshness
  - ある時刻  $t$  において最新である Web ページ数の割合
- age
  - ある時刻  $t$  において最新でない Web ページについて、取得からどのくらいの時間がかかっているかを平均した値

これらの指標に基づき、freshness を上げ age を下げる手法、あるいは、freshness のみを上げる手法が考えられている。

- ② オーストラリア・シドニー大学の Edwards らは2001年、Cho の研究[5]を踏まえた上で、更に一般的なモデル化を提案[4]した。Cho の研究では、PageRank を基準として重要なページを選び、それらのページからモデル化を行っていたが、Edwards らの研究では収集時に同時に更新頻度の調査を行うことで、収集対象となる全てのページのモデル化を行う。最初は全てのページを1ヶ月周期で収集し、更新されていたページに対しては、設定されている更新間隔の半分、すなわち2週間周期で収集を行う。さらに、2週

間後に更新されていたページに対しては1週間周期で収集を行うにより、全てのページの更新頻度をモデル化する方法を提案している<sup>12</sup>。

また、収集ロボットの能力を考慮し、どのページを優先して収集を行えば、システム内のデータの全体的な鮮度が最も高くなるかを考察し、全体のデータの新鮮さをもっとも高く保つ為のスケジューリングのモデル化を行っている。図7に示すように収集対象のWebページを以前に収集した時 ( $t_0$ ) からの間隔で分類し、各Webページの優先度を決定している。一度収集したWebページ全体を  $b$  と定義し、ある時点 ( $t_0 + \Delta t$ ) において更新が発生したWebページを  $y$  とする。また、 $\Delta t$  の間に再収集されたWebページの数を  $x$  とする。また、 $a$  を  $\Delta t$  の間にWebページが更新される確率とする。そして、実際に各ページを収集したときに更新間隔を測定することにより、各ページの  $a$  を求め、収集されたWebページの新鮮さを最大に保つようにスケジューリングする手法を提案している。

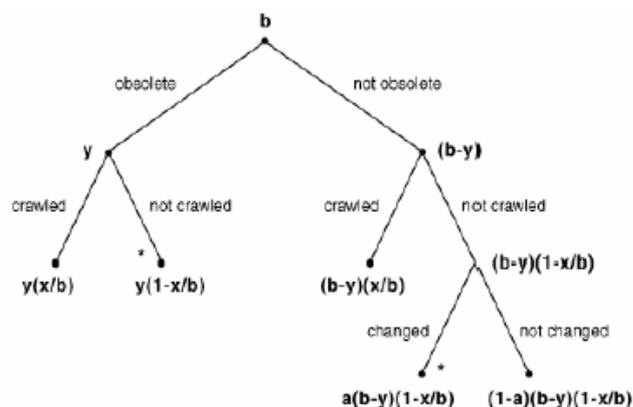


図7 収集対象となるWebページの分類 ([4]より引用)

## (2) Web ページのスコアを用いる手法

各Webページのスコア(重要度)を定義し、重要だと思われるページを優先的に収集することにより、効率的に収集を行う手法を紹介する。

- ① Stanford大学のChoらは、1998年に、(1)Webページへのリンクを見つけた順番に収集していく収集ロボット、(2)バックリンク(あるいはinlinkと呼ぶ)の多いWebページへのリンクを優先して収集を行う収集ロボット、(3)PageRankが高いWebページへのリンクを優先して収集するロボットの3つを用いて、どの方法が最も効率よく収集できるかを比較した[6]。実験は、Stanford大学内のWebページを用い、重要なページ

<sup>12</sup> 更新されていない場合は、逆に更新間隔を2倍に設定する。

を、「コンピュータという単語を含み、バックリンクが多いページ」と定義し、実験を行った。その結果、PageRankに基づいて収集を行う方法<sup>13</sup>は安定して高い成果を得られることがわかった。図9に示す通り、全体の80%のページを収集した時点で、収集対象のページ（重要ページ）のほぼ全てを収集し終えている。図中、x軸が全対象Webページに対する収集率、y軸が全需要Webページに対する収集率を示す。幅優先の収集方法は、PageRankを基にした方法には及ばないものの全体の80%を収集し終えた時点で、対象のページの約90%を収集しており、バックリンクを基にした収集方法では、対象のページの約80%しか収集されていなかった。

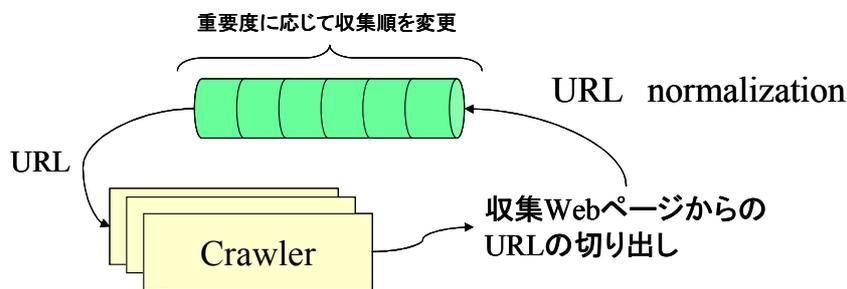


図8 収集順の変更方法

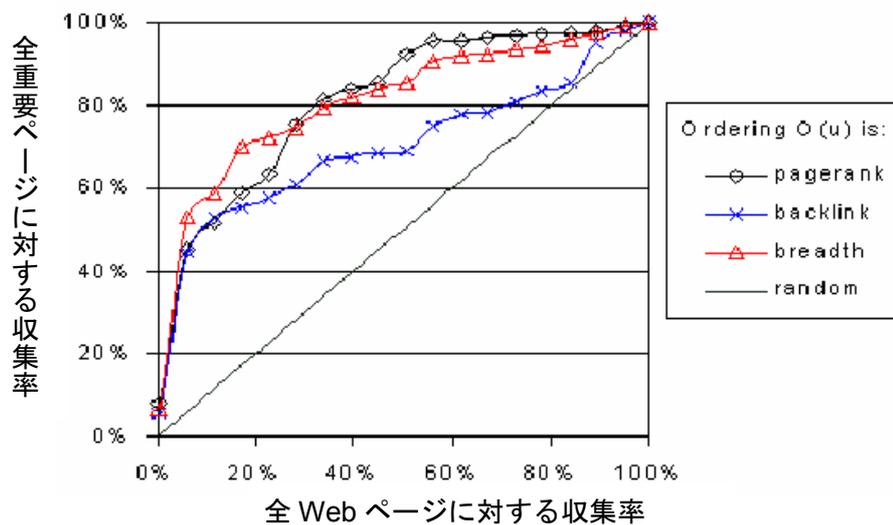


図9 各収集方法の比較

(computer という単語を含むページを対象とする) ([6]より引用)

<sup>13</sup> 図9に示すように、queueに入っている収集候補URLの内、どのURLを優先して収集するかを、その時点までで収集完了しているページからPageRankを用いて重要度（スコア）を算出し、高い重要度を持つURLに対してqueueの中での追い越しを許す。

- ② Compaq System Research Center の Najork らは、2001 年に、一般的に用いられている幅優先の収集方法で、十分に効率的な収集ができることを示した[9]。通常の収集ロボットでは、リンクが見つかった順に収集を行うが、幅優先の収集方法では、違うサーバに交互にアクセスするように収集を行う。幅優先に基づく収集方法により、同一 Web サーバへの連続的なアクセスを減らすことができる。また、リンクを順に辿っていく収集方法では、同じサイト内のページを連続して収集してしまうが、幅優先の収集方法では、違うサイトに同時にアクセスすることから、幅広いトピックを効率的に収集することが可能であるという利点も持つ。

Najork の研究では、重要なページの評価基準として PageRank を用い、1 日ごとに収集されたページを用いて PageRank を更新し、各日に収集した全ページの PageRank の平均を求めた。60 日間収集を行った結果、幅優先の収集方法で集めたページでは、1 日目に集めたページの PageRank が 2 日目以降に収集したページに比較し高いという結果が得られた (図 10)。また、2 日目以降に収集されたページの PageRank は平均して、1 日目の 4 分の 1 以下だった。この結果から、幅優先の収集方法では、収集の初期の段階で、重要だと考えられるページを多く収集できることを示す。

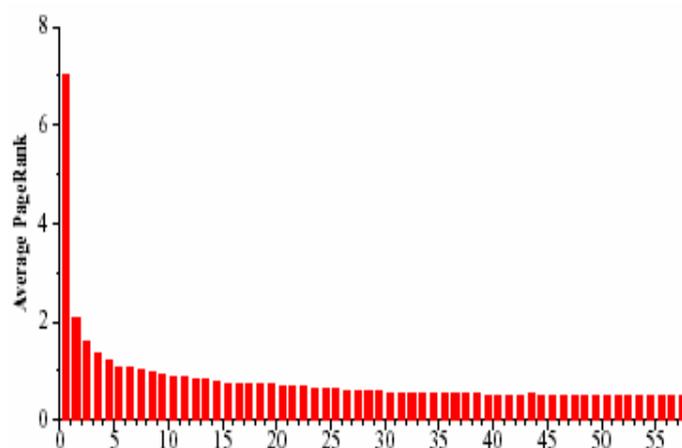


図 10 収集を開始した時からの経過時間(hour)ごとの PageRank の分布  
(x 軸は一日単位) ([9]より引用)

#### 3.4.3.1.2 収集ロボットを分散化させる手法

本項では、収集ロボットを分散させることにより収集速度を向上させる研究を紹介する。

- ① Transarc 社の Bowman らが 1994 年に提案した Harvest[10]では、収集と検索を共に分散して行うことにより、システムへの負荷を分散させる方法を提案している。

Harvest のシステム構成を図 11 に示す。Harvest では、Gatherer と呼ぶ収集ロボットをネットワーク上に分散配置し、収集を行う。しかし、Gatherer が収集する URL は事前に静的に決められており、スケジューリング方法については特に考慮されていない。ただし、分散して収集を行うという世界で初めての提案となった。

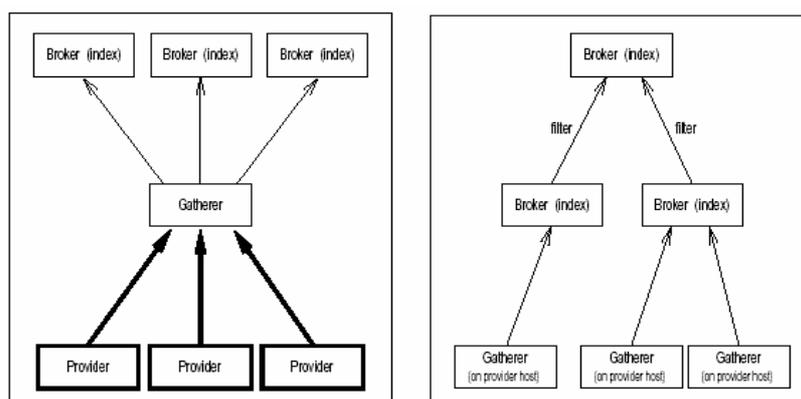


図 11 Harvest の構成 ([10]より引用)

- ② 東工大の内藤らが 2000 年に提案した超分散サーチエンジン[11]では、各 Web サーバに各 Web サーバが持つページをインデックス化する機能を持たせ、Web サーバ自体がクエリーを受け付けることのできる機能を実装している。つまり、Web サーバ自体が小規模の検索システムとして動作する。さらに、サーバのハードウェアスペックに余裕がある時には、自 Web サーバ内のページからリンクが張られているページを収集し、検索できるデータを増やす手法も提案している。超分散サーチエンジンを用いた情報検索は、リンクを辿ることによって実現される。論文では 6 つのモデルが提案されているが、いずれも、クエリーをある Web サーバに送り、検索結果として得られたページが存在するサーバに順次クエリーを伝えて探索する手法をとる。このため、検索のターンアラウンド時間は長くなる。
- ③ 早稲田大学の村岡らが 1998 年に提案した分散型ロボット[12]は、地理的に分散して設置した各ロボットの負荷を均等化することにより、Web ページの収集時間を短縮しようとする方法である。各収集ロボットにランダムに Web サーバを割り当てた場合に比較して、ネットワーク及び Web サーバの負荷を均衡化することによって、数倍高速に収集を行えることを示している。村岡らの方法では、同じ Web サーバに対するアクセスでも、収集ロボットのネットワーク環境の違いにより、負荷を均等化することにより高速化できることを示している。また、各収集ロボットの制御用に管理用サーバを

置き、動的に各収集ロボットが収集する URL を動的に決定している。管理用サーバには、各収集ロボットとサーバの距離や新規に発見されたサーバの情報が送られ、その情報を基に各収集ロボットの負荷を均等化する。収集ロボットとサーバの距離(データ転送速度)の計測は、サーバのトップから 50URL を収集し、収集にかかった時間から判断している。また、JP ドメインからランダムに抽出した 6,500 サーバを、国内 17 箇所に配置した収集ロボットで、負荷を均等化して収集を行った結果を示している。実験の結果は図 12 に示す通りであり、ランダムに収集を行った場合の予測値より、約 11 倍の速度で収集が行えたとしている。

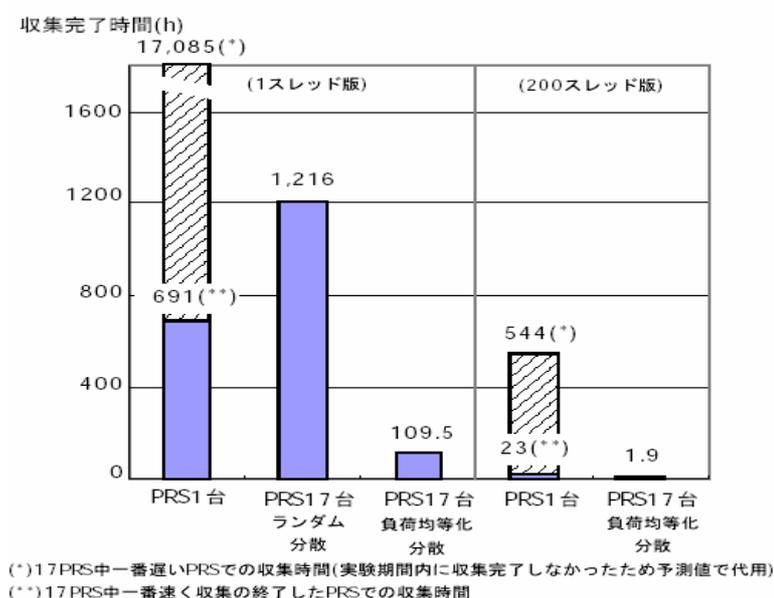


図 12 実験結果 ([12]より引用)

### 3.4.3.1.3 Web サーバを専用化する手法

本項では、物理的な収集時間を短縮させるために Web サーバを専用化する手法についての研究を紹介する。専用化する手法は大きく 2 つに分類される。一つは Web ページ更新時に変更された部分の元のページに対する差分を収集ロボットに収集させる(送る)ことができる機能を Web サーバに持たせる方法であり、もう一つは、いつ更新されたかを Web サーバ側が収集ロボットに対して通知するように Web サーバに機能を持たせる方法である。これらの手法は、一旦、対象となる Web ページを収集完了後に更新を効率的に行うために用いられる。

- ① Western Research Lab.の Mogul らが 1997 年に提案した Delta-Encoding[13][14][15][16]では、キャッシュを有効に行う為に HTTP プロトコルへの改良を提案している。Delta-Encoding では、Web サーバ内の各ページに対し、UNIX の diff コマンドを用いて求めた変更の履歴を保存する。クライアント(収集ロボット)からリクエストがあった時には、Etag からクライアントが保持しているページのバージョンを判断し、クライアントが保持しているページと最新のページとの差分を送信する。クライアントでは、送信された差分とキャッシュされているページから最新のページを作成する。従来の一般的なキャッシュでは、ページが更新されていると判断した場合には、そのページの全てを取得し直している。しかし、Douglass の調査[8]で示されているように、HTML ファイルは更新頻度が高く、全体の 50%程度は一ヶ月以内に更新されるため、従来のキャッシュでは、大きな効果を上げることはできない。これに対して Delta-encoding では、HTML ファイル内の、更新された部分だけをサーバから取得し直し、クライアント側で保存されているページと更新された部分から動的にページを生成することにより、更新頻度が高いページに対しても効率的に収集することを可能とした。
- ② Stanford 大学の Brandman らの 2000 年の研究[17]では、先に Cho が行った Web ページの更新頻度の調査[5]から、通常の収集頻度ではサーバに対して無意味なリクエストが行われてしまうことを示している。例えば、1ヶ月に 1回収集を行ったとすると、全体の 5割のページは更新されておらず、無駄な収集を行ってしまうことになる。その解決策として、サーバ側がサーバ内の各ページの更新情報をまとめ、収集ロボットは定期的にサーバの更新情報を取得し、更新があったページだけを収集する方法を提案した。Brandman の方法により、無駄なアクセスが無くなり、また、Web ページの更新間隔に見合った収集が可能になり、収集データの新鮮さを保つことが可能になる。
- ③ Illinois 大学の Gupta らは 2000 年、Brandman の提案した方法[17]をベースに、更新するのに最適な時期がきたら Web サーバ側から収集ロボットに通知するというシステムを提案している[18]。更新時期(Web サーバが収集ロボットに通知する時期)は、収集にかかるコスト (ページのサイズ) やページの重要度、ページが更新されてからの経過時間から総合的に判断する。
- ④ 筑波大学の加藤らが 1998 年に提案した Planet[19]は、分散環境の為にミドルウェアである。Planet を導入している Web サーバ群内では、対象の Web サーバにエージェントプログラムを送り込むことによって、当該 Web サーバと同一 LAN の Web サー

バ上のデータ収集を行う。収集したデータは、いくつかのファイルにまとめると共に、圧縮して送り返す。これにより、ネットワーク負荷を減らして収集を行うことが可能になる。

- ⑤ ブラジルの Minas Gerais 連邦大学の Silva らが 1999 年に提案した CoBWeb[20]は、既存のプロキシを用いた手法を提案している。検索システム内のデータの更新を、プロキシを用いて行う **Cache parasite** という方法を提案している。提案されているキャッシュ機能を持ったプロキシでは、キャッシュされている Web ページにユーザからリクエストがあると、リクエストされた Web サーバから当該 Web ページのヘッダ情報を取得し、更新されていると判別された場合には、新しくページを取得する。また、更新されていないと判別されたときには、キャッシュされているページをユーザに送信する機能を持つ。**Cache parasite** では、このようなキャッシュ機能を持ったプロキシと連動し、プロキシで更新を感知したページを検索システムに送信することで、収集を行うことなく検索システム内のデータの鮮度を保つ事を可能にしている。

### 3.4.3.2 インデックス化の高速化

収集した Web ページを検索可能にするためには、インデキシングを行う必要がある。このインデキシングには膨大な時間がかかることが知られており、例えば、2001 年 3 月時点で、Google はインデキシングに 1～2 ヶ月をかけていた<sup>14</sup>。このようなインデキシングを高速化する手法として、Stanford 大学の Melnik らは、2001 年にパイプライン化する手法を提案している。図 13 に示すように、インデキシングの行程を、「loading」「processing」「flushing」の 3 つに分割し、図 14 に示すようにこれらの行程が 3 並列で動作するようにスケジューリングを行う。これによって、図 15 に示すように 3 割程度の高速化が達成できることを示した。

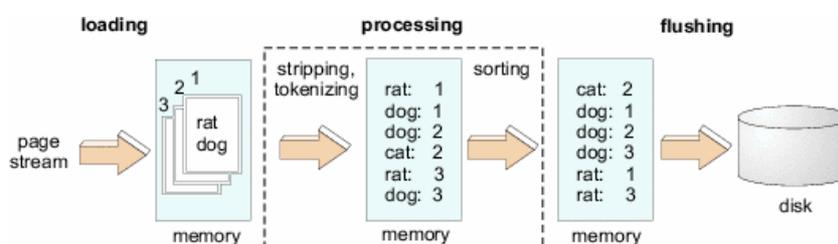


図 13 インデキシングのパイプライン化 1 ([21]より引用)

<sup>14</sup> 2001 年当時までは、Google は「収集→インデキシング」を約 3 ヶ月で実現しており、この処理を 3 つ並列に走らせることで 1 ヶ月毎にインデックスを更新(つまり約 3 ヶ月前のデータが検索可能となる)していた。以降、Google では、頻繁に更新されるページに対しては毎日インデックスを更新している。

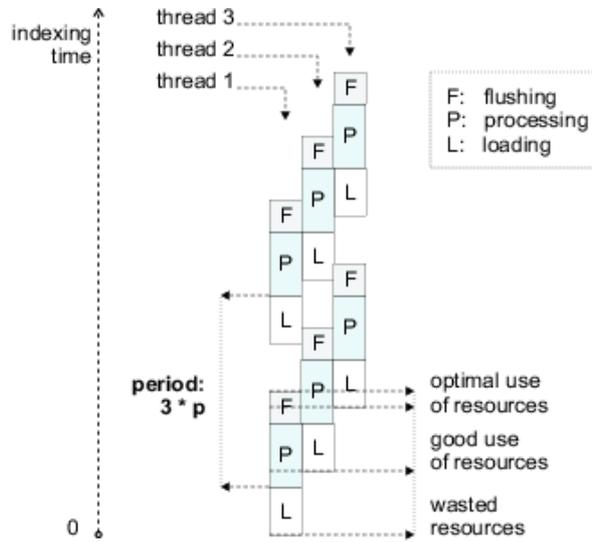


図 14 インデキシングのパイプライン化 2 ([21]より引用)

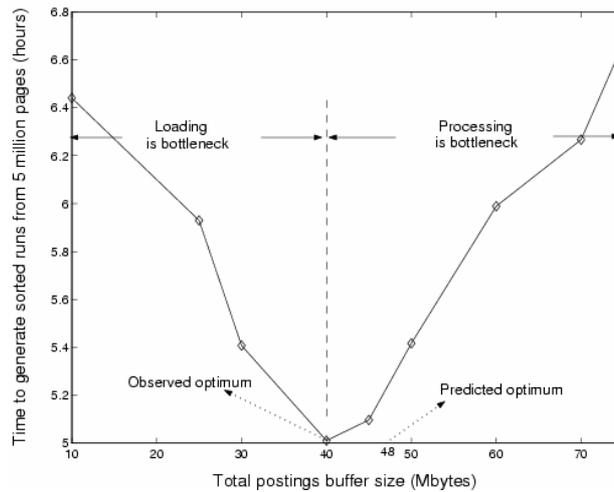


図 15 パイプラインの各行程の調整による処理時間の変化([21]より引用)

### 3.4.4 検索結果のランキングに関する研究動向

検索結果のランキング手法としては、従来の IR<sup>15</sup>研究で用いられているベクトル空間モデル等、各 Web ページの内容に基づいた手法に加えて、WWW サーチエンジンでは、リンク情報を用いてランキングを行っている。本節は、リンクを用いた主要なランキング手法について紹介する。

<sup>15</sup> Information Retrieval

## ① HITS[22][23]

HITS は、1997 年に IBM の Kleinberg らにより提案されたコミュニティ発見のためのアルゴリズムであり、その中で Web ページの重要度を表す Authority 度 (スコア) と Hub 度 (スコア) が定義された。これらの値は、Web ページの重要度を表す指標である。

HITS では、ハイパーリンクをリンク元からリンク先への推薦行為と考え、多くのページからリンクされているページを有用なページとして、Authority ページと定義し、また、多くの Authority ページにリンクしているページを良質な Hub ページと定義した。

### Authority スコア・Hub スコアの計算

まず、計算対象となる全てのページ  $p$  に対して、Authority スコア ( $Authority(p)$ ) と Hub スコア ( $Hub(p)$ ) を与え、まず、1 で初期化する。次に、以下の計算を  $Authority(p)$  及び  $Hub(p)$  の値が収束するまで繰り返す ( $p \rightarrow q$  はページ  $p$  がページ  $q$  に対してリンクしていることを示す)。概念図を図 16 に示す。図中の  $\circ$  がページを表し、 $a$  が Authority スコアを、 $h$  が Hub スコアを表す。

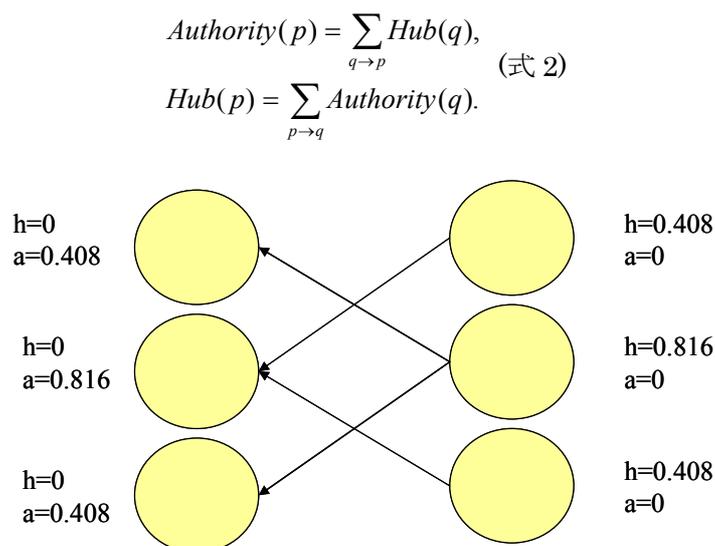


図 16 HITS の概念図

( $h$  が  $Hub(p)$  を、 $a$  が  $Authority(p)$  を表す)

## ② PageRank[24][25]

PageRank は、1998 年に Stanford 大学の Page と Brin が提案した Web ページの重要度を表す指標である。PageRank の基本的な考え方は、ネットサーフィンをする人のモデル化 (ランダムウォーク) に基づく。ネットサーフィンをする人は、Web 上の各ページのリンクをランダムに辿ると仮定し、各ページに辿りつく確率から各ページのランキングを

行う。ランキングに用いられるスコアの計算を(式 1)に示す。イメージとしては図 17 のようになる。

$$R(p) = \frac{\varepsilon}{n} + (1 - \varepsilon) \cdot \sum_{(p,q) \in G} \frac{R(q)}{\text{out degree}(q)} \quad (\text{式 1})$$

- $R(p)$  : ページ  $p$  の PageRank
- $R(q)$  : ページ  $p$  へのリンクを持つページ  $q$  の PageRank
- $n$  : 対象とするグラフ  $G$  のノード総数 (Web ページ数)
- $\text{outdegree}(q)$  : ページ  $q$  からの外向きリンク数
- $\varepsilon$  : dampening factor (0.1~0.2)

$R(p)$  をページ  $p$  の持つ PageRank、 $R(q)$  はページ  $p$  へのリンクを持つページ  $q$  の PageRank を示す。N は対象とするグラフ  $G$  (Web ページをノードとし、Web ページ間のリンクをエッジとしたグラフ) のノード総数 (Web ページ数)、 $\text{outdegree}(q)$  はページ  $q$  からの外向きリンク数である。また、グラフ  $G$  内の全ての Web ページの PageRank の合計は 1 であり、 $\varepsilon$  は通常 0.1~0.2 の間に設定される。

モデルとしては、グラフ  $G$  のノード  $p$  から  $(1 - \varepsilon)$  の確率で  $p$  上のリンクを辿り、ノード  $q$  に進むが、 $\varepsilon$  の確率でリンクとは関係なくノード  $r$  に進むランダムウォークのモデルである。つまり、ユーザは  $(1 - \varepsilon)$  の確率で現在の Web ページからのリンクを辿り、 $\varepsilon$  の確率で全く無関係な Web ページへジャンプするとの仮定のもとでネットサーフィンのモデル化を行っている。他にもリンク構造をランキングに応用している検索システムは多数あるが、PageRank の特徴としては ScamWeb に強いという点が挙げられる。Scam Web とは、あるページのランキングを上げるために、複数のダミーサイトからのリンクを張っているページのことである。PageRank では、このようなダミーサイトのスコアを小さくすることが可能である。

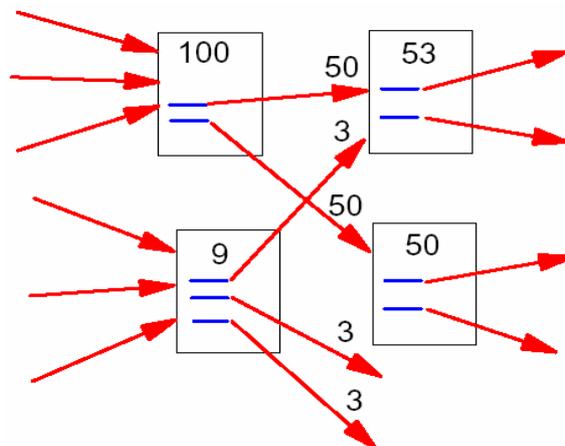


図 17 PageRank のイメージ図 ([24]より引用)

### 3.4.5 まとめ

WWW サーチエンジン構築上重要となる、①膨大な Web ページの効率的な収集とインデックス化、②検索結果のランキング、という 2 つの技術について紹介した。なお、紙面の都合上紹介できなかつたが、WWW

サーチエンジンのランキングを評価する指標として DCG(Discounted Cumulative Gain)[26]、MRR(Mean Reciprocal Rank)[27]等が提案されている。また、国立情報学研究所(NII)が主催した第 3 回 NTCIR Web Task<sup>16</sup>の評価として文献[28]でこれらの値が紹介されているので、利用者の立場としての評価指標として参考にされることをお勧めする。

#### [参考文献]

- [1] Steve Lawrence and Lee Giles: "Accessibility and Distribution of Information on the Web", Nature, vol. 400, pp. 107-109 (1999)
- [2] A. Broder, R.Kumar et.al. : "Graph structure in the Web : experiments and models", Proc. of the 9th WWW Conf. (2000)
- [3] 西村真幸, 山名早人 : "ドメイン毎の Web ページ数の偏りを考慮した日本の Web ページ数推定調査" , 第 64 回情処全大, 2X-6 (2002.3)
- [4] Jenny Edwards, Kevin McCurley, John Tomlin: "An Adaptive Model for Optimizing Performance of an Incremental Web Crawler", Proc. of the 10th WWW Conf., <http://www10.org/> (2001)
- [5] Junghoo Cho, Hector Garcia-Molina: "The Evolution of the Web and Implications for an Incremental Crawler", Proc. of the Conf. on Very Large Databases (2000.9)
- [6] Junghoo Cho, Hector Garcia-Molina, Lawrence Page : "Efficient crawling through URL ordering", Proc. of the 7th WWW Conf., pp. 161-172 (1998)
- [7] Marc Najork, Janet L.Wiener : "Breadth-first search crawling yields high-quality pages", Proc. of the 10th WWW Conf., <http://www10.org/> (2001)
- [8] Fred Dougilis, Anja Feldmann, Bachander Krishnamurthy : "Rate of Change and other Metrics: a Live Study of the World Wide Web", Proc. of the USENIX Symp. on Internet Technologies and Systems (1997.12)
- [9] Marc Najork, Janet L.Wiener: "Breadth-first search crawling yields high-quality pages", Proc. of the 10th WWW Conf., <http://www10.org/> (2001)
- [10] C.Mic Bowman, Peter B.Danzig, Darren R.Hardy : "Harvest: "A Scalable,

---

<sup>16</sup> <http://research.nii.ac.jp/~ntcadm/workshop/work-ja.html>

- Customizable Discovery and Access System”, Technical Report CU-CS732-94, University of Colorado-Boulder (1994)
- [11] 内藤清一郎,小林亜樹,山岡克式,酒井善則:“超分散サーチエンジンを用いた効率的情報探索について”, 信学技報, SSE99-200, pp.123-128 (2000.3)
- [12] 村岡洋一,山名早人、田村健人、河野浩之,森英雄,浅井勇夫,西村英樹,楠本博之,篠田洋一:“Internet 広域分散サーチロボットの研究開発”,第 20 回 IPA 技術発表会論文, <http://www.yama.info.waseda.ac.jp/~yamana/IPA/> (2001.10)
- [13] Jeffrey C. Mogul, Fred Douglass, Anja Feldmann, Balachander Krishnamurthy :“Potential Benefits of Delta Encoding and Data Compression for HTTP”, Proc. of SIGCOMM, pp. 181-194 (1997.9)
- [14] - :“What is HTTP Delta Encoding”, <http://www.webreference.com/internet/software/servers/http/deltaencoding/intro/>
- [15] James J.Hunt, Kiem-Phong Vo, Walter F.Tichy : “An Empirical Study of Delta Algorithms”, IEEE Software Configuration and Maintenance Wks., (1996)
- [16] Fink Dmitry, Sloutsky Alexander : “Delta Encoding Extension for HTTP/1.1”, <http://www-comnet.technion.ac.il/~cn33s00/>
- [17] O.,Brandman,Junghoo Cho,Hector Garcia-Molina,Narayanan Shivakumar: “Crawler-Friendly WebServers”, Proc. of the Workshop on Performance and Architecture of Web Servers (2000.6)
- [18] Vijay Gupta,Roy Campbell :“Internet Search Engine Freshness by Web Server Help”,UIUCDCS-R-2000-2153 (2000.1)
- [19] 加藤和彦:“モバイルオブジェクトシステム Planet: ネットワークコンピューティングのためのソフトウェア基盤”, テレコムフロンティア (1998)
- [20] Altigran S.da Silva, Eveline A. Veloso, Paulo B.Golgher: “CoBWeb-A Crawler for the Brazilian Web”, Proc. of the Sixth Symp. on String Processing and Information Retrieval (SPIRE'99), Cancun, Mexico, pp.184-191(1999.9)
- [21] S.Melnik, S.Raghavan et.al.:“Building a Distributed Full-Text Index for the Web”, Proc. of the 10th WWW Conf. (2001)
- [22] David Gibson, Jon M. Kleinberg, and Prabhakar Raghavan: "Inferring Web Communities from Link Topology", Proc. of the 9th ACM Conf. on Hypertext and Hypermedia (Hypertext'98), (1998)
- [23] J. M. Kleinberg: "Authoritative sources in a hyperlinked environment", In Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms,

pp. 668-677 (1998)

- [24] Lawrence Page, Sergey Brin, Rajeev Motwani, Terry Winograd, "The PageRank Citation Ranking: Bringing Order to the Web" , Stanford Digital Libraries Working Paper (1998)
- [25] S. Brin, L. Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine", <http://www-db.stanford.edu/~backrub/google.html>
- [26] K.Jarvelin and J.Kekalainen: " IR Evaluation Methods for Retrieving Highly Relevant Documents", Proc. of ACM SIGIR 2000, pp.41-48 (2000)
- [27] E.Voorhees: "The TREC-8 Question Answering Track Report", Proc. of TREC-8, NIST Special Publication 500-246, pp.77-82 (1999)
- [28] J.Eguchi, K.Oyama, E.Ishida, N.Kando and K.Kuriyama : " Evaluation Methods for Web Retrieval Tasks", Proc. of the Third NTCIR Workshop, pp.415-422 (2002.12)

## 3.5 デジタルライブラリとメタデータ - 概要と課題

杉本 重雄 委員

### 3.5.1 はじめに

今のデジタルライブラリにつながる活動としてどのようなものがあったかを、1990年代のはじめに戻り、筆者の知る範囲で振り返って考えてみたい。

- WWW はソフトウェアとしてはあったが、Mosaic が現れるより以前であり、今のような広がりはまだない。Gopher や WAIS が使われていた。
- エルゼビア社の tulip プロジェクトでは、アメリカの 9 大学と共同してページイメージで提供された雑誌論文のためのシステムの試作とその利用性の評価が行われた。
- アメリカ議会図書館の歴史資料のデジタル化プロジェクト American Memory が CD-ROM を配布媒体として使う前提で始まっていた。
- 1994 年 5 月に開催された Rutgers 大学での会議では California 大学、AT&T、Springer の協調による RedSage プロジェクトなどの報告がなされ、同年 6 月の Texas A&M での会議では多くのデジタルライブラリの研究プロジェクトについて発表がなされた。(Rutgers 大学の会議はその後 IEEE-CS、Texas A&M 大学の会議は ACM のデジタルライブラリの国際会議となり、さらにそれが統合されて現在に至っている。)

その後、アメリカでは 1994 年から NSF 他による共同助成で Digital Library Initiatives が始まり、イギリスの eLib などヨーロッパでも、インターネット上での新しい知識と情報の基盤としてのデジタルライブラリに関心が集まり、研究助成が進められた。一方、1995 年には、インターネット上での重要な開発課題として G7 で認められたトピックの中にデジタルライブラリが含まれ、国立図書館を中心とする Bibliotheca Universalis が進められた。

NCSA Mosaic の出現は、誰もがネットワークを利用している色々なコンテンツの利用ができることを直感的に理解させてくれた。これは、デジタルライブラリがどのようなかを実感できるという意味で、デジタルライブラリの研究開発を進める上での大きなサポートになった。90 年代中ごろのインターネットの爆発的な発展で、利用者の環境が整備され、ネットワークは情報ライフラインとでも言うべきくらい日常の仕事と生活の中

---

<sup>1</sup> デジタルライブラリに関するより詳しいレポートは平成 13 年度の HCIS-WG 報告書、あるいはそれを基に作った筑波大学附属図書館主催の電子図書館シンポジウムのチュートリアル資料[8]を参照していただきたい。デジタルライブラリやメタデータに関する参考資料のリストを筆者のページに置く[9]。本文中で紹介している組織や活動のページや関連論文等はそちらを参照していただきたい。

で利用されるようになった。こうした環境変化にともない、出版の電子化、特に学術雑誌の電子化が進んだ。現在では学術雑誌の場合、冊子体とネットワーク上でのオンライン出版を並行で行っているものが、多数を占めている。

また、インターネットとデジタルライブラリの発展とともに注目されるようになったものにメタデータがある。メタデータの定義は、単純に「データに関するデータ (data about data)」とされる。そのため、メタデータには、利用目的によって様々なものがある。インターネット上の情報資源 (Information Resource あるいは単に Resource) の発見のために開発された Dublin Core はインターネットやデジタルライブラリの分野ではもっとも広く認知されているメタデータ規則である。このほかにも、教育情報資源を指向したもの、電子商取引を指向したものなど、様々なメタデータ規則がある。インターネット上では、情報資源の検索、アクセス、課金、取引などネットワークを介した人間あるいはソフトウェアによるやり取りが行われる。こうした様々なやり取りを支えるためにメタデータが用いられている。メタデータの利用の広がりにはメタデータの相互利用性、相互運用性の必要性をも高めている。

デジタルライブラリ分野の主要な国際会議は、現在、ACM と IEEE Computer Society がスポンサーとなっている Joint Conference on Digital Libraries (JCDL) とヨーロッパの European Conference on Digital Libraries (ECDL) がある。JCDL は前述の ACM と IEEE がそれぞれ主催していた国際会議 (DL と ADL) が 2001 年から統合してできたものである。ECDL は 1997 年に ERCIM が主催して始められたものである。アジア地域をベースにするものとして、International Conference on Asian Digital Libraries (ICADL) がある。我が国では、1995 年、97 年、99 年に図書館情報大学で主宰した国際シンポジウム、2000 年に京都大学が英国図書館および NSF と共同で主催した国際会議がある。デジタルライブラリの研究分野は、データベースや WWW 関連の技術を中心とする要素情報技術を総合することが求められる応用分野であり、また社会科学的な側面をも含んでいる。

### 3.5.2 デジタルライブラリの研究開発の活動を振り返る

大きな流れとして、情報技術の研究開発助成プログラム、図書館によるデジタルライブラリサービスの開発、そして出版から流通に至る過程での新しいサービスの開発である。これらを概観してみたい

#### 3.5.2.1 新しい技術の研究開発の視点を中心とした研究助成プログラム

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デジタルライブラリのための要素技術の研究を中心に進められた。一方、DLI-2では少数の大規模プロジェクトに加えて、多数の中小規模のプロジェクトが進められた。また、DLI-2では、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。

ヨーロッパでもネットワーク上での学術情報環境に関連してデジタルライブラリの研究開発が進められた。イギリスで進められたeLibプログラムは助成母体となったJISCが高等教育機関における図書館や情報環境を担当する組織であるため、DLIに比べると、図書館よりの研究が多く進められた。ドイツではGlobalInfoがすすめられた。EUでは第5および第6フレームワークの中でInformation Society Technology (IST)プログラムが進められ、Digital Libraryに関するプロジェクトはその中に含まれている。たとえば、Digital Libraryに関する国際会議などの活動を進めているDELOS (DELOS Network of Excellence on Digital Libraries)はISTプログラムの助成を受けている。

#### 3.5.2.2 図書館をベースにした活動

図書館におけるデジタルライブラリの開発は下記のように大きく分類できる。

##### (1) 電子出版物の提供

電子ジャーナルやデータベースなど、電子的に出版された資料を収集し、利用者に提供する。

##### (2) 既存資料の電子化と提供

図書館自身が既存資料（たとえば、冊子体資料）を電子化し、蓄積提供する。電子化の方式は、資料の種類と利用目的の違いに応じて、ページイメージ、SGML (XML) を用いた電子テキスト、OCR で取り込んだテキストとページイメージの組み合わせなどさまざまである。

原資料の保存性を高めるとともに、かつネットワークを介したデジタルコピーへのアクセス性を高めるという観点から貴重資料や歴史資料のデジタル化が進められる。また、学位論文やテクニカルレポート等の大学などの組織が発信する資料の電子化が進められている。貴重資料の場合は「保存とアクセスの両立」、学位論文等の場合は組織の情報発信とアクセス支援が基本的な目的であると言える。

### (3) ネットワーク上の資料のアクセス支援サービス

有用な資源を収集し、資源に関する情報を適切に組織化して提供する。たとえば、サブジェクトゲートウェイと呼ばれるサービスはその典型である。

WWW やメールを利用したネットワーク上でのレファレンスサービスも行われている。また、レファレンスサービスの記録の再利用を行うことも進められている。

### (4) 図書館環境のハイブリッド化

従来型の資料、従来型のサービスとネットワーク型の資料とサービスを適切に組み合わせ、総合的な利用環境を提供する。

上の項目に加えて、デジタル資料の保存が図書館での大きな関心を集めている。従来型の資料と同様に、ネットワーク上で提供されるデジタル資料を収集し、保存する。デジタル資料の保存は、後でも述べるように、いろいろな困難な問題を多く含んでいる。

各国の国立図書館では、歴史資料や貴重資料の電子化、ネットワーク上で提供される有用な資源へのアクセス支援、納本制度に基づくネットワーク資料の保存といったことが中心的な話題である。大学図書館では、電子ジャーナルなど電子出版物の提供、貴重資料や歴史資料の電子化、大学が発信する資料の電子的提供とアクセス支援といったことが中心的话题になっている。公共図書館でも地域資料の電子化やネットワークの利用支援などが行われる他、ビジネス支援の一環として電子資料の提供への取り組みが見られる。

## 3.5.2.3 新しいサービスの開発

### (1) 電子ジャーナルについて

90年代の後半から、学術雑誌の電子的提供が進み、現在では非常に高い比率で電子的に提供されている。大学図書館での電子ジャーナルの提供はすでに一般化している。一方、価格の高騰や出版社による雑誌タイトルの囲い込みが進むとともに、図書館側でもそれに相対する動きがあった。ひとつは Stanford 大学を核とした HighWire press で、大学が学術雑誌の出版社と組んで電子ジャーナルの出版を進めるものである。また、SPARC の取り組みは図書館を中心とするコンソーシアムを形成し、電子ジャーナルを中心とする学術電子出版物に関する情報の共有や意見交換を進めている。

### (2) 第三者による雑誌の電子化について

JSTOR は雑誌の創刊号から全ページをイメージデータ化し、提供する遡及的なサービスを行っている。国立情報学研究所の電子図書館サービス NACSIS-ELS は国内の学協会が発行する学術雑誌を提供しており、遡及的な入力も行っている。こうしたサービスは、雑誌毎に完全なセットの内容を提供することができ、ネットワークを介して簡単にアクセ

のできるアーカイブの役割も持っている。また、JST の JStage のように、雑誌の編集段階から提供までの全過程を支援する取り組みもある。

#### (3) e-Print リポジトリについて

学位論文やテクニカルレポート、プレプリントなど、商用の流通ルートには乗らない資料を電子的に蓄積し、提供するサービスが実現されている。査読プロセスなど、雑誌論文との違いはあるが、学術論文が無料でアクセスできるという利点を持っている。しかも、査読や出版プロセスによる遅れが無いので、より早く研究成果にアクセスできる。研究成果の速報性が求められる分野では、雑誌論文は内容の保証された記録としての意味しか持たないという可能性もある。

#### (4) 電子ジャーナルや e-Print リポジトリの協調

電子ジャーナルや e-Print リポジトリを結ぶことによって、従来の出版形態とは異なるサービスが可能になる。例として、Virtual Journal、CrossRef、Open Archives Initiative(OAI)といったものがある。

Virtual Journal は American Institute of Physics と American Physical Society が出版する雑誌をソースジャーナルとして、特定分野の記事を集めて作る雑誌である。CrossRef は文献への参照により論文へのアクセスを可能にする。これらは論文や記事単位でのアクセス可能性を利用したサービスである。また、OAI は e-Print リポジトリが協調してメタデータを収集 (Harvesting) できるようにすることによって、リポジトリ単位のサービスからリポジトリにまたがるサービスを可能にしている。

#### (5) サブジェクトゲートウェイ

社会科学や、工学、医学といった何らかの主題分においてネット上の有用な資源を収集し、それらを適切に組織化し、分類して検索機能とともに提供するサービスである。ネットワーク上の資料を提供する仮想的な図書館と考えることもできる。資源の内容の評価など、人手に負うところが多い。

#### 3.5.2.4 デジタルライブラリにおける課題

デジタルライブラリの実現には様々な情報技術を総合することが求められる。その視点からは、情報技術の研究領域全体がデジタルライブラリに関連するということができる。一方、デジタルライブラリにおけるコアな話題は何であるかを考えてみたい。JCDL や ECDL といった国際会議の話題を見ると、情報技術の視点からのデジタルライブラリの研究領域が見えてくる。情報検索、情報資源の可視化、情報資源からの情報抽出といった技術要素を多様な情報資源に対して適用すること、デジタルライブラリとしての新しいサービスモデルとそれを支える技術などがある。また、実際に「使えるのか」という観

点からの評価、社会科学の視点からの研究も見られる。

NSF は DELOS (EU) と共同して、97 年から 98 年にかけてデジタルライブラリの研究戦略を議論するワーキンググループを設けた[1]。これは下記の 5 つのワーキンググループからなりたっている。

- (1) Intellectual Property and Economics
- (2) Global Resource Discovery
- (3) Interoperability
- (4) Metadata
- (5) Multilingual Information Access

この領域がデジタルライブラリを情報技術の観点から見た際に重要な領域であると思われることができる。

Digital Library Federation (DLF) は Strategic and Business plan のレポート[2]の中で、現時点での活動項目について下記の項目を挙げている。

- (1) Digital Library Architecture, Technologies, Systems, and Tools
- (2) Digital Collections
- (3) Use, Users, User Support, and User Services
- (4) Digital Preservation
- (5) Standards and Practices
- (6) Institutional Roles and Responsibilities of the 21st Century Digital Library

最後の項目は DLF 自身の役割について述べている。図書館における実際のデジタルライブラリサービスを構築して維持していくという観点からの中心的な話題として理解することができる。

筆者は昨年度の HCIS-WG 報告[3]の中でデジタルライブラリに関する課題として、Interoperability と Preservation を中心に述べた。Interoperability は、デジタルライブラリに限らずネットワーク上でのサービスにとって重要な課題である。また、図書館サービスとして特徴的な機能であると言えるデジタルコンテンツの保存 (preservation) の問題は、技術的にも、制度的にも様々な問題を含んでいる。

DLI を進めた部門のディレクタであった Y.T. Chien 氏は、ネットワーク上のビジネスモデルとしてのミドルウェアが重要であると述べている[4]。WWW における Web Services の取り組みは、Web 上のサービスのミドルウェア化と捉えることもできる。デジタルライブラリが本当に利用できるようになるには、デジタルライブラリのサービスが個別の閉じた応用システムとして利用できるのみならず、ミドルウェア的に利用できるようになることが求められる。

Digital Library とは何かということは最近ほとんど議論されなくなったように思う。その一方、多様なデジタルライブラリサービスが提供されるようになり、我々自身も日常的にデジタルライブラリサービスを利用するようになってきた。ヴァージニア工科大学の Edward Fox 教授は、ストリームモデル (stream model)、構造モデル (structural model)、空間モデル (spatial model)、シナリオモデル (scenarios model)、社会モデル (societies model) の 5 つの階層からなる 5S モデルを提案している[5]。デジタルライブラリのための情報技術の研究開発と実際のサービスを結び付けていく上で、こうしたモデル化の取り組みが求められる。

#### 3.5.3 メタデータ

##### 3.5.3.1 メタデータの概観

最近、メタデータに関する話題をよく聞く。メタデータの定義は、「データに関するデータ(Data about Data)」といたってシンプルであるため、メタデータには様々な目的のものを含むことになる。メタデータは、ネット上で資源を探し、アクセスし、利用し、料金を支払う、あるいは資源を維持、管理、保存するといった様々な目的で利用される。図書館の目録や索引は典型である。衣料品についての素材の情報、出版物の知的財産権表示、テレビ番組表やビデオの内容記述などもメタデータである。ネットワークの QoS 制御をメタデータによって行おうという取り組みもある[6]。このように、メタデータには様々なものがある。下に、いくつかのメタデータの例を挙げよう。インターネット上での Resource Discovery (Dublin Core)、教育学習資源の記述 (IEEE Learning Object Metadata)、出版物の電子商取引におけるコンテンツやエージェントなどの記述 (indecs)、デジタルコンテンツのためのユニークな識別子 (DOI: Digital Object Identifier)、マルチメディア (オーディオ・ビデオ) コンテンツの記述 (MPEG-7)、政府行政情報 (出版物) の記述 (GILS) などのほか、地理情報や環境情報など多様な分野でメタデータ規則が開発されている。また、デジタルコンテンツの長期保存に関して、メタデータの観点からも検討が進められ、OAIS (Open Archival Information System) のレファレンスモデルが国際標準化されている。Lagoze 等が提案した ABC メタデータモデルは、デジタルコンテンツの状態の変化 (イベント) に注目したものである[7]。

図書館分野では、目録や索引などのメタデータを維持管理してきた。MARC (MACHINE-Readable Cataloging) を代表とする書誌データの記述形式や Subject Heading、10 進分類など統制語彙の開発も行ってきた。デジタル資料への対応とネットワーク上でのサービスの提供のために新しいメタデータ規則の開発が進められてきた。EAD (Electronic Archival Description) はデジタルコンテンツの記述と検索利用を目的とし

て SGML を用いて定義されている。Making of America II 等のプロジェクトの経験を基礎にして、METS (Metadata Exchange and Transmission Standard) はデジタルコンテンツの記述、管理、および長期保存のためのメタデータとして開発された。MODS (Metadata Object Description Schema) は書誌データを表すためのエレメントセットと XML 上での表現を決めている。MODS は、非常にたくさんの要素をもつ MARC (MARC21) のサブセットとして定義されるとともに、XML 形式としてデータ交換のために利用しやすくなっている。

World Wide Web コンソーシアムが進めるセマンティックウェブ (Semantic Web) の活動からは、現時点では、Resource Description Framework (RDF) や Web Ontology Language (OWL) といった話題をよく目にする。こうした技術はネットワーク上でのメタデータの流通性、共有性を高める上で、とても重要な役割を持っている。

### 3.5.3.2 メタデータに関する課題

メタデータに関しては Interoperability が基本的でかつ重要な課題であると考えられる。XML 形式によるメタデータの提供、RDF の利用が進むことによって、形式的な表現上の Interoperability はかなり達成されると思われる。RDF Schema のようなメタデータ規則に含まれる記述要素や値 (もしくはそのクラス) の名前とその意味の定義 (メタデータエレメントの語彙定義) を行う技術を利用することで、メタデータエレメントの意味定義をネットワーク上で共有することも可能である。(すくなくとも、あるメタデータ規則のあるエレメントを一意に識別することは可能である。) このように XML を媒介とした Interoperability の基盤はできつつあると思われる。一方、応用毎に必要とされるメタデータ規則は、標準的なメタデータ規則をそのまま良い場合もあるが、いくつかの規則を組み合わせたり、応用毎の拡張が行われることが多くある。異なるメタデータ規則に基づく応用間の Interoperability を高めるには、メタデータ規則の中で用いられる記述要素や値の関係性を定義することが求められる。Interoperability を得るためには、はじめから統一されたメタデータ規則を用いることができればよいが、それは現実的ではない。そのため、異なるメタデータ規則の下で開発されてきたシステム同士の Interoperability を得る方法が求められる。その場合、応用毎のメタデータ規則の持つ表現能力に比べてレベルを落としても Interoperability を得るようなことが求められることになると思われる。Dublin Core における Dumb-Down の概念、OAI における Simple Dublin Core を用いたメタデータハーベスティングのプロトコルはその例である。

Interoperability を高める上でメタデータ規則のネットワーク上での共有が求められる。形式的なメタデータ規則の記述による人間と計算機の両者による共有が求められる。この

視点から、メタデータスキーマレジストリの開発が求められる。データ項目を定義し登録するためのメタデータレジストリの国際規格 (ISO 11179) も進められている。メタデータスキーマレジストリにはメタデータ規則の流通に加えて、メタデータエレメント間の関係の定義を与えることなど、メタデータの **Interoperability** を支えることが期待される。また、メタデータを利用した応用システムの開発の支援やメタデータ規則を応用する上での参考となる実践例などを提供することも有効であると考えられる。

DCMI ではレジストリワーキンググループをつくり、プロトタイプの開発を進めてきている。DELOS はメタデータレジストリに関するワーキンググループを作り要求要件等に関する議論を進めてきている。IST プログラムで助成された Schemas Forum プロジェクトはメタデータ規則の収集と意見交換を進めた。それに続く、CORES プロジェクトでは Dublin Core、MARC、ONIX、IEEE LOM などのメタデータ規則の開発を進める幾つかの組織から代表を招いたメタデータスキーマレジストリに関する意見交換と合意のためのミーティングを開いている。また、2003 年 1 月には Open Forum for Metadata Registries という会議が開かれている。ネットワーク上でのメタデータの相互利用性を高める上で、異なるメタデータコミュニティの間でのメタデータレジストリの相互利用性を高めることが求められる。

また、別の視点からの課題として、メタデータに関わるコストの問題がある。情報資源の分類や抄録作成の自動化、自然言語や画像、音声理解技術を利用したマルチメディアデータの索引付けの自動化などは情報技術分野における以前からの課題である。Semantic Web や DAML などの活動は、オントロジー (語彙) のレベルでの利用可能な資源が豊富になると思われる。そうした活動の成果によってこうした分野の研究がさらに進むと期待したい。

#### 3.5.4 デジタルライブラリを取り巻く環境—情報環境の変化とそこで求められる技術

この 10 年ほどの間に我々の情報環境は大きく変化した。デジタルライブラリとそれを取り巻く環境の変化の視点からデジタルライブラリに求められる技術について述べてみたい。

##### (1) 学術論文や記事の出版方法と流通経路

学術雑誌の電子ジャーナル化が進み、非常に多くの雑誌が電子的に提供されている。国内の大学図書館でもここ数年の間に電子的に提供される学術雑誌のタイトル数が飛躍的に増えた。一方、プレプリントやテクニカルレポートなど、商用の流通経路には乗らず、以前は入手に手間取っていた文献 (e-Print) が、インターネット上で多く提供されるようになった。学術雑誌は査読され、編集された品質の高い内容を提供する一方、出版までに時

間がかかること、雑誌の価格が高いために必ずしも多くの人の目に触れないという現実的な問題を抱えている。e-Print は一般に無料でアクセスすることができ、かつ出版までの時間がかからない一方、内容の評価は読者に任されることになる。価格の問題は図書館にとっての大きな問題であるばかりではなく、図書館を介して論文にアクセスする研究者、学生にとってとても大きな問題である。評価の高い雑誌に論文が掲載されること、数多く参照されることは研究者としての評価につながっている。仮に論文として出版される成果の多くが、論文になる前にプレプリントやテクニカルレポートとして出版されているとすると、e-Print の段階での評価を行うことができれば、これまでの学術雑誌の役割が変わる可能性がある。

また、学術雑誌の電子ジャーナル化によって、雑誌の巻や号といった単位が必ずしも意味を持たなくなり、また文献間の参照関係がハイパーリンク化できることになった。複数の雑誌の中から記事単位で取り出して作る仮想ジャーナル (virtual journal) といった取り組みも現れている。また、CrossRef のようにオンライン出版された文献間の参照関係を結ぶサービスもある。また、有料の雑誌の場合であっても、書誌情報 (タイトル、著者、キーワード、アブストラクトなど) と参考文献が無料で公開されているものも多くある。e-Print の場合は書誌情報のみならずフルテキストが公開される場合が多い。このように、論文・記事単位での情報が豊富にネットワーク上に提供されることになると、それらを利用した付加価値サービスが可能になる。Open Archives Initiative はそうした取り組みの代表である。OAI では e-Print を提供するリポジトリからメタデータを収集するためのプロトコル (Protocol for Metadata Harvesting, OAI-PMH) を決め、メタデータの協調的な利用を可能にしている。メタデータを収集 (Harvest) して、それに何らかの価値を付加して提供することが可能であり、これまでのリポジトリ単位でのサービスに比較して、広い範囲の資源へのアクセスが可能になることに加えて新しい付加価値サービスの可能性を持っている。たとえば、前に述べた論文や記事の評価のための基盤となる巨大な情報源となり得ると考えられる。

論文の評価のために Citation Index が用いられてきている。CiteSeer のようにネット上で流通する論文に関して自動的に Citation Indexing を作ろうとするものもある。電子ジャーナルや e-Print リポジトリに加えて著者によってネット上に直接発信される論文や記事をも含めた巨大な情報源の中から、価値があると評価されるものを組織化して、利用者に提供することは学術情報を指向するデジタルライブラリにとって重要な要素であると思われる。

## (2) デジタル資料の収集と保存

電子資料の保存はデジタルライブラリにとっての重要な課題である。現在、非常に多

数の資料が電子的に作られ、電子的にのみ出版されるものも多くある。現時点では、印刷物が主で電子的なものが補助的な役割を持っているものであっても、将来はその役割が逆転、ないしは電子的なものだけになる可能性もある。商用の出版ルートに乗らない学術資料や政府行政資料などはこうした傾向が強いと思われる。そのため、電子資料、特にデジタル資料の保存は将来に向けた大きな課題である。

資料の保存は、資料を利用可能な状態に保つということに加えて、保存すべき資料を収集するという面でも困難な問題を持っている。資料の保存は、「もの」として資料を保存するというだけでは不十分であり、資料の「内容」を人間が利用できる状態に保つことが求められる。したがって、CD-ROM や磁気テープを「もの」としておいておくだけでは保存とはいえない。また、インターネット上でのみ発信される資料の場合は、「もの」はなく、「内容」を保存することが求められる。

非電子資料の場合、基本的に資料が傷まないように保つことで保存することができた。ところが、電子資料の場合、利用するには資料から内容を人間が視聴できるようにするための機材が必要とされる。すなわち、資料の保存は技術の保存でもある。この点はアナログ媒体でも、デジタル媒体でも同様である。たとえば、レコードやオープンリールのテープレコーダ、8 インチのフロッピーディスクなど、以前は広く使われたのに現在はほとんど見ないものや製品が非常に少なくなったものが多い。音や画像などの単純なデータの場合、内容を新しい媒体にコピーしなおすことで保存に適した形式にすることができる。これは紙の資料をマイクロフィルム化したり、デジタル化したりするのと同じである。デジタル媒体を用いた資料の場合、内容の利用にソフトウェアを必要とするもの、あるいは内容がもともとソフトウェアとして構成されているものが多い。CD-ROM を使った辞書や教材などはそうしたものの例である。Web 上のコンテンツの場合も plug-in ソフトウェアを必要とするものが多い。こうした動的な資料を保存しようとする、ソフトウェアが動く環境、ひいては環境を実現する技術を保存することが求められる。コンピュータシステムの場合は新しいシステムに置き換える際に後方互換に関する留意がなされ、場合によってはエミュレータが用いられることもある。デジタルコンテンツの保存に関してもエミュレーションを行う方法と、新しい環境への移植（マイグレーション）を行う方法が提案されている。しかしながら、ソフトウェアの多様性とコストを考えるとデジタルコンテンツの完全な（あるいはできるだけ完全な）保存は実現の困難な問題であることは容易に想像できる。

図書館は資料を収集し、組織化して、保存するという役割を持つ。従来、出版物の流通ルートに乗った出版物、大学や自治体など図書館が所属するコミュニティが作り出す、あるいは持つ資料を、図書館の主たる資料収集対象としてきた。ところが、インターネット

上での情報発信が進んだことで、資料へのアクセスが容易になった代わりに、どのような資料を収集対象とすればよいかを決めることが大きな問題となってきた。納本制度に基づいて納本図書館の役割を持つ図書館（日本では国立国会図書館）の場合、将来の世代に現代の文化を残すという役割を果たすにはインターネット上に発信される資料を収集対象から除外することはできない。国立国会図書館の場合、CD-ROMなどの媒体に固定して流通される出版物（パッケージ系出版物）の場合は納本制度にすでに収集を始めている。しかしながら、ネットワーク上で出版される出版物（ネットワーク系出版物）については納本制度審議会で現在検討が進められている。ネットワーク上の出版物を納本制度の下に（強制的に）収集することに関する議論、あるいは知的財産権などの制度に関する議論はここでは扱わない。

図書館でのネットワーク系出版物の収集は、網羅的収集と選択的収集に大別できる。網羅的収集については、Internet Archiveが行っているように、ロボット（あるいはCrawler）による自動収集方法が適用できる。ところが図書館としてふさわしくない内容、あるいは公開することが非合法であるような内容の資料まで収集し、公開する（あるいはアクセス経路を提供する）恐れがある。一方、選択的収集の場合、内容や品質の評価に基づいて収集するため内容に関する問題は生じないが、コストが高くかつ大量の資料を集めることが困難である。内容を適切に評価し、分類する機能が必要とされている。また、電子的な文書は修正が容易に行えるのでどの時点で文書が固定されたと判断するかという問題、動的に生成される文書のように収集そのものに困難がともなうといった問題もある。

冊子体の出版物の場合、「もの」として配布されるためにどこかの図書館には残っているという期待を持つことができた。ところが、電子ジャーナルのように出版社のサーバに蓄積され、図書館が利用のライセンスしか持たないことになると資料の保存に不安を感じるようになる。第三者（公的な機関）による電子化を行っている場合、あるいは電子的保存を認めている場合は、そうした機関が保存することが期待できる。一方、スタンフォード大学で行われているLOCKSSのように、出版者の側をあまり頼らず利用者（図書館）の側で電子的なコピーを残す試みもある。

### **(3) 出版から利用にいたる環境の変化と図書館のビジネスモデルの変化**

従来、図書館は出版者と利用者の上に位置付けられ、出版物を収集、蓄積して利用者に提供するという役割を果たしてきた。これまでは蓄積と利用の場所を提供することが図書館の付加価値として認められたのに対し、ネットワーク上では出版者と利用者が直接結びつくことが可能であるため、デジタルライブラリの場合、従来とは異なる視点からの付加価値を提供することが求められる。大量の資料を適切に組織化して提供するという、図書館の基本的な役割は変わらないと思われる。しかしながら、次に述べるように、ネット

ワーク上の情報流通環境は従来のものとは異なっており、デジタルライブラリのための新しいビジネスモデルが求められる。

学術分野では著者と読者のコミュニティはほぼ一致する。これまで学術論文の出版における投稿から利用に至る過程で、出版社、流通業者、図書館という中間に位置付けられた組織がそれぞれの間で上流（著者）から下流（読者）に至る出版物の流れを作ってきた。ところが、ネットワーク上では「もの」としての流れは意味を持たなくなる。現在の図書館が提供している電子ジャーナルのサービスは出版社のサーバへのアクセスを提供しているものが大半である。この範囲では、「電子図書館」は図書館ではなく出版社が提供していることになる。

先に述べた e-Print リポジトリの場合、リポジトリは著者と読者を直接つなぐ「電子図書館」の役割を果たしている。ヴァージニア工科大学を中心にいくつもの大学が参加して進めている学位論文の e-Print リポジトリである **Networked Digital Library of Thesis and Dissertations** の場合、学位論文を電子的に蓄積するためのツールの提供や著作権に関する規約を決め、学位論文の「電子図書館」としてサービスを提供している。学位論文の場合、著者によって完成されたものがそのまま公開されるという点で、出版のための編集作業を経て公開される雑誌論文とは少し異なった面をもつにせよ、学術出版物の中継ぎ役としてのひとつのビジネスモデルであると考えられる。また、OAI や CrossRef なども新しいビジネスモデルとして興味深い。

#### (4) コンテンツの変化とアクセス方法の変化

ネットワーク上で提供される資料と従来の出版物との大きな違いに、資料の粒度がある。図書館は基本的に出版物を単位として利用者に提供してきている。ところが、ネットワーク上での雑誌タイトルのような単位もあれば、1 件の論文、あるいはその中の図表と、管理とアクセス単位は様々である。そのため、大粒度から細粒度まで、様々な粒度でのコンテンツの検索とアクセスを支援することが求められる。また、コンテンツの検索から閲覧に至る過程を統一された環境で支援することが求められる。たとえば、デジタル化された 1 冊の本に関して、その知的内容に関するメタデータ、デジタルデータの構造に関するメタデータ、アクセス権限に関するメタデータ（一部分だけがアクセス制限されることがある）といった細粒度の管理が求められる。一方、利用者の環境も様々であり、携帯電話から高精彩ディスプレイをもつ端末環境まで多様である。こうした多様性に対応することが求められる。

#### (5) 隣接する領域：電子政府、教育と学習

デジタルライブラリには幾つかの主要な隣接分野がある。デジタルミュージアムは非常に近い隣接分野である。大きな隣接分野として電子政府分野、ネットワークを利用し

た教育と学習の分野がある。電子政府は政府行政情報から個人情報に至るまで非常に広い範囲の情報をカバーする。多様で大量の情報資源の発信から流通、保存に至る過程で必要とされる技術はデジタルライブラリと共通するものが多い。ネットワークを用いた遠隔教育、遠隔学習は大学等での教育だけに限らず、企業での教育にとっても今後の重要な課題である。教育学習コンテンツを適切に蓄積、組織化し、共有するにはデジタルライブラリの技術が求められる。NSFが進める National Science Digital Library (NSDL)では、デジタルライブラリの基盤作りのプロジェクトから、コンテンツ作りのプロジェクトまで、多数のプロジェクトの助成が進められている。また、単にコンテンツを蓄積提供するだけではなく、コンテンツの利用方法やコンテンツを結んで作るコースウェア（これ自身もコンテンツ）等を加えていくことが求められる。こうした機能は、教育や学習に特化しないデジタルライブラリにおいても利用者環境作りの上で重要な要素である。

#### **(6) 多様な相互運用性への期待**

現在、ネットワーク上には数多くのデジタルライブラリサービスが提供されている。ネットワーク上では地理的距離とは無関係にサービスを利用できるとは言っても、いくつものデジタルライブラリを渡り歩くのは面倒である。書誌データの検索では分散横断検索のためのプロトコル Z39.50 が広く用いられている。また、OAI ではメタデータハーベスティングによるメタデータの共有を進めている。OAI のおもしろい点は、Simple Dublin Core をメタデータ共有の基礎とすることで、メタデータの記述の詳細度を落とす可能性を許すことで、異なるメタデータ規則の下に作られてきたメタデータの共有を進めている点である。個別に作られたシステムをつなぐこと、異なる語彙の下に作られたシステムをつなぐことはかなり困難を伴う仕事である。システム（あるいはサービス）間の相互運用性に関し、深いレベルから浅いレベルまでいろいろなレベルを考えることが求められる。

#### **(7) その他**

インターネット上における全てのサービスと同様にセキュリティや利用者のプライバシー、知的財産権の保護はデジタルライブラリにとっても非常に重要な話題である。インターネットの自由度や利便性と安全性を同時に得ることは必ずしも容易ではない。たとえば、現在、国立国会図書館では法定納本制度に基づくネットワーク情報資源の収集に関する議論が進められているが、収集対象となる資料の性質の問題と収集のための技術の問題に加えて、従来制度ではカバーしきれない点も指摘されている。デジタルコンテンツの場合、簡単にしかも細粒度でのコピーができ、知識や情報の再生産が行いやすい反面、著作権の保護が問題になる。デジタルライブラリのパーソナル化サービスが進められているが、従来図書館では個人情報の保護は思想信条の自由の下に最も重視されてきたことのひとつである。こうした相反する要求を満たすための制度とそれを支える技術の開発が

求められる。

#### 3.5.5 おわりに

WWW と Mosaic によるインターネットの大爆発から 10 年近くが立とうとしている。この数年の間に、電子ジャーナルをはじめ図書館現場でのデジタルライブラリサービスはずいぶん進歩したと感じる。利用者の情報環境とそこでの行動はこの間にずいぶん変化した。「電子図書館というと古い資料を電子化してネット上で見られるようにすることですか」という認識はかなり減ったように思える。これは、電子ジャーナルをはじめ、図書館で提供されるコンテンツが増え、デジタルライブラリサービスが進歩したからであろう。また、「電子図書館ができると本や図書館はなくなるのですか」という質問もほとんど聞かなくなった。紙の本と同様な使い勝手（大きさ、軽さ、耐久性など）を持つ出版媒体が出てくるまで、紙の本がなくなるとは思えないので、この質問に対しては「そんな心配はないですよ」と答えることにしている。

図書館の機能、すなわち情報を必要とする人に適切な情報を提供することは社会の中のどこにでも組み入れられている。「図書館」という名前と呼ばれなくてもそうした機能を持っている組織は多くある。インターネット上でのサービスのミドルウェア化が進むとすれば、デジタルライブラリも同様にミドルウェア的に利用できるようになる必要がある。大きなサービスを提供するソフトウェアシステムを構成する一部として組み込まれる完全自動のデジタルライブラリも出てくるであろう。一方、デジタルライブラリにはエンドユーザ向けのサービスを提供する、いわば応用システムとしての役割のみならず、様々なサービスの影で働くミドルウェア的な役割を演じることも求められると思われる。そうしたばあい、人間をも含めて構成される社会システムの中でのミドルウェア、いわば社会的ミドルウェアとしての役割が求められるのであろう。

「理想の電子図書館」では本も雑誌もビデオもすべて電子化され、図書館員との対話や図書館でのイベントにもネットワーク経由で参加でき、「館」に行くのと同水準のサービスが受けられるようになるべきなのであろう。また、ことばの壁を越えて、ビデオデータであろうと、新聞記事であろうと、あるいはホームページであろうとメディアの違いを超えて、利用者に適切な情報を提供することができるのが「理想の電子図書館」なのであろう。その意味からは「理想の電子図書館」はまだ近くにきているとは思えず、今後もチャレンジすべきものであると思える。

最後に、筆者にとって人間主体の知的情報技術に関する調査研究ワーキンググループから得られる情報や議論は非常にありがたいものであった。こうした機会を下さった先端情報技術研究所の内田俊一所長、本ワーキンググループの奥乃博主査（京都大学大学院情報

学研究科・教授)をはじめとして、先端情報技術研究所の皆様にご感謝の意を表したい。

[参考文献]

- [1] Schauble, P., Smeaton, A.F., (eds), “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>, 1998.10
- [2] Greenstein, D., “Digital Library Federation, DLF strategy and business plan”,  
2000.9, <http://www.diglib.org/about/strategic.htm>
- [3] 杉本重雄, ” Digital Libraries, Metadata, Interoperability” ,  
<http://www.icot.or.jp/FTS/REPORTS/H13-reports/H1403-AITEC-Report3/AITEC0203R3-html/AITEC0203R4-ch3-6.htm>, 2002.3 (これをもとに加筆したものが筑波大学附属図書館・電子図書館シンポジウム報文集にある。  
<http://www.tulips.tsukuba.ac.jp/pub/dlsympo/tutorial.pdf>)
- [4] Chien, YT., Whither digital libraries? The case of a "billion-dollar" business, 筑波大学知的コミュニティ基盤研究センター研究談話会資料, 2002.11
- [5] Gonçalves, M.A., Fox, E.A., “5SL: a language for declarative specification and generation of digital libraries”, Proc. Of JCDDL2002, pp.263-272, 2002.7
- [6] DC-2002 の二つの論文 (Atarashi, R., 他, “Policy Control Network Architecture using Metadata” , Kawarasaki R., Kishigami.J., “Metadata associated Network Services and Capabilities” ,  
<http://www.bncf.net/dc2002/program/social.html#posters>, 2002.10)
- [7] Lagoze, C. and Hunter, J., “The ABC Ontology and Model, Journal of Digital Information”, vol.2 issue 2, 2001.11,  
<http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Lagoze/>
- [8] 杉本重雄, “電子図書館－概要と課題－” ,  
<http://www.tulips.tsukuba.ac.jp/pub/dlsympo/tutorial.pdf>, 2003.1
- [9] 杉本重雄, “Digital Library に関する参考資料リスト” ,  
<http://avalon.ulis.ac.jp/~sugimoto/References/>

## 3.6 セマンティック・トランスコーディング —より実用的な“Semantic Web”に向けて—

長尾 確 委員

### 3.6.1 はじめに

次世代の高度な知識共有のインフラであるSemantic Web に不可欠なものは、ユーザが自由にコンテンツを作成し、さらにその共有化を促進し、知識として再利用可能にするためにコンテンツを意味的に拡張するツールやプラットフォームである。われわれは、これまでセマンティック・トランスコーディングという枠組みにおいて、現在のWeb の若干の延長線上に、Semantic Web と同等の機能を有する仕組みを研究してきた。それは、主に以下の3つのシステムから成っている。

1. 既存のWeb コンテンツに対して、機械による内容理解を促進する補足情報(アノテーション) を付与するツール
2. アノテーションデータを管理・共有するためのサーバ
3. アノテーションデータを用いてコンテンツを動的にユーザに適合させるためのトランスコーディングプロキシ

これらのシステムを実現した経験に基づいて、Semantic Web の早期実現のために、近い将来に、われわれが何を、どのように行うべきか、に関する一つの提言を行う。また、動画コンテンツに対するアノテーションとそれを用いた応用についても述べる。

### 3.6.2 Semantic Webのためにやるべきこと

Semantic Web[10] はグローバルな知識共有のインフラを目指すアプローチであるが、そのような試みがそう簡単にうまくいくはずがないのは、これまでの人工知能(特に、知識工学)、あるいは経営工学における知識管理などの分野における方法論の多くの失敗から考えても明らかである。ではどうしたら、この困難な問題に対処できるだろうか。

筆者の考える「現在の」Semantic Web の主な問題点は以下の通りである。

1. RDF やOWL 等のSemantic Web を構成する記述言語の仕様が必要以上に複雑になっている。
2. メタデータの作成の方法論が確立していないため、何をどう作ればよいのかわからない。
3. 意味的な検索以外の具体的な応用についてほとんど述べられていない。
4. マルチメディアデータのようなテキスト以外のコンテンツの意味構造化についてはまったく述べられていない。

以下では、これらの問題についての具体的なアプローチについて述べる。キーとなる概念は、アノテーションによる情報の階層化とトランスコーディングによる情報の個別化である。当然ながら、これらはXMLのような、情報を相互運用するための枠組みを前提としている。

アノテーションとトランスコーディングによって、内容および利用の観点から拡張されたWebは、「現在における」Semantic Webの一つの具体例と考えられる。これは、人間と機械がよりよく助け合って利用するためのWebである。システムの内部に人間が上手に関わっていくための仕組みがないと、知識共有のような高度なシステムはうまく機能しないだろう。

アノテーションは、従来のデジタルコンテンツを知的コンテンツとするための最良の手段である。それは、人間が、自分自身あるいは他者の創り出したコンテンツを再評価し、価値あるものとそうでないものを見分ける良い機会が与えられるからである。コンテンツを人類共有の財産とするためには、やはり、責任を持って、そのコンテンツを吟味する人間が必要であろう。アノテーションとは、まさにそのような責任の所在を明らかにし、内容にさらなる価値を与えていく仕組みなのである。

また、トランスコーディングは、コンテンツのアクセシビリティ(ユーザの身体的特性やスキル、使用するツールなどによらずに適切にアクセスできること)を強化する手段である[1]。これによって、コンテンツは真に人類共有の資源となる。

アノテーションは人間と機械の構成するシステム全体が賢くなっていくための仕組みである。この場合の機械とは、あらかじめプログラムされた手続きを文脈に応じて選択的に実行する自律的なシステム、すなわちエージェントである。エージェントをある程度以上に複雑にする代わりに、コンテンツの方をアノテーションによって、人間がエージェントにとって都合の良い形に変えていければ、人間とエージェントとコンテンツが構成するシステム全体をより高度にすることができる。つまり、コンテンツそのものがより理解しやすくなれば、それを扱うエージェントが可能なタスクもより高度になるだろう。

エージェントはアノテーションの付与されたコンテンツを対象にすることによって、単純な手続きを繰り返すだけで、より高度なサービスを提供できる。これは、見かけ上、エージェントが賢くなったように見えるが、実際はコンテンツそのものが(人間の不断の努力によって)賢くなっているのである。

このようになって初めて、人々はエージェントの価値を認めて受け入れていくだろう。そして、情報の収集や分類などのタスクはエージェントに任せて、より創造的な仕事に専念できるようになると思う。

### 3.6.3 アノテーション：デジタルコンテンツの階層化

図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付け加えるメタ(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。

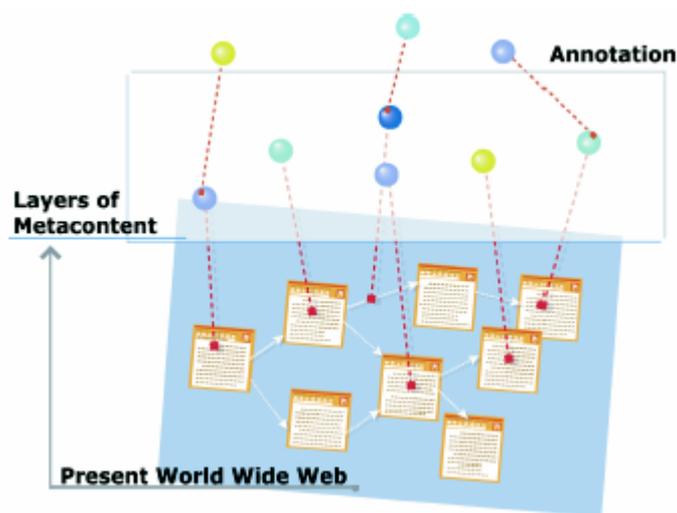


図1 アノテーションによるWebコンテンツの拡張

この図にあるように、従来のWebコンテンツは一枚の平面上に存在する要素群として捉えることができる。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XLink[7] 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。

アノテーションによる階層化の手法を用いて、具体的にはHTML 文書などのWeb コンテンツが抱える、以下の3つの課題を解消できるだろう。

1. HTML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、RDF[8] を用いることができるが、その方法論は。
2. 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。

3. 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。

#### 3.6.4 トランスコーディング：デジタルコンテンツの個別化

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、進歩を遂げていくことは間違いない。これまでは、とに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。

デジタルコンテンツの高度利用の主なものに、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。これらは、ともにコンテンツの個別化の例である。個別化はコンテンツの送受信がブロードキャストからポイントtoポイントになったことと大いに関係がある。

ここでは、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、オンラインコンテンツへのアクセスはPC経由で行なわれることが多い。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。たとえば、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なう必要がある。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他に、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。トランスコーディング技術を使えば、画面の表示機能やデータ伝送速度など、それぞれ違った仕様や制約をもつ多様な機器に対して、1つのコンテンツ・ソースから情報やサービスを提供できるようになる。コンテンツ・プロバイダやサービス・

### 3.6 セマンティック・トランスコーディングーより実用的な“Semantic Web”に向けてー

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。コンテンツ・プロバイダは、現状のようにPC 向けと携帯電話向けのコンテンツを作り分ける必要がなくなる。

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、筆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[4]。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構造や語彙情報をアノテーションとして関連付けることによって、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。たとえば、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付ける手間が増すが、誤解なく伝達すべき重要な情報にはアノテーションを付与して、より適切な形で伝達・共有すべきだろう。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して機能する。

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する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くこともできる。

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・翻訳を行なう。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。

#### 3.6.5 トランスコーディングの仕組み

セマンティック・トランスコーディングを実行する複数のソフトウェア・モジュール(トランスコーダ)は、HTTP プロキシ上で機能するプラグインとして実装した。トランスコーダを制御するHTTP プロキシをトランスコーディングプロキシと呼ぶ。

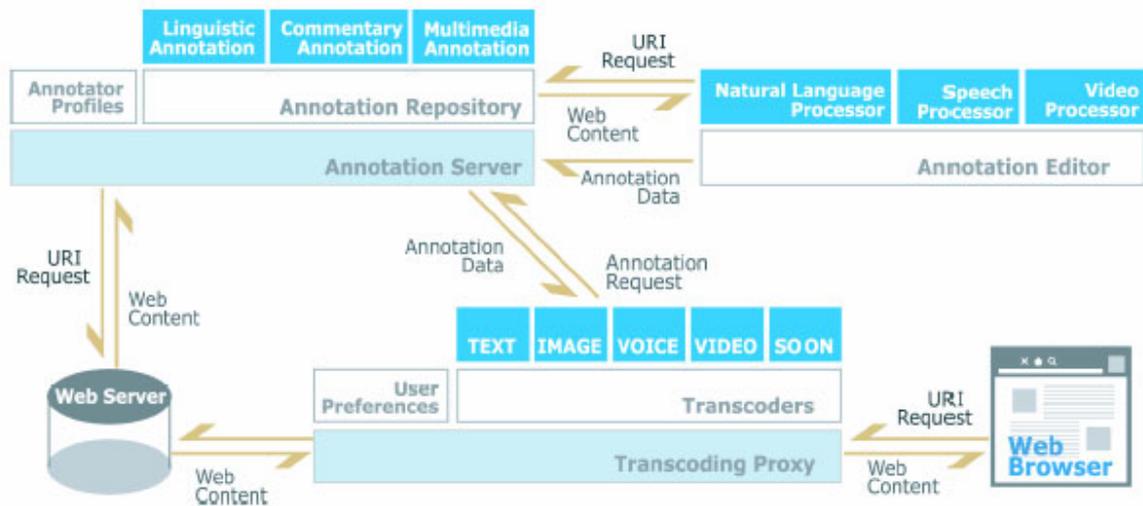


図2 セマンティック・トランスコーディングシステムの構成

図2はセマンティック・トランスコーディングシステムの構成を表している。トランスコーディングプロキシを中心とした情報の流れは次のようになる。

1. クライアントのWebブラウザからURL とクライアントID を受け取る。
2. Web サーバにURL の示すWeb ページをリクエストする。
3. Web ページを受け取ると、そのハッシュ値を計算する。
4. アノテーションサーバにURLに関連するアノテーションデータを要求する。もし、アノテーションデータが見つかったら、アノテーションサーバからデータを受け取る。
5. データを受け取ると、データのハッシュ値とWebページのハッシュ値と比較する。
6. 同時にクライアントID に基づいてユーザ情報を検索する。ユーザ情報がない場合は、ユーザから与えられるまでデフォルト設定を使う。
7. ハッシュ値を照合したら、アノテーションデータとユーザ情報に基づいて適切なトランスコーダを起動する。
8. 加工したコンテンツをユーザのWeb ブラウザに送信する。

トランスコーディングプロキシは、実装環境としてIBM Almaden Research Center の開発したWBI (Web Intermediaries) を使用した[2]。このWBIを利用したトランスコーディングプロキシには、以下の3つの主要な機能がある。個人情報の管理、アノテーションデータの収集と管理、そしてトランスコーダの起動と結果の統合である。

個人情報の管理を行なうには、まずアクセスしてきたユーザを特定する必要がある。ユーザの特定にCookieを使う。個人情報を管理するIDを、Cookieデータとしてユーザに渡す。

### 3.6 セマンティック・トランスコーディングーより実用的な“Semantic Web”に向けてー

これにより、ユーザのアクセスポイントに関係なくユーザの特定が行なえる。ただし、既存のWeb ブラウザは、Cookieをセットしたサーバに対して、そのCookieを渡すものであり、プロキシのCookie 利用は考慮されていない。通常プロキシは、ホスト名とIPアドレスのみによってユーザを識別する。そこで、ユーザが個人情報をセットした時に、Cookie 情報(ユーザID) と個人情報を関連付け、一方、アクセスポイントの変化ごとにIP アドレスとホスト名、Cookie 情報(ユーザID) を関連付け直す。これによりIP アドレスが変化してもユーザの特定が行なえる。

トランスコーディングプロキシは、アノテーションサーバと通信して、アノテーションデータを入手する。アノテーションサーバは複数存在することができるので、それぞれのサーバの管理するアノテーションデータのインデックスを定期的につけておく。このインデックスを、どのアノテーションサーバからデータを入手すべきかを判断するときに役立つ。トランスコーディングプロキシの最も重要な役割は、個人情報とアノテーションデータに基づいてコンテンツを加工することである。コンテンツの加工は、必要なトランスコーダを起動し、その結果を統合することによって行なう。現在、開発済みのトランスコーダは、テキスト文、画像、音声、映像にそれぞれ対応したものである。これらのトランスコーダは、直列あるいは並列に結合することで、複合的なトランスコーディングが実現できる。たとえば、文書を要約後に翻訳して、さらに音声化するなどの一連の処理をトランスコーダの使い分けにより行なう。

#### 3.6.6 提言

セマンティック・トランスコーディングで用いるアノテーションは、主に文書の言語構造、マルチメディアの内容に基づく構造化情報、任意のコンテンツに対するコメント情報などである。これらは、ある種のリテラシーがあれば誰にでも作成可能な情報である。そのようなリテラシーは、アノテーションエディタと呼ばれるツールを使っているうちに自然に獲得されていくような仕組みにすべきだと思っている。

一方、「現在の」Semantic Webにおける主な(メタ)コンテンツは、RDFによるグラフ構造によるメタデータは何をどう作ればよいのかよくわからないし、OWL[11]によるオントロジカルなデータは、さらに何をどう記述すればよいかわからない。やはり、具体的なコンテンツに関して、自然に追加できるような内容でないと動機的にも技量的にもとっかかりがないのである。

アノテーションは、コンテンツと乖離したトップダウン的なものであるべきではないし、段階的により高度なものに発展させていく必要があるだろう。そのために必要なのは、コンテンツをサーバ側で変換して配信する場合にも、オリジナルデータへのポインター(デー

データベースURL やレコードID など) を変換後のコンテンツの該当する部分に挿入し、アノテーションがオリジナルデータに直接関連付けられるように工夫することである。

また、当然ながら、**Semantic Web** は、現状の**Web**とシームレスに統合できるものでなければならない。トランスコーディングプロキシはサーバとクライアントの「中間」で処理を行なうため現在の**Web** のアーキテクチャに自然に統合される。ここで必要なのは、**URL** のようなコンテンツのポインタを要求するだけでなく、どのプロキシにどのような変換を必要するかということも含めて要求とすることである。これは、現在ではブラウザの機能とトランスコーディングプロキシのデータベースを用いることで解決しているが、たとえば、トランスコーディングのためのプロファイルを**XML** を用いて標準化して、**SOAP (Simple Object Access Protocol)**[9]等でリクエストを送るようにすれば、より一般化できるだろう。

### 3.6.7 マルチメディアコンテンツのアノテーションとその応用

**Semantic Web** には、具体的なコンテンツの意味構造化に関する視点がまったく欠けていると言わざるを得ない。文書に対する言語構造のアノテーションに関しては別稿に譲ることにして、ここでは、マルチメディアコンテンツへのアノテーションとその応用について述べる。

映像や音声を含むマルチメディアコンテンツは、テキストコンテンツに比べて、内容に基づく処理が極めて困難である。そこで、マルチメディアコンテンツの検索変換を行う上で必要となるインデックス情報を生成・加工し、これらをアノテーションとしてオリジナルコンテンツに関連付けておく、という手法が考えられる。われわれは、特にビデオデータを対象に研究を進めている[5, 18, 3]。

ビデオデータがさまざまな視聴環境で利用される状況を考える場合、その対処方法としては以下の2通りが考えられる。

- 異なる視聴環境ごとに適合させたデータをあらかじめ用意しておく
- オリジナルデータを環境に合わせてオンデマンドに変換する

セマンティック・トランスコーディングが採用しているのは、後者の方法である。すなわち、コンテンツ提供者はデータを1種類のみ用意すればよく、利用者側の環境や要求に応じてサーバがコンテンツ変換を行うようなシステムである。この場合の変換は、画面サイズやビットレートなどの視聴環境のパラメータを用いて行なう単純な変換の場合と、要約や翻訳のようなユーザの好みと(アノテーションによってもたらされる)コンテンツの意味情報を同時に考慮して行なう複雑な変換の場合がある。セマンティック・トランスコーディングは、特にアノテーションに基づくさまざまなコンテンツ加工を特徴としている。

### 3.6 セマンティック・トランスコーディングーより実用的な“Semantic Web”に向けてー

ただし、これを可能にするためには、コンテンツ変換を容易にするアノテーションを適宜付与しておく必要がある。コンテンツ変換を容易にするアノテーションとは、すなわちそのコンテンツの内容記述である。ビデオデータであれば、データ中のシーンやシーンに含まれる対象、人物の発話内容などである。

以下では、われわれが試作しているマルチメディアコンテンツ検索システムの概要について述べ、アノテーションに基づいたビデオ検索を実現する手法を現状のシステムを例に提案、解決すべき問題について考察する[18]。

#### 3.6.7.1 先行研究・事例

完全なビデオデータを対象とした研究や事例は少ないが、単語と動画像の相互検索(ただし、音声ではなくクロズドキャプション(字幕)のような映像に付加されているテキストを利用)[15]、話者と発話内容の同時検索[17]、Web上の文書と画像のクロスメディア検索[14]、画像から画像の検索、ニュース音声のトランスクリプトに対する検索[16]など、これまでに多くの研究がなされており、岡らのグループが研究・開発したCrossMediator[13]は、現在実用化されているマルチメディア検索システムの1つである。彼らは、音素や濃度ヒストグラムといったデータの時系列特徴量をインデックスとして用いて検索を行っている。

音声データを音素系列として記述する方法は、音声認識誤りによる検索精度の劣化や未知語に対して強いという利点を持つ一方で、同音語や単語境界誤り、短い単語による精度劣化や、内容に基づいた検索や要約が難しいという欠点がある。

個人が家庭で録画したデータに対して検索を行う場合には、このような手法が適していると考えられるが、Web上で無数のデータが公開されるような場合、それらに対して検索・要約を行うためには十分な内容記述が必要である。

#### 3.6.7.2 ビデオ検索におけるアノテーションの有用性

##### 3.6.7.2.1 テキスト検索・イメージ検索との比較

例として、次のような検索要求を考えてみる。「テロで航空機がビルに激突したらしい」テキスト検索ではどうだろうか。検索エンジンGoogle([www.google.com](http://www.google.com))でAndをとる単語を増やしながら検索した結果を表1に示す。

表1 Google によるテキスト検索例

検索キーワード	ヒット件数
テロ	376,000
テロ, ビル	76,500
航空機, ビル	30,700
テロ, 航空機	25,500
テロ, 航空機, ビル	11,400
貿易センタービルに航空機が激突	90

検索キーワードとしてユーザが与える語数は平均1,2 語と言われているが、テキストコンテンツの量がビデオコンテンツとは比較にならないほど多いことを考慮しても、検索結果としてユーザが確認・視聴するためには、相当の絞り込みや提示方法の工夫が必要であることが予想される。

一方、イメージ検索について考えてみると、この例の場合、「航空機」、「ビル」、「爆発・炎上」といったイメージあるいは単語列から検索を行うことになる。イメージからの検索は精度面から考えて現状では非常に困難であるため、ここでは単語列からの検索のみについて考える。同様にGoogleのイメージ検索を利用した結果を表2 に示す。

表2 Google によるイメージ検索例

検索キーワード	ヒット件数(正解)
テロ	1,700 (計測不能)
テロ, ビル	117 (15)
航空機, ビル	44 (6)
テロ, 航空機	37 (6)
テロ, 航空機, ビル	11 (4)

上記の正解件数は、テロに関係するものをカウントしており、この中で実際に航空機が写っているものはわずかであった。これは、イメージの説明文としてニュース記事の文章を利用していることにより、航空機という単語が補完された結果であると考えられる。逆に、ニュースのように公共性の高い情報でなく十分な説明がなされていない場合には、イメージの検索は容易ではない。これはビデオ検索についても同様に当てはまる。

### 3.6 セマンティック・トランスコーディングーより実用的な“Semantic Web”に向けてー

#### 3.6.7.2.2 ビデオデータの特徴

ビデオデータの最大の特徴は、映像が持つ過去の事象の時間的連続性である。テキストには、それを扱う人間の知識や状況が反映され、時間軸の前後やスキップも容易である。イメージには、瞬間の情報は凝縮されているが、時間情報が欠落している。

ビデオデータをテキスト(実際には音声)とイメージの合成と考えると、補完の最も難しい情報は時間の経過によって我々が得ることのできる真実である。

先ほどのテロの例を考えてみる。同様の例として、イタリア・ミラノで起きた小型機事故と合わせて、テロという事実情報に着目してみる。

米国同時多発テロの場合は、事件の初期段階では、テロという言葉は断定的には使われていない。大統領の声明により、その瞬間から事件はテロであるとの認識が定まったわけである。このため、1機目と2機目の激突のシーンに対してテロという情報を与えるには、その真実を理解した人間の介在が必要になる。

一方、イタリア・ミラノ小型機事故の場合は、事件の第一報で、「テロ攻撃の可能性が非常に高い」との見方が示されたと報じたが、その後に事故説と自殺説が出た。このため、事件直後のテロという情報が誤っていることを与えるためには、やはり同様に人間の介在が必要である。

このように、ビデオデータの場合は、必ずしもその時点での音声が真実を表しているとは限らない。そのため、後からアノテーション情報を付与することが可能な仕組みが非常に重要であると考えられる。

#### 3.6.7.3 アノテーションを利用したコンテンツ変換例

アノテーションが付与されているデータは、様々な自然言語処理が可能である。



図3 ビデオドキュメント



図4 要約されたビデオドキュメント



図5 PDA 用に変換されたビデオドキュメント



図6 携帯電話用に変換されたビデオドキュメント

その例として、ビデオデータからHTML ドキュメントへのコンテンツ変換を行った例を図3 に示す。また、図4 は、このビデオドキュメントを要約した例を示す。さらに、図5 と図6 に、それぞれPDA 用と携帯電話用に変換されたビデオドキュメントを示す。

このように、検索されたビデオデータを、ユーザの利用環境や嗜好に応じて適宜変換して提示することも、アノテーションを付与することによって可能になる。一度に多くの検索結果を視聴することのできないビデオデータにとって、これらの処理を可能にするアノテーション情報は必要不可欠である。

### 3.6.8 アノテーションに基づくビデオ検索システム

#### 3.6.8.1 システム概要

我々が提案するアノテーションに基づいたビデオ検索システムの概要を図7 に示す。

システムは、コンテンツ提供者側によるアノテーション生成処理、ユーザからの検索要求に対する検索処理、ユーザ環境に応じた検索結果の変換処理、の大きく3つに分けられる。検索結果の変換処理については、セマンティック・トランスコーディング[4] によって行う。

### 3.6 セマンティック・トランスコーディングより実用的な“Semantic Web”に向けて

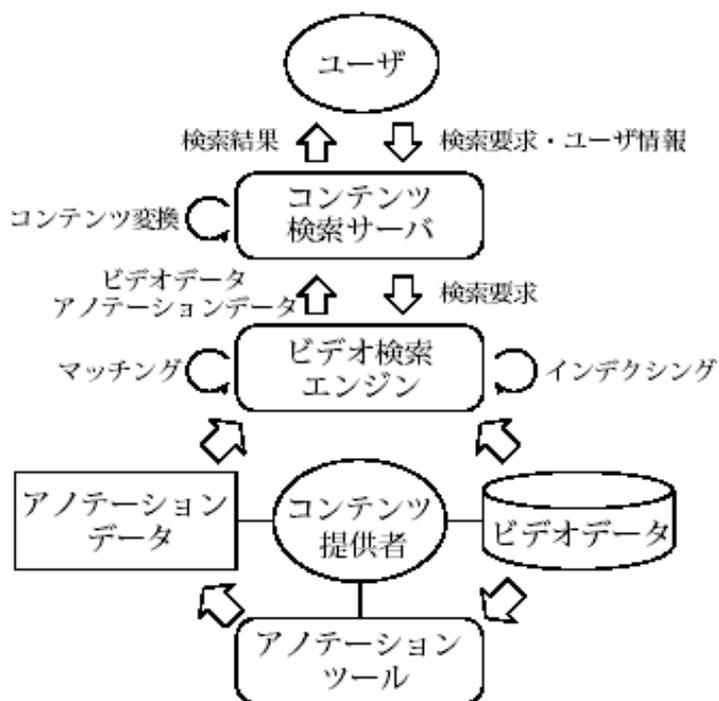


図7 ビデオ検索システムの概要

#### 3.6.8.2 アノテーション作成

我々は、ビデオデータ中の発話情報、シーン情報、シーン内オブジェクト情報をアノテーションとして生成・編集・関連付けすることを可能にするツールとして、多言語ビデオトランスクリプタを開発している(図8)。

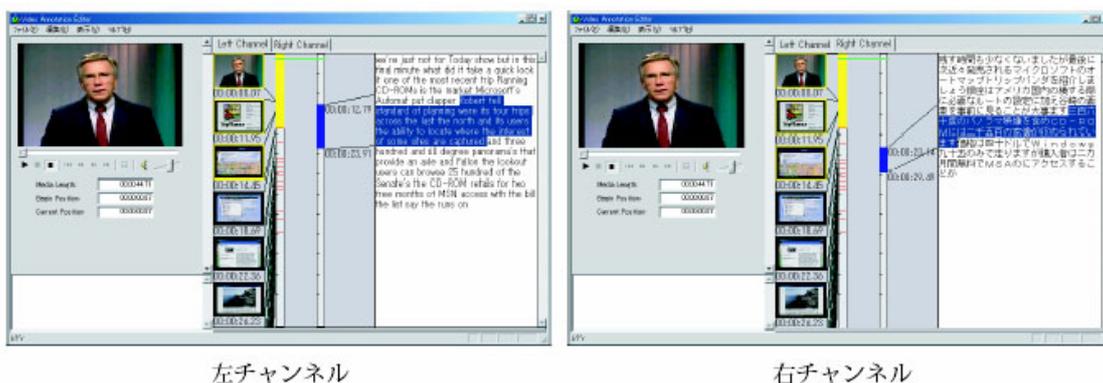


図8 多言語ビデオトランスクリプタ

ビデオデータへのアノテーションの付与は、図9に示される手順にしたがって行われ、作成したアノテーションデータは、XMLで記述される。

アノテーションによって扱う内容記述は、シナリオにおけるト書きのような情景描写(シ

ーン) と登場人物の台詞に相当する発話文(トランスクリプト)、およびフレーム内に登場するオブジェクトの記述から構成される。

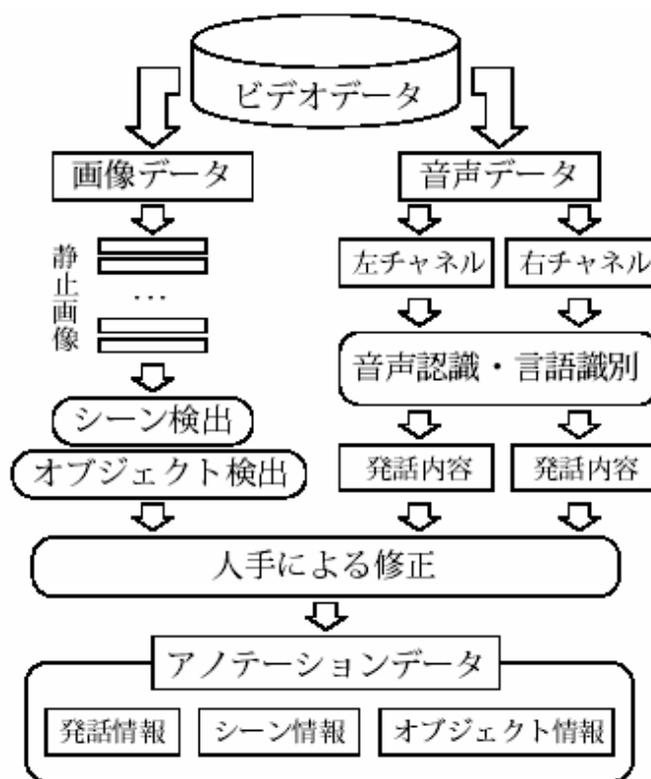


図9 ビデオアノテーションデータ生成までの手順

発話情報にはタイムコードと発話内容、シーン情報にはタイムコードとインデックスタイトル、オブジェクト情報にはタイムコードと名称、説明、矩形、リンク情報等が含まれる。

現在、音声認識、言語識別、シーン検出は自動的に処理されているが、認識誤り訂正、シーン統合、オブジェクトの切り出し、シーン記述等は、人間が行う仕組みになっている。

### 3.6.8.3 アノテーションに基づく検索処理

生成されるアノテーションデータには、代表シーン画像やオブジェクト画像も含まれるが、これらは検索後のコンテンツ変換や要約時に使われ、検索時には利用されない。検索の原理としては通常のテキスト検索と同じであり、代表的なベクトル空間モデルを用いて類似度を計算する。

ビデオデータの場合は、類似度の高いデータの中から目的のシーン(タイムコード)を検出することが要求されるため、類似度計算はビデオファイル全体とビデオファイル中に含

### 3.6 セマンティック・トランスコーディングーより実用的な“Semantic Web”に向けてー

まれる全シーンに対して行われる。

ビデオファイル  $V_i$  が  $m$  個のシーンから構成されるとする。

$$V_i = \{S_1, S_2, \dots, S_m\}$$

このとき、各シーンの特徴ベクトルを  $\overline{Vi(Sj)}$  とすると、ビデオ  $V_i$  の特徴ベクトルは、シーンの時間長とターム数によって正規化された次の式で表される。

$$\overline{Vi} = \sum_{j=1}^m \frac{l_j}{L} \cdot \frac{n_j}{N} \cdot \overline{Vi(Sj)}$$

( $l_j$ : シーン時間長、 $L$ : ビデオ時間長、 $n_j$ : シーン中ターム数、 $N$ : 全ターム数)

各シーンの特徴ベクトル  $\overline{Vi(Sj)}$  は、発話情報  $Sj(t)$ 、シーン情報  $Sj(s)$ 、オブジェクト情報  $Sj(o)$  中に出現するタームのタームベクトルと、各々の情報に対する重み係数  $\alpha$ 、 $\beta$ 、 $\gamma$  によって以下の式で表される。

$$\overline{Vi(Sj)} = \alpha \cdot \overline{Sj(t)} + \beta \cdot \overline{Sj(s)} + \gamma \cdot \overline{Sj(o)} \quad (0 \leq \alpha, \beta, \gamma \leq 1)$$

タームベクトルを構成する各タームの重みは、 $TF \cdot IDF$  法によって求められる。検索は、まず全ビデオファイルに対してビデオの特徴ベクトルを用いて類似度計算を行い、次にその中の上位  $R$  件に含まれる全シーンに対して同様に類似度計算を行いランク付けする。最終的に、ランキング上位の  $r$  シーンを検索結果として出力する。

#### 3.6.8.4 システムの問題点

多言語ビデオトランスクリプタにより、人間の介在する半自動的なビデオアノテーション作成の支援が可能になったが、アノテーションを付与する際の問題として、次の2点が挙げられる。

- 内容記述量、精度
- 内容記述コスト

これらは、処理の自動化と検索・要約等の実用性の意味において互いにトレードオフの関係にある。そこで、以下ではそれぞれの問題を解決する上で今後取り組む必要のある技術的課題について考察する。

##### 3.6.8.4.1 内容記述量・精度

すでに述べたとおり、マルチメディアコンテンツは内容に基づく処理が極めて困難であることから、アノテーションツールを用いた内容記述は、ビデオ等のコンテンツの検索・要約を実現するために重要な役割を果たすと考える。しかし、内容記述が乏しかったり、機械処理に頼ることで内容記述の精度が悪いと、その後の検索や要約結果の精度を低下さ

せる原因となる。

図10 に示すグラフは、音声認識誤りをシミュレーションして、実際の検索精度にどの程度影響を及ぼすかを分析するとともに、検索において期待されるアノテーション情報の質を予測したものである。具体的には、RWCP 検索・要約用ニュース音声データベース(192 記事、3,517 異なり単語)[12] を用いて、検索対象となる記事中に含まれる単語をランダムに除去した場合と、文書の特徴付けに有効でない単語を *IDF*(逆文書頻度) 値により優先的に除去した場合の2 通りについて検索実験を行い比較した。

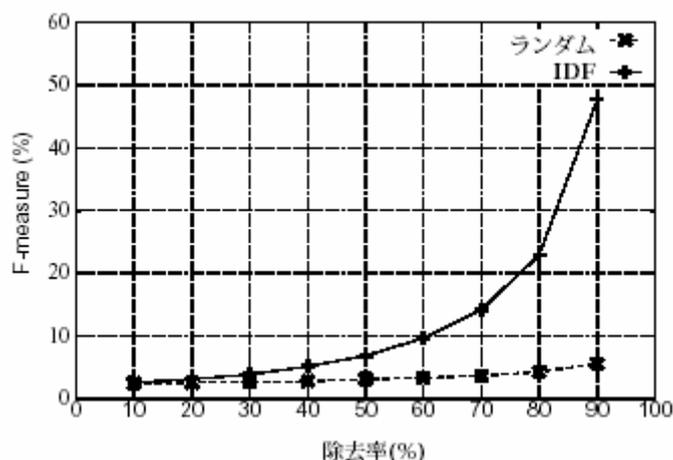


図10 単語除去率と検索精度の関係

評価尺度には、再現率と適合率から求められるF-measureを用いた。入力クエリーにオリジナルの記事を用いており検索対象数も少ないことから、除去率が小さいときには両者に大きな差は見られないが、除去率90%になるとその差は40%程度になる。すなわち、検索に有意な10%以上の単語を音声認識の脱落誤りによって失うと、検索精度に多大な影響があることがわかる。さらに、置換誤りや挿入誤りが起こると検索精度に対する信頼性は著しく低下する。

そこで、内容記述量を向上させると同時に内容記述精度そのものを向上させる技術が必要である。前者については、アノテーションツールによる内容記述処理の自動化によってある程度は達成されているが、オブジェクト認識・トラッキング等の画像処理や発話内容に基づく談話構造化など、改良の余地はまだ多く残されている。後者については、発話文認識精度向上のための言語モデルの修正や、認識結果に対する事後処理としての認識誤り訂正処理が挙げられる。

特に、入力音声については、話題に応じた言語制約を加えることにより精度改善が期待される。音声対話システムの場合はリアルタイム性を重視するために、通常音声認識処理

### 3.6 セマンティック・トランスコーディングより実用的な“Semantic Web”に向けて

は1度しか行われませんが、本研究のようなオフライン処理の場合は、音声認識と言語モデル修正を繰り返し行うことにより、認識結果の精度改善を図ることが可能であると考えられる(図11)。言語モデルの修正には、認識結果以外に、クローズドキャプション、Webテキスト、アノテーション情報等の外部知識が利用できる。

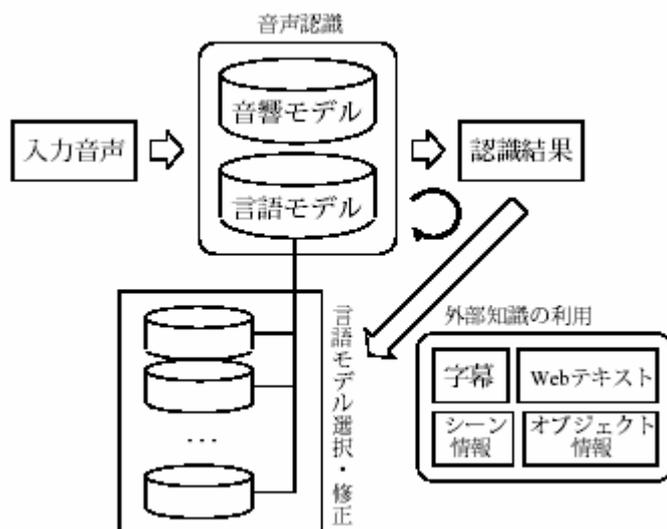


図11 話題に応じた言語モデルの選択・修正処理例

また、内容記述量・精度を客観的に示す評価尺度が必要である。コンテンツ時間長に対する内容記述量や、発話文認識結果の認識精度、構文としての正しさ、等を統一的な評価尺度で測った上で検索時にその評価量を導入することにより、検索結果に対する信頼性を向上させることが可能であると考えられる。

#### 3.6.8.4.2 内容記述コスト

アノテーション付与において内容記述にかかるコストは、本システムにおける最大の課題である。内容記述にともなう機械処理の精度を極限もしくはタスク達成に必要な精度まで向上させることが最も重要であることは言うまでもないが、現実的に実用レベルのアノテーションを付与するためには人間の介在が不可欠であるため、人間がスムーズに作業を行うための入力支援やインターフェースの改良も必要である。

現時点の多言語ビデオトランスクリプタを用いて、Windows PC (CPU: PentiumIII 800MHz、Memory:512MB) 上でアノテーションを行った場合にかかるコストを表3に示す。アノテーション作業は、まずツールの使用方法を説明した後にテストファイルを用いて10分間操作に慣れてもらい、その後、新しいビデオファイル(88秒)に対して作業を行ってもらった。これを計3人分計測した。シーンとオブジェクトについては、作業による

って結果が異なるため、作業時間をシーン数、オブジェクト数で平均してある。

表3 内容記述コスト

分類	作業者			平均
	A	B	C	
(a)	58			
(b)	88	70	74	77
(c)	23	30	17	23
(d)	528	394	425	449
シーン数	5	4	4	
オブジェクト数	2	3	2	
(b-d)	1014	764	755	844
計	1072	822	813	902

(認識結果の単語正解率78.5%, 正解精度73.9%)

(a) 機械処理(音声認識、シーン検出) 時間[sec]

(b) シーン統合・内容記述時間[sec]

(c) オブジェクト切り出し・内容記述時間[sec]

(d) 発話内容修正時間[sec]

表3より、音声認識精度が70~80%の場合、アノテーション作業を行う人間にかかる時間的なコストは、機械処理にかかる時間の約15倍、アノテーションを付与する対象データ時間の10倍程度であることがわかる。誰もが手軽にアノテーションを付与できるようにするためには、今後さらなる改良が必要であると思われる。

### 3.6.9 まとめと今後の課題

より実用的なSemantic Web に向けて、筆者の考えと提言を述べた。また、筆者のグループで研究開発しているマルチメディアコンテンツのアノテーションとその応用システムの概要について説明し、アノテーションに基づいたビデオ検索を実現する手法を具体的なシステムを例に提案した。また、現状のシステムの問題点として、アノテーションにおける内容記述量・精度と内容記述コストを挙げ、これら解決すべき問題について考察した。今後は、システムの問題点を改善しながら個々の技術を検討し、大量のデータに対して提案した検索手法の有効性を評価していきたいと考えている。

### 3.6 セマンティック・トランスコーディングーより実用的な“Semantic Web”に向けてー

また、昨年度のTREC-2001よりVideo Retrieval Track が導入され[6]、ビデオデータにおける類似検索の研究促進が期待されているので、こちらの動向にも注目しながら研究を進めていく予定である。

今後の課題には、もちろん、アノテーションの作成コストを下げることが含まれるが、その他に、アノテーションに基づく、コンテンツからの知識発見を実現することがある。近い将来には、Web上の情報検索には、既存の検索エンジンから、複数のコンテンツから新たな知識を得てその結果を要約して出力するような、いわば知識発見エンジンを使うようになるだろう。それによって、ハイパーリンクを集めた大量のリストの代わりに、短時間で容易に理解できるように要約されたコンテンツを読むことができるようになると思われる。また、映像や音声といったマルチメディアデータを含むデジタルコンテンツの効率的な検索も重要である。この場合の検索の質問には単なるキーワードではなく、音声あるいはテキストの自然言語文を用いることができるようになるだろう。

#### [参考文献]

- [1] Chieko Asakawa and Hironobu Takagi. Annotationbased transcoding for nonvisual Web access. In Proceedings of the Fourth International ACM Conference on Assistive Technologies (ASSETS 2000), pp. 172–179, 2000.
- [2] Steven C. Ihde, Paul P. Maglio, Joerg Meyer, and Robert Barrett. Intermediary-based transcoding framework. IBM SYSTEMS JOURNAL, Vol. 40, No. 1, pp. 179–192, 2001.
- [3] Katashi Nagao, Shigeki Ohira, and Mitsuhiro Yoneoka. Annotation-based multimedia summarization and translation. In Proceedings of the Nineteenth International Conference on Computational Linguistics (COLING-02), 2002.
- [4] Katashi Nagao, Yoshinari Shirai, and Kevin Squire. Semantic annotation and transcoding: Making Web content more accessible. IEEE MultiMedia Special Issue on Web Engineering, Vol. 8, No. 2, pp. 69–81, 2001.
- [5] Shigeki Ohira, Mitsuhiro Yoneoka, and Katashi Nagao. A multilingual video transcriber and annotation-based video transcoding. In Proceedings of the Second International Workshop on Content-Based Multimedia Indexing (CBMI-01), 2001.
- [6] Text Retrieval Conference (TREC). TREC-2002 Video Track, 2002. <http://www-nlpir.nist.gov/projects/trecvid/>.
- [7] W3C. XML Linking Language (XLink) Version 1.0, 2001.

<http://www.w3.org/TR/xlink/>.

- [8] W3C. Resource Description Framework (RDF) Model and Syntax Specification, 2002. <http://www.w3.org/TR/REC-rdf-syntax/>.
- [9] W3C. Simple Object Access Protocol (SOAP) 1.1,2002. <http://www.w3.org/TR/SOAP/>.
- [10] W3C. The Semantic Web Community Portal, 2002.<http://www.semanticweb.org/>.
- [11] W3C. Web-Ontology (WebOnt) Working Group,2002. <http://www.w3.org/2001/sw/WebOnt/>.
- [12] 伊藤克亘, 田中和世, 中沢正幸, 岡隆一. ニュース音声コーパスの構築. 日本音響学会講演論文集, pp. 171–172,1999.
- [13] 岡隆一, 西村拓一. パターン検索のアルゴリズム・マップ–CrossMediator を支えるもの–. 人工知能学会研究会資料, SIG-J-A101, pp. 1–6, 2001.
- [14] 森靖英, 岡隆一. WWW 上の文書・画像混在データのクロスメディア検索. 人工知能学会研究会資料, SIG-CII-2001-MAR-06, 2001.
- [15] 森靖英, 高橋裕信, 岡隆一. 画像と単語の相互検索手法. 人工知能学会研究会資料, SIG-CII-2000-NOV-17, 2000.
- [16] 西崎博光, 中川聖一. 音声入力によるニュース音声検索システム. 電子情報通信学会技術研究報告, SP99-108, pp.91–96, 1999.
- [17] 西田昌史, 緒方淳, 有木康雄. 話者と発話内容の同時検索に関する検討. 人工知能学会研究会資料, SIG-CII-2000-NOV-12, 2000.
- [18] 大平茂輝, 長尾確, 白井克彦. アノテーションに基づくビデオ検索システムの提案. 情報処理学会研究報告, SLP-43-6, pp. 33–39, 2002.

## 3.7 多感覚情報のデータベース化と情報サービスへの応用

加藤 俊一 委員

### 3.7.1 はじめに

多感覚を用いた情報通信・情報サービスの基盤技術として、多感覚を用いたヒューマンインタフェース、多感覚の情報のデータベース化、多感覚の情報の内容に基づく連想・検索の技術の必要性が高まっている。

本稿の範囲で扱う「感性」とは、人間が実世界のさまざまな事物に接した時に、どのようにセンシングし、これを類型化し、どのような概念のイメージや主観的な評価と結びつけるかの過程をいうものとする。

感性をモデル化する場合、従来は、あるメディアのコンテンツを対象に、それぞれのコンテンツごとに、かなりの数の教示用データを利用者から得て、統計的学習を行っていた。従って、利用者（被験者）の心理的・身体的負担は大きく、感性のモデル化技術を応用する上でのボトルネックとなっていた。また、人間は多感覚を利用して人とコミュニケーションを行っているが、1つの感覚を分析・モデル化し、ヒューマンインタフェースに適用している研究がほとんどであり、システムの開発は非常に個別的・限定的に行われている。このような現状をブレークスルーして、感性情報技術（本来は応用分野に横断的な技術である）を汎用性・一般性のある技術として確立し、モノ作りやサービス、ヒューマンコミュニケーションなどの様々な場面で利用可能とする新しい枠組みが必要となってきたのである。

本稿では、現行のモデル化の枠組みを踏まえながら、感覚情報の幅を視覚・聴覚・触覚・味覚・嗅覚の五感に拡大するとともに、より精密な感性のモデル化を行う上で重要な視点を整理して検討する。

### 3.7.2 感性の工学的なモデル化の新しい視点

多感覚の感性情報を工学的にモデル化し、種々のアプリケーションに応用できるようにするためには、以下のような新しい視点からの感性情報処理の枠組みが必要となる。

(a) 「多感覚＝マルチメディア」と短絡的にとらえて、画像情報・音情報・触覚情報のように、複数の感覚情報を並列的に処理するだけでは、多感覚を統合的に知覚する過程をモデル化することは難しい。それぞれの感覚の特性や、実世界の対象がそれぞれの感覚を通じてどのように知覚されるかなどの、個別の感覚と複数の感覚の統合的な枠組みの確立が必要である。

私見であるが、視覚情報・聴覚情報・触覚情報など、それぞれのメディアの特殊性と共通性を総合的に進めている研究は少ないように思われる。研究開発の効率を高めるためにも、個別のメディア、個別のコンテンツに対する感性情報処理アルゴリズムの適用可能な範囲、一般化の方法論などの検討が必要であろう。

(b) 感性のモデル化でしばしば用いられる統計的な学習法では、対象が多様で複雑になると、多数の学習用データに対して被験者の主観的な解釈を教示する必要があり、被験者の心理的・身体的な負荷は非常に大きくなる。多種多様なコンテンツに共通な概念・固有な概念に対応して、新しいコンテンツに対しても、既存の感性のモデルを利用して、出来るだけ少ない教示用データから、感性のモデルを構築できるようにする必要がある。

私見であるが、現状の感性情報技術の研究には、実際に運用する場面まで想定して被験者の負荷を考慮しているものは少ないように思われる。感性のモデル化のアルゴリズムが高性能でも、一人一人の利用者からの感性的な情報（教示のための情報）を心理的・身体的な負荷をかけずに得ることができなければ、モノ作りやサービスの現場に感性情報技術を導入することは難しい。

(c) 人間の感性的な応答は、同じ刺激に対して常に同じではなく、直前までに受容していた刺激や、同時に受容している他の刺激の影響を受けて、異なった応答を示す場合が少なくない。また、経験を積んでいくことにより、感性そのものが変化していく。このように感性の時間的な特性もモデル化の枠組みの中にも含める必要がある。

このように状況によって変わる応答の規則性を統計的な学習（あるいはデータマイニング）により求めるには、莫大な数の事例および教示データが必要となる。私見であるが、「何でもかんでも教示学習」的なアプローチでは、このような壁はブレークスルー出来ないと思われる。知識・記憶と感性の関係や、インタラクション・相互作用の定式化など、感性の構成的な分析法の確立も必要となろう。

### 3.7.3 感性の工学的なモデルの枠組み

以上の考察より、感性の工学的なモデル化の手法として、以下に示すような枠組みを考える。

感性の工学的なモデルの基本的な構造を図1に示す。

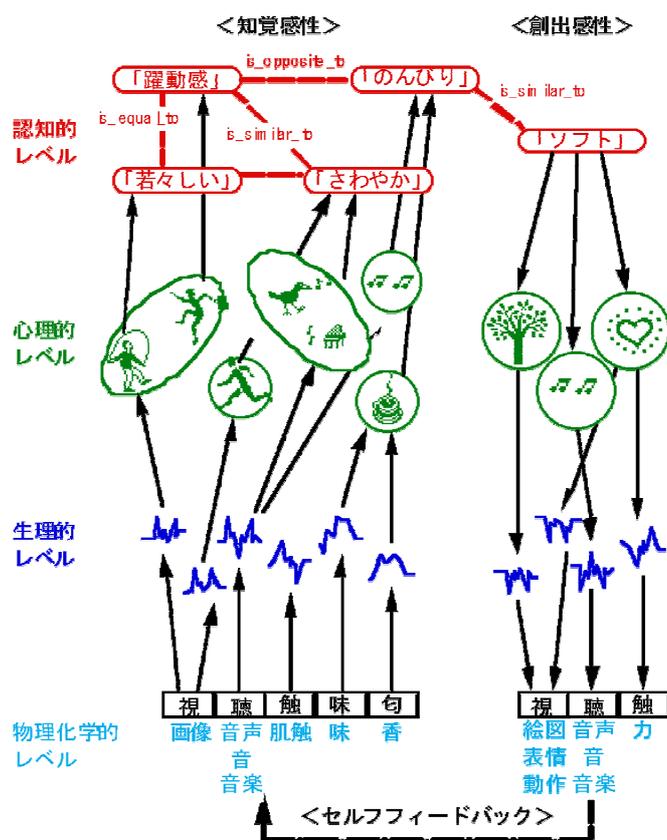


図 1 感性の工学的なモデルの基本構造

(1) 感性の静的な構造

(1-a) 感性のマルチメディア性

人間は、視覚・聴覚・触覚・味覚・嗅覚の五感のチャンネルをもつ。それぞれの感覚のチャンネルは、同時並列的に、その感覚に特有のセンサー（感覚器）で対象や外界からの刺激を受容し、それぞれの感覚器から感覚野にいたる過程で特徴を抽出・分析し、最終的には対象や外界の状況を総合的に認識し理解する。このように、個々の人間が感覚情報を受け取り、様々な特徴抽出機構を経て、これを解釈する過程（入力に相当）を知覚感性と呼ぶことにする。

一方、人間は、自身の発声や身振りや、あるいは、ペン・絵筆・楽器などの様々な道具を使用し、また、様々な素材（繊維・木材・石材や食材・香料など）を加工することにより、自分がイメージした様々なメッセージや事物を、他の人の五感を通じて伝えようとする。このように、個々の人間が頭の中でイメージする情報を、具体化して、実際に他の人間や計算機が五感を通じて知覚できるような情報に形作る過程（出力に相当）を創出感性と呼ぶことにする。

感性を工学的にモデル化する上では、それぞれの感覚のチャネルは、マルチメディア情報とそれぞれのメディアの処理過程としてモデル化できる。モデルには、複数の感覚に共通な処理の枠組みと、センシングの方法や信号の特性などそれぞれの感覚に固有な特徴抽出の仕組みとを含める必要がある。

### **(1-b) 感性の階層性**

次に、知覚感性や創出感性が、どのような段階を経て実現されているのかに着目して、感性の構造を整理してみよう。

知覚感性の過程では、個々の人間が感覚情報を受け取り、様々な特徴抽出機構を経て、これを解釈するまでの段階を経ると考えられる。

このような知覚感性の過程は、実世界に存在する対象から発せられる様々な信号を、感覚器が物理的・化学的な刺激として受容し、神経回路上の電気信号に変換するところからはじまる。これを物理レベルの感性と呼ぶことにする。これらの信号は、網膜などの感覚器から一次視覚野などの感覚野にいたる過程で、様々な特徴抽出がなされ、多次元ベクトルの的に数量化される。これを生理レベルの感性と呼ぶことにする。これに続いて、様々な対象やその状態を認識・理解するために、これら対象から抽出された種々の特徴に注目して、これらの信号を類似性に基づいて直感的に分類・類別化していると考えられる。これを心理レベルの感性と呼ぶことにする。ついで、これらの分類・類別化された対象やその状態と言語（イメージ語など）が対応付けられて、言葉による解釈が行われる。これを認知的なレベルの感性と呼ぶことにする。

創出感性の過程では、個々の人間が頭の中でイメージする情報を、具体化して、実際に他の人間や計算機が知覚できるような情報に形作るまでの段階を経ると考えられる。

このような創出感性の過程は、頭の中でイメージした言葉、あるいは、イメージそのもの（認知的レベル・心理的レベル）を、より具体的で直感的な表現・イメージ群に置き換えることから始まる。次いで、このような表現・イメージの中から、最も適切と判断される表現・イメージを選び出し、具体的な行動（発話・ジェスチャー・行動・造形など）として表出（生理的レベル・物理的レベル）する。

感性を工学的にモデル化するために、感性情報処理の過程で得られる（あるいは必要となる）種々の情報が感性の物理・生理・心理・認知の4つの階層のどのレベルの情報に相当するのかを整理する。同時に、それぞれの階層での情報処理や、特徴抽出や具体化などの階層間での情報の対応関係をモデル化する。

### **(1-c) 感性のマルチモーダル性**

次に、視覚・聴覚・触覚・味覚・嗅覚の五感の感覚情報を総合的に処理する過程に着目して整理してみよう。

人間は、多種類の感覚（視覚・聴覚・触覚・味覚・嗅覚）を同時に受容しながら、様々な様態（マルチモーダル）の情報を感知取り、また、相互の関係付けをしつつ、外界を知覚する。

ある対象から得られる視覚刺激から、様々な様態の情報を感知取ることができる。例えば、ある人物の顔画像から、個人の識別だけでなく、顔や視線の向き、表情や感情、顔色などの情報を得ることができる。一方、様々な感覚の情報からはほぼ同様の内容の情報を得られる場合もある。例えば、「さわやか」というイメージは、言葉や音声・音、図柄・映像、皮膚感覚や香りなど様々な感覚を通じて知覚することができる。

逆に、自分がイメージする情報を人に伝える場合、様々な表現手段（スケッチ・表情・音声・ジェスチャーなど）を選んだり、これらを同時に組み合わせて、イメージを伝えることも自然に行っている。

感性を工学的にモデル化する上では、感性情報のマルチモーダル性（感性情報のどのような性質がどのような情報を含んでいるか）や、多様なメディア間での相互の関係を分析することによりモデル化できる。

#### (2) 感性の主観性

同じ対象を見ていても、人によって異なった解釈・異なったイメージ語との結び付けを行う場合がある。また、頭に浮かぶ同じイメージを伝える際にも、人によって異なった伝え方をする場合がある。このように感性の知覚過程・創出過程には、個人ごとに特性の異なる主観性が認められる。

人間の感覚器の性能には大きな個人差はないものとみなすと、感性の知覚過程における主観性は、感性の階層的な構造の中では、感覚情報を受容し多次元ベクトル化する物理的レベルや生理的レベルではなく、類似性に基づいて分類・類別化する心理的レベルや認知的レベルで生じると考えられる。

感性を工学的にモデル化する上では、感覚情報を多次元ベクトル化する過程には個人差はないものとし、感覚情報を分類・類別化する際の規準・対応関係に個人差があるものとしてモデル化できる。また、このような個人差を検出するために、統計的な教示学習などの仕組みが必要となる。

実際に運用する際に、利用者（被験者）の心理的・身体的な負荷を軽減するためには、例えば、類似のプロファイルを持った人々の平均的な感性のモデルを利用し、これを順次、追加学習によりカスタマイズしていくなどの方策を考える必要がある。そのためには同時に、データベースに登録された多数のプロファイルの中からどのプロファイルを選択するかなどでのプロファイルの感性的な取捨選択の技術も必要となる。

### **(3) 感性の動的な構造**

#### **(3-a) 感性の相互作用性**

人間は、多種類の感覚（視覚・聴覚・触覚・味覚・嗅覚）のチャンネルを同時に働かせながら、それぞれの感覚情報相互の関係付けをしつつ、また、外界に能動的に働きかけたりしながら外界を知覚している。このような過程で、感覚のチャンネル間や知覚と創出の過程間に相互作用が認められる場合がある。

複数の感覚のチャンネルにまたがる現象として、視覚と聴覚に関しては、マガーグ効果 (McGurk Effect) が知られている。被験者に「ga」という口の動きの映像に重ねて「ba」という音を聞かせると、「da」という音を感じ取ってしまう。視覚情報によって聴覚情報の認識が影響を受ける場合がある。また、視覚と触覚に関しては、GUI 上でのポインタの動きに合わせてマウスや指先に振動を与えると、背景のテクスチャ・ウィンドウ境界・チェックボックスの別が容易に知覚できて操作性が向上したと考えられる報告もある。

感性を工学的にモデル化する上では、それぞれのチャンネルで感覚情報を並列に処理し、また、これらを統合する過程に、相互作用の仕組みを組み込む必要がある。

#### **(3-b) 感性の状況依存性**

感覚情報を知覚する過程においては、同じ情報を受け取っても、場合によって異なった応答や解釈を示す場合がある。

生理的なレベルでは、順応機構が知られている。明るい屋外に長時間いると、視細胞は明るい画像（強い光刺激）に順応し、明るい中での微妙な明るさの違いを知覚できるようになる（明順応）。この逆の順応現象を暗順応という。また、ある香りにしばらく接していると、その香りを感じなくなる。

心理的なレベルでは、あいまいな知覚の現象が知られている。若い女性の肖像画を数枚見た後と、老婆の肖像画を数枚見た後とでは、同じ肖像画（あいまいな画像）を見た時に、若い女性と知覚するか老婆と知覚するかが変わる。しかし、ある程度の時間が経過した後では、前に見た画像の影響は弱まっている。

認知的なレベルでは、状況が情報の価値判断に影響する場合がある。例えば、我々は、同じメッセージ・同じ音楽を聞いても、それが誰からの言葉・誰の演奏かの状況によって、主観的な解釈が左右されるのは、日常、よく経験することである。

感性を工学的にモデル化する上では、順序・短期的な時間の経過（数秒～数時間）を考慮に入れた情報処理過程としてモデル化する必要がある。

#### **(3-c) 感性の経時変化**

状況依存性も考慮に入れば、我々の知覚感性・創出感性は、短期的には一貫した規準に基づいて情報処理されていると考えられる。しかし、このような規準は、学習や経験の

蓄積に伴い、長期的には徐々に変化していくと考えられる。

例えば、絵画に対する視覚的な感性の場合、様々な絵画を多数鑑賞し、絵画の作法や歴史に関する知識も得ていくのに伴い、素人の解釈から専門家の解釈へと変化してゆく。

布地に対する視覚的・触覚的な感性の場合、例えばはじめは、布の色やテクスチャ、手触りの固さなどに対する比較的単純な解釈であっても、経験をつむことにより、布の光沢や、張り、ドレープなどの微妙な特性もとらえた解釈を行うようになる。

感性を工学的にモデル化する上では、長期的な時間の経過（数日～数年）を考慮に入れた情報処理過程としてモデル化する必要がある。そのためには、一度、統計的な教示学習で個々の被験者の感性をモデル化するだけでなく、長期的なスパンでの解釈の規準の変化などを検出し、統計的な学習に反映させるための仕組みが必要となる。

#### 3.7.4 多感覚感性データベースの技術的ニーズ

以下の節では、味覚センサ・嗅覚センサなどの開発により、近年急速にニーズの高まってきた多感覚感性データベースシステムとこれを利用した多感覚コミュニケーションシステムを開発する上での、技術的な課題を検討する。

感性の工学的モデル化の枠組みに関しては、マルチモーダル感性のモデル化に焦点を当て、複数の感覚チャンネル間での知覚の相互作用と階層的な類型化・解釈との対応付けの過程をモデル化する必要がある。このような実験・分析を高精度に行うため、仮想現実空間を利用した実験環境の構成・制御法を開発する必要がある。また複数の感覚チャンネルの具体例として、視覚・聴覚・嗅覚などを取り上げて詳しく分析すると共に、多感覚情報をコンテンツとした情報提供サービスを試作・評価することも必要となろう。

多感覚感性データベースシステムを実現するためには、次のような技術課題でのブレークスルーが必要となると考えられる。

##### (1) 階層的なマルチモーダル知覚過程の工学的なモデル化に関する研究

多感覚それぞれを別々の感覚のチャンネルとして、符号化することや通信することは可能となりつつあるが、五感（あるいは複数の感覚）が同時に現れる場合には、それらの感覚の間での相互作用（＝マルチモーダリティ）が人間の知覚に大きく影響する（例：視覚と聴覚におけるマザーグ効果）。

これら複数の感覚の間での相互作用の仕組みを工学的にモデル化すると共に、これらのモデルを利用して、データベースから多感覚の情報の内容に基づく連想・検索を行い、多感覚に訴える工業製品や都市空間の設計技術を開発する。

##### (2) 味や香り刺激の味覚・嗅覚情報データベースの開発とその応用に関する研究

味や香り刺激の味覚・嗅覚情報データベース技術開発のために、生理的な応答特性だけ

ではなく、利用者の解釈や主観的なイメージ語から味や香りを推定できる技術を開発する。特に、味や香り刺激の化学的特性、刺激を受けた人間の生理的な指標、人間の主観的な味や香りの解釈間の相関を分析・モデル化する。そして、味・香りと視覚・聴覚の相乗効果（マルチモーダル効果）の分析・モデル化を行う。

### （3）マルチモーダルな実験環境構築技術に関する研究

リアルタイムに、利用者への種々の情報提示を動的に行い、レスポンスにあわせて、次々に情報提示を変化させられる空間を構築する。システム側から利用者へ提示した情報に対して、利用者が示す感性的な振る舞いを動的に計測し、モデル化する。

### 3.7.5 海外の研究動向

米国では、2000年以降、ITT（IT21世紀）計画やITR&D計画において、「ヒューマンコンピュータインタフェースと情報管理」がこれからの情報通信技術の最重要課題の一つとされ、「人に優しいシステム」の研究開発が活発になっている。

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ISTプログラムが1998年以来進行中で、「市民情報サービス、マルチメディアコンテンツ、電子商取引、コンピュータシステム」が最重要分野となっている。さらに、英国においても、独自に、PACCITプロジェクトを進めるなど、当該分野の研究開発は非常に活発になっている。

味覚や嗅覚情報データベース技術に関する事例は見当たらない。自動化および実時間で処理できる定量化された嗅覚の測定および解析のためのシステム開発に関してはKobal(独)の嗅覚評価および感度測定技術が部分的に行われている程度である。HCIの分野では人間中心の味覚・嗅覚基盤のヒューマンインタフェースにおいて有効な応用例は少ない。

### 3.7.6 多感覚感性データベースの応用分野

味や香りの化学的特性と心理的な要因との関係を分析モデル化することにより、新たな味や香りの生成への応用が可能になる。さらに、味覚・嗅覚情報に基づいた味覚・嗅覚関連製品や人間中心の感性製品および居住環境の設計に応用できる。特に、適切な香りはリフレッシュ効果があることが知られており、例えば、VDT症候群およびストレス予防にも応用することが可能になる。特に子供によく見られるゲーム症候群を予防できると同時に学生の場合は集中力の向上、学習効果を高めてパソコン作業の際、能率高い環境作りに期待できると思われる。

また、人間の生理・心理的指標と車内環境との関係を分析・モデル化することにより、運転者にマッチした情報や環境を提供することが可能になり、自動車操作の快適性と安全性を向上することが可能になる。

## 3.8 感性のモデル化技術とEコマースへの応用

加藤 俊一 委員

### 3.8.1 はじめに

検索技術の向上により、膨大な商品群の中から自分が購入したいと考えている商品を検索することは、ある程度可能になった。しかし、「何かいいものがないかな?」というようにウィンドウショッピングするような感覚で、自分が興味を持ちそうな商品を探し出すのは、非常に大変である。しかし、「自分の興味の対象になりそうな商品の情報が欲しい」という思いは、消費生活の上では非常に基本的な要求である。このような要求に対して、消費者の主観的な商品の選択基準をモデル化することができれば、消費者ごとに興味を持つ可能性のある情報を適切にレコメンドすることができる。そうすることで、消費者は自分の選択基準にあったより多くの情報に触れることができる。

本稿では、電子社会時代に、市民生活に最も近いBtoC (Business to Consumer)型の電子商取引 (Eコマース) として、商品情報のレコメンデーションシステムを取り上げ、これを実現する上での技術的課題を考察する。

### 3.8.2 BtoC 型のEコマースの問題点と解決のアイデア

一人一人の消費者を対象としたBtoC (Business to Consumer)型のレコメンデーションシステム<sup>4)</sup>では、消費者本人のアクセス履歴や購入履歴、他の多数の消費者のアクセス履歴や購入履歴などの情報を用いて、消費者本人の興味を判断・予測し、商品のレコメンドを行っている。さらに、消費者がレコメンドを受けるためには、個人を特定できる情報 (名前、住所、E-mail アドレス) も登録しなければならない。

このようなシステムの構成と運用には、以下のような問題点がある。

(1) 同じ商品を見ても、消費者一人一人で、その商品のどの性質に注目してどのように評価するかの基準が異なるため、他人の商品に対する評価を利用して、消費者にレコメンドするには限界がある。

(2) まだ誰も買っていない商品や新商品は、その商品が消費者にどのように評価されるかわからないため、それを必要としている消費者にレコメンドされることはない。

(3) 嗜好情報や個人情報を、レコメンデーションシステムサービスを提供している会社に表示しないと、商品のレコメンドを受けることができないことに抵抗感を持つ消費者 (特に若い女性) も多い。

このような問題点を解決するためには、次のような技術的なブレークスルーが必要となる。

### (1) 消費者の商品に対する感性のモデル化技術

「商品のどの性質に注目して、どう評価するか」という消費者によって異なるプロセス（商品の観察から、主観的な評価を経て、購入に至るプロセス）をモデル化する。特に、個々の消費者が、特定の商品やある商品群のどのような特徴に注目し、それをどのような基準で解釈し、取捨選択（つまり購入するか否か）するかを、個々の観察事例・購入事例などからモデル化する必要がある。

このモデルには、商品の素材・形状・質感（木質・布、丸い・角ばった、半透明、光沢など）などの商品の客観的な特徴を表現する客観的イメージ語、消費者が主観的に商品の特徴を表現する主観的イメージ語（かわいい・シックなど）や、商品の様式を表現する知識的イメージ語（ブリティッシュトラッド・アールヌーボー・ポストモダンなど）の相互の関係の強さを表現した「イメージ語ネットワーク」が含まれる。同時に、ログからとれる Web アクセス履歴、商品購入履歴から、消費者が興味を持つ商品のカテゴリ・商品イメージなども含める。

このような感性モデルを利用することで、商品情報提供者は消費者の評価基準に合った商品のレコメンドが可能になり、消費者は自分の判断基準にあった商品の情報を手に入れることができる。

### (2) 新しい商品のレコメンド

上述のイメージ語ネットワークがあれば、新商品がどのような特徴を持つかを分析することで、個々の消費者が新商品についてどう思うかを推定できる。従って、新商品であっても、これに関心を持つ（潜在的な）消費者にレコメンドすることができる。

### (3) 感性モデルの匿名化

消費者の立場からは、商品のレコメンドは、個人のプライバシーを守りつつ、適確な情報だけを受け取りたい。一方、売り手の立場からは、購入する可能性の高い（潜在的な）消費者に選択的に情報を送りたい。両者のニーズは、消費者のプライバシーを開示することなく満足させることができる。

そのために、個人の特特定を防ぐために感性モデルを匿名化し、これをマッチメイキング・レコメンドサービス会社のサーバー上で匿名の私書箱的に公開する。この匿名感性モデルを用いれば消費者は匿名性を保ちながら、興味を持ちそうな情報を私書箱宛のメールとしてより多く得ることができるようになる。

## 3.8.3 感性的レコメンドシステムのシステム構成

以上のような感性モデルとこれを利用したレコメンドシステムの構成は、図1のようになる。

レコメンドシステムは、消費者の感性モデルに基づいて消費者の商品選択の基

準に合う商品を既存の商品データベースや新商品リストから検索し、 Recommend する。本システムは以下の手順からなる（図1）。

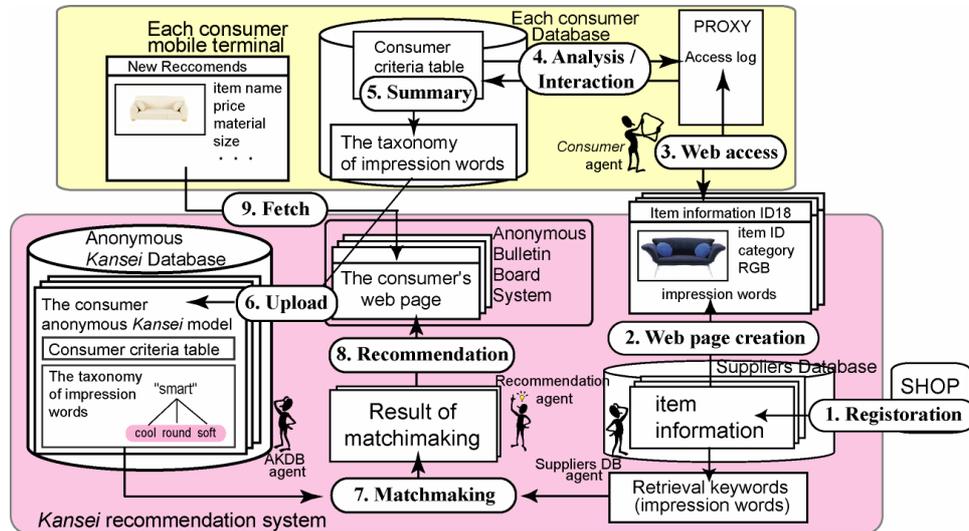


図1 匿名化された感性モデルを利用した商品のレコメンデーションシステムの基本構成

- (1) 商品情報提供者は新商品に関する基本的な情報（商品名、カテゴリ、画像、色、素材）と共にその商品の雰囲気や言葉を登録する。
- (2) 消費者の感性モデルや商品情報提供者の登録する情報は更新などにより、時間と共に変化する。そうした時系列データを管理するために、各データベースにはエージェント機能が組み込まれている。商品データベースエージェントは商品情報提供者が登録した商品の情報をXMLで記述してWebページを自動的に生成する。
- (3) 消費者は興味を惹くような商品を探す際に、モバイル端末内に立てられたプロキシサーバを経由して商品ページを閲覧する。このプロキシサーバでURL、閲覧頻度、閲覧時間、イメージ語などのログを取得する。
- (4) 一人一人のモバイル端末に組み込まれている消費者エージェントはWebアクセスで得られたログをチェック・解析し、インタラクションテーブルに蓄積する。消費者エージェントはインタラクションテーブルを参照して、一人一人の消費者ごとに異なる商品に対する評価基準を「スマートな」「クールな」といったイメージを表す言葉を用いて階層化し、「個人用イメージ語ネット」をインタラクションを通して作る。
- (5) 消費者エージェントはインタラクションテーブル中のイメージ語間の関係を整理してイメージ語ネットを作成する。このイメージ語ネットは商品の雰囲気や主観的イメージ語と物理的特徴を表す客観的イメージ語からなる。同様に、商品情報提供者のイメージ語ネットも商品データベースエージェントが作成する。

- (6) 消費者イメージ語ネットを匿名感性データベースへ格納する。
- (7) レコメンデーションエージェントは、匿名感性データベースにある消費者のイメージ語ネットと商品データベースにある商品情報提供者のイメージ語ネットをマッチメイキングして一人一人の消費者にあわせた商品情報を検索する。
- (8) 個々の消費者はレコメンド結果を入れてもらうために、持ち主が誰かは匿名化されている掲示板を持つ。レコメンドの結果はこの匿名掲示板に格納される。
- (9) 消費者エージェントは、消費者ごとにある専用の匿名掲示板をチェックして、レコメンドされた情報を取得し、消費者の感性モデルに合わせて表示順序を編集し、レコメンド画面を表示する。消費者はそれを見ることで自分の判断基準と同じ基準によりレコメンドされた商品の情報を得ることができる。

以上の手順により、感性レコメンデーションシステムは、消費者の感性モデルに基づいて選択の基準に合う性質の商品をレコメンドする。

### 3.8.4 感性的レコメンデーションシステムの技術課題

感性のモデル化技術とこれを利用した感性情報通信技術の研究開発を核に、感性的な情報提供サービスの具体例としてショッピング支援システムの試作を通じて、基盤技術の整備を進める必要がある。

以下のような技術課題があげられる。

#### (1) 利用者の心理的・身体的負担の少ない感性のモデル化技術の研究開発

従来の感性のモデル化技術では、対象とするマルチメディアコンテンツごとに、かなりの数の教示用データを利用者から得て、統計的学習を行っていた。従って、利用者（被験者）の心理的・身体的負担は大きく、感性のモデル化を進める上でのボトルネックとなっていた。

このような研究開発では、様々なカテゴリの商品それぞれを対象に、感性の階層的なモデル化（物理→生理→心理→認知）とイメージ語ネットの統合化・高度化を図る。特に、多種多様なカテゴリの商品に共通な概念・固有な概念（イメージ語表現）に対応して、新しいカテゴリの商品に対しても、既存の感性のモデルを利用して、出来るだけ少ない教示用データから、感性のモデルを構築できるようにする。

商品の素材・形状・質感など、それぞれの属性と主観的なイメージ語との関係を分析する必要がある。さらに、これらの属性が複合された状態で消費者は商品を知覚しているので、視覚と聴覚、視覚と嗅覚のように、複数の感覚・複数の属性が複合された状態で知覚する感性（マルチモーダルな感性）のモデルの構築技術の研究開発も必要となる。

## (2) マルチメディア情報の感性的な翻訳技術の研究開発

商品の作り手・売り手と消費者の感性が異なる場合、作り手・売り手の意図を正確に消費者に伝えることは難しい。両者の感性モデルを踏まえて、適切に翻訳する仕組みが必要となる。

様々なカテゴリの商品の物理的・知識的な特徴と、言葉で表現される主観的な解釈との対応関係に注目して、

- ・ ある商品と同様の解釈をすると推定される異種のカテゴリの商品の選択(すなわち、あるカテゴリの商品から別のカテゴリの商品を連想する。)
- ・ 作り手・売り手の考えるイメージ語表現に対して、それと結びつく商品群を作り手・売り手の感性モデルから検索し、次いで、個々の消費者の感性モデルに照らして、それらの商品群に対して消費者がするであろう解釈をイメージ語表現の形で生成する(すなわち、主観的イメージから主観的イメージへの翻訳)

などの技術の研究開発も必要である。

本研究開発の後期では、翻訳技術の範囲をマルチモーダルな感性に拡張する。

## (3) 感性的な相互理解の研究開発

ヒューマンコミュニケーションにおいて、「相手を理解する」のは、次に示す2つの方式に類型化できると考えられる。どちらの方式を採用するかは、当事者である人間の判断に任される。

1. 独立感性モデルの構築： 送り手の意見を理解してもらうためには、送り手の感性モデルを受け手側に再現する必要がある。送り手の感性モデルの特徴が現れる事例とその主観的な解釈を選び出し、順次、受け手側に伝える。受け手側では、一連の情報に基づいて、送り手の感性モデルを統計的学習などにより構築する。送り手の感性モデルが正確に再現されているかを判断(評価)するためには、同様に、送り手の感性モデルの特徴が現れる事例を受け手に伝え、受け手側で推定した解釈との一致度を評価する。この方式では、受け手は、自身の感性とは別に、知識として送り手の感性をモデル化するので、受け手の感性モデルそのものは変化しない。
2. 従属感性モデルの構築： 送り手の意見を元に受け手が学習し直す場合、送り手と受け手間の感性モデルの相互作用が重要になる。具体的な事例に基づいて、受け手が再学習することにより、受け手の感性モデルが変化する。特に、受け手の考えが明確に定まっていない場合には、インタラクションの過程を通じて、受け手がイメージする商品の仕様や特徴などがはっきりし、当初と違う考えに至ることがある。このような思考の明確化を **concept articulation** と呼び、そのアルゴリズムを開発する。

ここで、消費者がショッピングで「面白い・楽しい」と感じて満足感を得るのは、単に、品質の良く自分の嗜好にマッチした商品を安価に購入するだけでなく、ショッピングの

過程を通じて、自身の知識が増え、より広い視点・専門的な視点から、商品を吟味できるように消費者自身が成長すること（「目が肥える」）を実感するときであることを指摘しておきたい。いわゆるリピーターとして消費者を確保してビジネスを発展させていくためには、単に商品の情報を伝えるだけのサービスではなく、消費者がそこに面白さ・楽しさを実感できるようなサービスとして、システムを構築する必要があると考える。

#### **(4) 感性モデルの匿名化と共有技術の研究開発**

感性モデルを高度に利用するために、利用者の感性モデルを完全に秘匿する部分と、匿名で公開・共有する部分に分けて管理し、プライバシーを守りながら適切にネットワーク上の情報サービスを享受できるようにする技術を開発する。

高度情報化社会は、インターネット、モバイルネット、ユビキタスネットが重畳した、複合的な情報通信環境を実現しつつある。このような複合的な情報通信環境も前提として、個人情報 の 携 行 や 開 示 ・ 保 護 の た め の 技 術 を 開 発 す る 必 要 が あ る べ し。

#### **3.8.5 期待される波及効果**

本研究開発により、モバイル環境での新しい電子商取引市場の生成、自動車における状況に応じた快適な情報環境提供サービス、製品設計の意思決定時間の短縮などが期待できる。モバイルインタフェースエージェントや感性的なプライバシーの保護技術は、携帯電話を利用したサービスや感性工学を利用したシステム設計やサービスを得意とする我が国をはじめ、世界的にも大きな市場があると期待できる。

本研究の成果を利用することにより、例えば、産業面では、消費者のプライバシーを高度に保護しながら、ある感性的なプロファイルを持つ消費者層の嗜好などのマーケット情報を産業界に提供することにより、よりの確できめの細かい情報サービスと、これを利用した商業活動が活発になると期待できる。

教育面では、学童生徒一人一人のプライバシーを保護しながら、適切な教材の提供や学習指導が行えるようになり、自主的な学習や学校教育の質の向上が期待できる。

個人の知的な資質・技能的な資質を、個人レベルから社会に向けて発信することも容易になり、高齢者の社会参加・社会貢献に道を開くものと期待できる。

#### **[参考文献]**

- [1] T.Ishida (ed.) "Community Computing: Collaboration over Global Information Networks", John Wiley & Sons, 1998
- [2] 矢野, 北野, 末吉, 篠原, ピンヤポン, 加藤, "消費者の感性モデルを利用したレコメンドーションシステムの構築", 情報処理学会, DBWeb2002, pp.283-289, 2002

## 3.9 テレイマージョンとビジュアルデータマイニング

宮田 一乗 委員

### 3.9.1 はじめに

CG 技術の進歩とネットワーク速度の高速化にともない、地理的に離れた空間をコンピュータとネットワークで構成される仮想空間内で統合する、テレイマージョン (Tele-Immersion) の研究が現実味を帯びてきた。一方で、大量の画像データの中から意味のあるデータを発掘するビジュアルデータマイニングの研究は、バイオサイエンスや気象学などにおける多元解析のツールとして、今後重要度が増すと考えられる。本報告では、これらの技術動向に加えて、テレイマージョンの環境下で、協調しながらビジュアルデータマイニングする手法なども紹介する。

### 3.9.2 テレビ電話とアウェアネス

遠隔地にいる人たちがリアルタイムで映像と音声でコミュニケーションをとるインフラとしては、テレビ電話の歴史が古い。コンピュータの普及により、テレビ電話はその形態を変えて、ISDN 回線を用いたテレビ会議システムや、PC と CCD カメラによるインターネットを介した動画チャット、さらには小型 CCD カメラ搭載の移動通信端末によるモバイルコミュニケーションへとダウンサイジング、およびカジュアル化してきている。

これらのテレビ電話 (とその発展形) のインフラは、どちらかという音声によるコミュニケーションが主で、映像はコミュニケーションの補助として用いられることが多い。ただし、映像は多くを語るのも、情報の付加価値は大きく、伝わる感情 (好き嫌いの感情など) のダイナミックレンジは拡大される。その一方で、会話に付随するノンバーバルな情報は、うまく伝えることが出来ない。そのひとつの例として、視線の不一致が挙げられる。すなわち、カメラとディスプレイを同一の場所に設置することが困難であるために互いの視線が微妙にずれることになり、正確なアイコンタクトを取ることが出来ず、意思の疎通が難しいものになっている。また、空間的な位置関係を感じ取ることが困難であるため、アウェアネスを共有できないことも問題点として挙げられる。例えば、会議に積極的に参加しているのは誰か、審議事項に難色を示しているのは誰かといった情報は、空間を共有していないとなかなか伝わらない事柄である。

テレイマージョン (Tele-immersion) とは、地理的に離れた空間にいる人々が、あたかも同室にいるかのように自然に対話できるように、コンピュータとネットワークで構成される仮想空間を提供する仕組みである。テレビ電話の枠組みと大きく異なるのは空間を共有するという点にあり、これを可能にしたのが、ネットワークとメディア技術である。な

お、テレ-immージョンは、米国における次世代ネットワークのコンソーシアムであるインターネット 2 (<http://www.internet2.edu/>)での重要なアプリケーションのひとつとして考えられている。

### 3.9.3 テレ-immージョンとメディア技術

テレ-immージョンを支えるメディア技術の主たるものとしては、バーチャルリアリティ (VR) 技術とコンピュータビジョンが挙げられる。

#### 3.9.3.1 VR 技術

VR 技術は、CG の父と呼ばれる I. Sutherland が 1968 年に試作したヘッドマウントディスプレイ (HMD) を発端としており、臨場感のある視覚情報の表示法や、触覚や力覚のフィードバックの手法など、人間の五感へ作用するデバイスの開発を中心に研究がなされている。

体験者を取り囲む多面投影による没入型ディスプレイである CAVE が 1991 年にイリノイ大学にて開発された。これは、体験者の周囲に 3m x 3m の平面スクリーンを正面、左右面、床面の 4 面に配置し、ステレオ液晶プロジェクタでそれらの投影面に映像を表示する広視野角の表示システムである。日本でも、これを拡張した 5 面投影の CABIN (CAVE に加えて天井に投影)、6 面投影の COSMOS が東大で開発された。HMD や CAVE は、もともとスタンドアロンの VR 空間の表示用デバイスとして利用されてきたが、これらを高速ネットワークで結合し、1つの共有空間を実現しようというのがテレ-immージョンの大きな目標である。

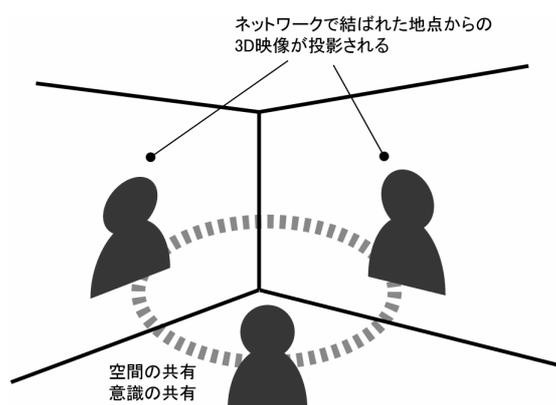


図1 テレ-immージョンの概念図  
高速ネットワークで結合された仮想空間により、空間の共有、および意識の共有を行い、遠隔地間の自然なコミュニティを実現する。

人間の両眼での水平視野角は 240°ほどであるが、この視野をどれだけ映像でカバーできるかが臨場感の要因となる。HMD や CAVE などは、体験者の眼や身体を物理的に映像空間で取り囲むことで、視野角の広い映像を表示しているが、一方で、大型のスクリ

ーンを前面に設置し、視野角の広い映像を表示する試みもなされている。

ボストン大学の Scientific Computing and Visualization グループでは、Deep Vision Display Wall (DVDW : <http://scv.bu.edu/Wall/>) と呼ぶ、複数の PC と液晶プロジェクタによる大型高精細ディスプレイを開発している。図 2、3 で用いられている液晶プロジェクタの解像度は 1024x768 であり、2x2 枚の投影面を持つ場合、DVDW 全体としては 2048x1536 の解像度を持つ。DVDW の 1 つの特長に、スケーラブルな解像度の向上があり、現在では 4x3 枚の投影面で 4096x2304 の解像度を持ち、4.6m x 2.4m の大きさのスクリーンが開発中である。なお、各プロジェクタは 1 台の LINUX PC でドライブされている。ステレオ画像生成の生成には、従来のような高価なグラフィックワークステーションを持ちいておらず、LINUX PC のクラスタを用いることで廉価なシステムを実現していることも特筆すべきことである。この場合、PC 間の描画のシンクロを取らないとバラバラな画像になってしまう。DVDW では、FrameLock と呼ばれるソフトウェアによるシンクロを取っており、また、PC クラスタ内の分散レンダリングのプラットフォームには、WireGL を用いている。



図 2 Deep Vision Display Wall  
複数の液晶プロジェクタ(2x2)により、大型で高精細な画像を表示できる。  
(画面はサメの遊泳の CG アニメーション)



図 3 Deep Vision Display Wall の映像  
投影部  
2 台の液晶プロジェクタが対になり、Deep Vision Display Wall のスクリーン上における 1 ブロックの映像を投影する。

### 3.9.3.2 コンピュータビジョン

カーネギー・メロン大学 (Carnegie Mellon University) の金出武雄教授の研究グループでは、1993 年以来、Virtualized Reality と呼ばれるコンピュータビジョンの技術を開発している ([http://www.ri.cmu.edu/labs/lab\\_62.html](http://www.ri.cmu.edu/labs/lab_62.html))。この技術では、スポーツのような時系列で変化する動きのある場面を同時に複数台のカメラで取り囲むように撮影し、撮影されたデジタル映像に「見え方」の情報を付加して、時間的および空間的に自由な位置からの再生を可能にした。これは、複数のカメラで撮影された画像から、三角測量の原理で3次元データを求め、それらのデータをマージして、撮影されたシーンの3次元モデルを復元するものである。データのマージングの際には、3次元空間をボクセル (Voxel: サイコロのような小立方体) に分割し、そのうちのどこに表面があるか (どこが見えているか) を求めている。この技術は、アメリカンフットボールの試合に適用され、タッチダウンの瞬間の多視点からのシーン再現などを可能にした[1]。



図4 Virtualized Reality の3D Room の壁面

4面の側面の壁にそれぞれ10台、天井に9台、合計49台の CCD カメラが、図中の丸い穴を通して室内に向けられて設置されている。3D Room の大きさは、20 feet x 20 feet x 9 feet ある。



図5 Virtualized Reality の3D Room の裏側

壁の裏側に CCD カメラが設置されている。CCD カメラは PC に接続され、リアルタイムで動画を取り込める。また、音も同時に取り込めるようになっている。

図4、5は、3D Room と呼ばれる CMU 内の施設で、部屋の4面の側面の壁にそれぞれ10台、天井に9台、合計49台の CCD カメラが、図中の丸い穴を通して室内に向けられて設置されている。この施設内で、Virtualized Reality の研究が進められており、ダンスのシーンやバスケットボールを楽しむシーンなどを取り込んで実験を進めている。

ノースカロライナ大学チャペルヒル校 (The University of North Carolina at Chapel Hill) のフックス(Henry Fuchs)教授の研究グループは、多視点カメラと特殊な光 (Imperceptible Structured Light: ISL) の組み合わせで、物体表面の3次元の凹凸形状を計測するシステムを開発した[2]。このシステムで用いられている ISL とは、人間の眼に知覚出来ない速さで明滅する光 (高速に明滅するので連続した光に見える) である。

3次元形状を計測するには、複数のカメラで撮影された画像間の対応点を求め、カメラからの距離を求める必要がある。物体表面になんらかの目印 (例えばチェッカーパターンや、ゼブラパターンなど) が付いていると、対応点が比較的簡単に求められる。フックスらは、この目印となるパターンを対象物体に ISL で投影し、リアルタイムで距離画像を求めた。ISL は人間の眼に知覚出来ない光の明滅なので、視覚的に非浸襲である。このシステムを用いて部屋の中を測定することで、壁面の凹凸をリアルタイムで計測し、柱やパイプなどの出っ張りがある壁に対しても、歪みのない映像を投影できるようになった。すなわち、部屋の壁全体を大きな平面ディスプレイとして扱うことが出来るのである。

#### 3.9.3.3 イメージベースドレンダリング

実世界をコンピュータ内に取り込む技術であるイメージベースドレンダリング (Image Based Rendering : IBR) やイメージベースドライティング (Image Based Lighting : IBL) の技術を用いて、仮想空間を構築する研究が進められている。

IBR および IBL の研究成果は、Paul Debevec (<http://www.debevec.org>)を中心にした研究グループのものが有名である。Debevec らは、多数の写真を元に撮影時の照明条件を仮想空間内に再構築して、CG 物体のリアルな合成を試みた。また、2002年の SIGGRAPH においては、仮想空間での照明条件をコンピュータ制御の LED 光源で再現して実世界での照明に用いる、LightStage3 と呼ばれるシステムを発表した[3]。このシステムにより、現状のクロマキー合成にとって変わる、新たな実世界と仮想世界とのシームレスな統合が可能になると考えられる。これらの研究も、テレイマージョンを支える技術となる可能性を秘めている。

IBR, IBL では、撮影された風景を光源と仮定して、仮想空間内の CG モデルの照明を行うが、そのためにはダイナミックレンジの広い画像フォーマットが必要である。そこで、HDR image (High-Dynamic Range image) と呼ばれる、各ピクセル値を小数値で表す画

像フォーマットが提案されている。

IBL の手法は、最新の RenderMan (CG 映像を生成するためのソフトウェア: RenderMan Release11, <https://renderman.pixar.com/>) に採用されており、また、一部のグラフィックカード上では、リアルタイムに IBR の画像を表現できる。

### 3.9.4 ビジュアリゼーションとビジュアルデータマイニング

“seeing is believing” - 「百聞は一見にしかず」。人の視覚情報による理解力を示すことわざである。コンピュータを用いて大量のデータを画像に変換し(ビジュアリゼーション)、視覚化されたデータから意味のある情報を探り出す(ビジュアルデータマイニング)技術は、多次元情報からの確に必要情報を抽出し、人間の判断を補助するものとして重要な技術であるとする。

#### 3.9.4.1 コンピュータによるデータの視覚化

コンピュータによる大量データの高速処理は可能であるが、人間がそのような膨大なデータを瞬時に的確に理解することは困難である。一方、画像化されたデータに対する人間の理解力は、きわめて高い。例えば、天気図から明日の天気を予想したり、CT 画像から病巣を見つけ出したりすることができる。

コンピュータを用いて、複雑な現象や大量のデータ、目に見えないデータを人間に理解しやすい形に視覚化する技術が、コンピュータ・ビジュアリゼーション(**computer visualization**)である。コンピュータ・ビジュアリゼーションは、その目的に応じて、以下のような分類ができる[4]。

- サイエンティフィック・ビジュアリゼーション:  
気象情報や流体解析の結果などの複雑な数値データの可視化など
- エンジニアリング・ビジュアリゼーション:  
振動解析や応力解析のような、設計に関わる測定・解析データの可視化など
- メディカル・ビジュアリゼーション:  
CT 画像の表示や器官の視覚化のような医療診断のための情報の可視化など
- ビジネス・ビジュアリゼーション:  
プレゼンテーション支援を主とした統計結果などの可視化など
- エデュケーショナル・ビジュアリゼーション:  
わかりやすい教材の提供のための可視化など

コンピュータビジュアリゼーションは、対象の時空間のスケールに依存せずに、好きな見方で、対象を壊さずに、高い再現性で繰り返し探ることが可能である。したがって、理想的には、コンピュータ上で理論と実験を対置することが可能になる。

## 3.9.4.2 ビジュアルデータマイニング

単なるデータの可視化ではなく、人の見やすい画像や動画にデータを変換したり、変換後の画像データから再度情報を抽出するとともに、大量のデータを再度人間に理解しやすい情報に変換する、一步進んだビジュアリゼーションが必要である。すなわち、人間とデータ、可視化画像間におけるインタラクションが必要とされる。このインタラクションがビジュアルデータマイニングである。

図6、7にビジュアルデータマイニングの例を挙げる。これらの例は、ウェブサイトへのアクセス数や更新日時を視覚的に表したものである。このような膨大なデータも視覚化することで、ビジュアルに大局を把握することが可能になり、また、全体での特殊なケースも浮き彫りになってくる。

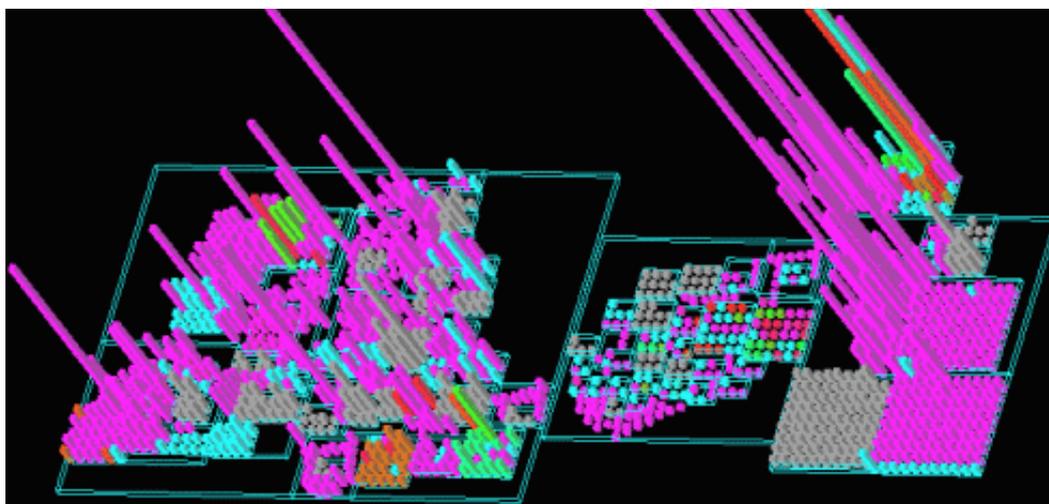


図6 ウェブサイトの全体像を表現した例  
(日本アイビーエム(株) 東京基礎研究所 伊藤貴之氏 提供)

個々の棒グラフが個々のウェブページに相当する。アクセス数を棒グラフの高さで表現し、更新日時で棒グラフが色分けされている。



図7 ウェブサイトへのアクセス分布の表現例  
(日本アイビーエム(株) 東京基礎研究所 伊藤貴之氏 提供)

灰色のドットは、該当する時間帯にアクセスがなかったウェブページを表している。赤色の棒は、該当する時間帯にアクセスがあったウェブページを表しており、その高さがアクセス数を表現している。灰色や赤色のドットを囲む長方形の境界線は、ディレクトリ階層を表わす。

例えば、図7の右上にある緑丸は、あるディレクトリ内部の100個以上のウェブページが、ある一定時間内（1時間以内）にすべてアクセスされたことを示している。その後、該当するウェブページ群は、すべてオープンソースのAPIマニュアルのページであることが判明した。このような熱心な潜在的閲覧者によるアクセスは、従来のアクセス分析ツールによる棒グラフ表示やランキング表示で発見することは困難であるが、視覚的にデータマイニングすることで容易に発見できる。

(<http://www.trl.ibm.com/projects/webvis/index.htm>)

### 3.9.5 遠隔コラボレーション

以上で述べたテレマージョン技術とビジュアルデータマイニング技術を有機的に結合することで、有効な遠隔コラボレーションが実現可能になる。すなわち、地理的な制約条件を取り払い、コンピュータの補助による知的で効率的な協同作業の環境が提供できるのである。

遠隔コラボレーションに関しては、既にいくつかの実証実験が進められている。例えば、GMでは、車のデザインにあたり、Detroit, Brazil, Germany (Opel) の3拠点を結んで、バーチャルプロトタイピングを行っている。

遠隔コラボレーションを発展させることで、新たな遠隔教育のプラットフォームを提供できるかもしれない。例えば、GeoWall Consortium ([www.geowall.org](http://www.geowall.org))では、VR技術を教育分野に応用し、地球科学の体験的学習環境の実現を目指している。この環境をネットワークで結合すれば、地球環境に対して全世界で考える「学びの場」を提供できるであろう。

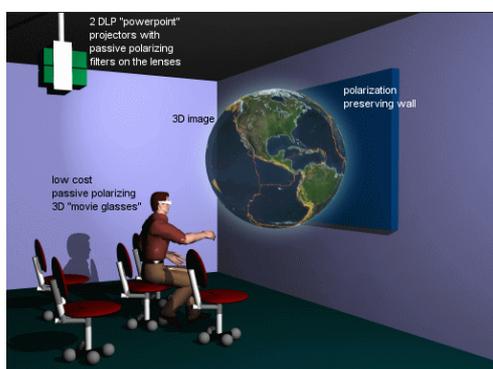


図8 GeoWallの概念図（Webより転載）  
安価なVR表示技術により、データの可視化技術を教育の場に導入している。

### 3.9 テレイマージョンとビジュアルデータマイニング

平成 14 年度より文部科学省が新たに開始した、情報通信分野における「IT プログラム」<sup>1</sup>の委託事業のうち「e サイエンス」実現プロジェクト<sup>2</sup>の実施テーマの 1 つである「スーパーコンピュータネットワーク上でのリアル実験環境の実現」の研究プロジェクトが進行中である。この研究開発は、北陸先端科学技術大学院大学、富士通のほか、京都大学、広島大学、日本原子力研究所、金沢医科大学の計 6 機関による産学官共同プロジェクトとして実施されている。

このプロジェクトでは、時間・空間・組織・分野の違いを超えた新たな研究開発環境として、疑似体験共有によるコラボレーションや遠隔共同実験などを実現するスーパーコンピュータネットワーク上でのリアル実験環境の構築を目指している。

実応用分野における適用例としては、分野横断的な循環器疾患診断治療支援が挙げられる。ここでは、動脈瘤を対象とし、CT/MRI 画像を元にした血管（構造計算）と血流（流体計算）の連成解析とその統計データ解析により、診断や治療の支援を行う。例えば、患部の実画像に高速ボリュームレンダリングによる高精細 CG 画像を重ねあわせる（リアルタイムテクスチャ）ことにより、外科手術をアシストしたり医療教育に利用する仕組みを提供したり、診断データ及び解析結果をデータベース化し、データマイニング手法を用いた診断の IT 化を実現することを想定している。

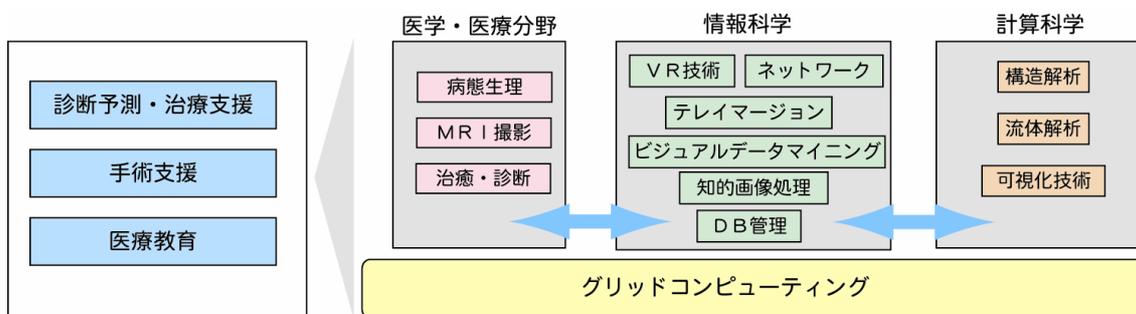


図 9 分野横断的な循環器疾患診断治療支援  
グリッドコンピューティングのインフラとテレマージョン技術の融合による、新たな医療環境の提供を目指す

<sup>1</sup> IT プログラム

我が国が取り組むべき情報通信分野の国家的な研究開発課題について、産学官の最も能力の高い研究機関を結集し、総合力を発揮できる体制により取り組むことを目的として、平成 14 年度から文部科学省が新たに開始した委託事業。

<sup>2</sup> 「e サイエンス」実現プロジェクト

研究開発現場に高速研究情報ネットワーク等の高機能 IT を活用することにより、研究開発スタイルを革新し、新たな研究分野（融合研究領域等）を創出する「e サイエンス」の実現に向けた、研究情報基盤技術の開発・整備・実証等を行うプロジェクト。

### 3.9.6 おわりに

テレマージョン技術の発展により、地理的にバリアフリーな環境が実現できるだろう。それはすなわち、物流以外の物理的な交通量（人の移動量）の削減につながり、安全性や、環境問題、地域格差などの問題を解決してくれる手段のひとつになる可能性がある。また、ビジュアルデータマイニングは、可視化技術にさまざまな「知識」を付加することにより、人間のソフト面でのバリアフリーを実現してくれるかもしれない。画像や映像は言語と違って、世界共通である。テレマージョン技術とビジュアルデータマイニング技術の有機的な結合により、次世代のビジュアルコミュニケーション環境の実現が期待される。

#### [参考文献]

- [1] <http://www.ri.cmu.edu/events/sb35/tksuperbowl.html>
- [2] R.Raskar, et.al., “The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays,” Computer Graphics Proceedings(SIGGRAPH 98), pp.179-188
- [3] P. Debevec, et. al, “A Lighting Reproduction Approach to Live-Action Compositing,” ACM TOG (SIGGRAPH2002), Vol.21, No.3, pp.547-556
- [4] 中嶋、藤代編著, “コンピュータビジュアリゼーション”, 共立出版

## 3.10 昨今の音楽情報処理における研究プロジェクトについて

平田 圭二 委員

### 3.10.1 はじめに

本節では、昨今の新しい音楽情報処理研究の流れを象徴するような国際的研究プロジェクトを 5 つ紹介する。

昨年の報告書では、新しい音楽情報処理研究の特徴は、(a) 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、(b) 応用システムをインターネット/Web 上に展開すること、(c) 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の 3 点であると述べた。これらを念頭に置いて、今年 3 つの音楽情報処理関連の国際会議に参加した (International Conference on Music and AI (9 月, Edinburgh)、International Computer Music Conference (9 月, Gothenberg)、International Conference on Music Information Retrieval (10 月, Paris))。得た所感は以下の通り。

(a) の試みは地道に続けられている。論文はある程度の件数発表されているものの、決定的なブレークスルーには至ってないようだ。現在、確立しつつあるサブ研究分野としては、メロディ分割、メロディ類似度、楽曲構造分析、声部分離、演奏表情分析/合成、音楽データベースのインデキシングなどがある。

(b) に関しては実際に多数の事例を見出すことができたが、アプリの動作環境としてモバイル端末を想定したものも発表されていた。例えば、ユーザの嗜好を反映した **playlist** (ラジオ局やレコード店の推薦盤/曲リストのこと) の自動作成が目をつけた。

(c) に関しては対象タスクが中粒度に変化しただけではないように感じた。音楽情報処理におけるタスクは大きく生成系 (制作系) と認識系に分類できるが、現在注目されているのは中粒度でかつ認識系のタスクである。信号レベルのパターン認識技術に限定し、(a) のような意味レベルまで踏みこまなくとも、有用な商用アプリケーションが組めるのではないかという実績と期待の反映である (これは、昨今の統計的自然言語処理の隆盛を連想させる)。また、生成系 (制作系) に敢えて取り組まないことで、他分野の研究者の参入が容易になり、一般の人でも利用可能な音楽システムの研究開発であることがより明白にアピールできるため、研究リソースが集まってきている。

現在欧米では、上述した音楽情報処理の新しい研究の流れを象徴するような研究プロジェクトが、NSF、FP5 (FP6) 等から支援を受けて進行中である/あった。本報告では、その中から OMRAS、CUIDADO、WEDELMUSIC、MUSICNETWORKS、SALIERI/GUIDO を取り上げ紹介する。

なお、FP5<sup>1</sup>、FP6 関連のプロジェクト (CUIDADO, WEDELMUSIC, MUSICNETWORK) の詳細は <http://www.cordis.lu/ist/projects/projects.htm> のページより調査できる。

### 3.10.2 OMRAS (Online Music Recognition and Searching)

表 1 OMRAS プロジェクトの概要

研究費	米国 NSF の Digital Library Initiative Phase 2 と英国 Higher Education Funding Councils の Joint Information Systems Committee (JISC) が共同で運営する International Digital Libraries programme より研究費助成を受ける
研究組織	米側: Massachusetts Univ. Center for Intelligent Information Retrieval (CIIR), Indiana Univ. Johns Hopkins Univ.、英側: King's College London (KCL), Oxford Univ., Queen Mary College, London
研究代表者	米側 Bruce Croft (CIIR), Donald Byrd (Indiana University)、英側 Tim Crawford (元 KCL、現 City University, London), Matthew Dovey (元 KCL、現 Oxford University)
プロジェクト期間	1999 年 6 月からの 3 年間
参考資料	<a href="http://omras.dcs.kcl.ac.uk/">http://omras.dcs.kcl.ac.uk/</a> <a href="http://www.elec.qmul.ac.uk/research/projects/nsf_9905842_omras.html">http://www.elec.qmul.ac.uk/research/projects/nsf_9905842_omras.html</a>

#### 3.10.2.1 プロジェクト概要

OMRAS は、大量の音楽情報が様々な形式でインターネット上のオンラインデータベースに格納されているような状況において、コンテンツに基づく効率の良い音楽情報検索の技術及び使い易いプロトタイプシステムを研究開発する米英共同プロジェクトである。

特徴は、一般ユーザでも利用でき、2~3 年以内に産業利用可能な技術開発を目指している点である。しかし、プロジェクトには企業は参画していない。

<sup>1</sup> Fifth Framework Programme の略。欧州の産業振興のために EC (European Community)、EU (European Union)が支出する研究開発費は、1984 年以降 framework と呼ばれる単位で管理されるようになった。1つの framework の期間は 4~5 年で、その期間中いくつかの重点的な研究テーマが設定される。Fourth Framework は 1994 年~1998 年 (総額 132 億 ECU)、Fifth は 1999 年~ 2002 年 (総額 150 億 Euro) であり、そのうちの約 24%が 4 年間の内に IT 関連、通信関連研究に支出された(年間 1200 億円)。FP6 は 2002 年~2006 年で総額 163 億 Euro である(その内 IT 関連は 22%程度)。



図 1 OMRAS のロゴ

### 3.10.2.2 研究テーマ

#### (1) 多種のデータフォーマット

インターネット上の音楽データにはオーディオデータ、MIDI データ、楽譜データ等がある。従来の音楽検索システムでは、楽曲に関するテキスト情報（曲名、演奏者、作曲者等）を用いることが多く、そのような多種のデータに渡って十分な性能を出しているシステムは少ない。

OMRAS で問合せ可能なデータフォーマットは以下の 3 つである：

- a. 楽譜データレベル（市販楽譜エディタ等で用いられているもの）、
- b. 演奏データレベル（MIDI）、
- c. サウンドデータレベル（デジタルオーディオ、例えば wav）。

この 3 種類に対応した音楽データ表現を設計し、データベース中やシステム内部で利用する。

#### (2) UI

従来のテキストベースの音楽検索システムにおける UI の欠点として、問合せの作成が面倒、結果表示が分かりにくい、システム利用に習熟しなくてはならない、メンタルモデルがアドホック等が指摘されている。OMRAS の検索問合せと結果表示では、実際の音や可視化を用いた直感的なインタフェースを採用する。検索問合せを投入するインタフェースでは、問合せの新規作成、再利用、修正が容易に実現でき、ユーザは検索問合せの結果をチェックし次の問合せに反映させることができる。また、内部表現データを可視化、再生、モニタする機能も開発され、特に非専門家のための GUI 構築に注力している。

#### (3) 検索手法

検索手法に関連した研究テーマを挙げる：インデックス生成法、インデックスの最適長

決定法、インデックス・ファイルの検索法/管理法、音楽的（近似）照合法の設計等。

#### (4) 評価用ベンチマークの作成

ISMIR 2002 では Stephen Downie (University of Illinois at Urbana-Champaign) が chair を務めた "Creation of Standardized Test Collections, Tasks, and Metrics for Music Information Retrieval (MIR) and Music Digital Library (MDL) Evaluation" というワークショップが開催された。MIR/MDL 業界における TREC (Text Retrieval) 風の共通標準評価枠組みの構築を目指している。本ワークショップ参加者が寄稿した原稿からなる白書が公開されている (<http://music-ir.org/evaluation/wp.html>)。考慮すべき問題点として挙げられているのは、多種フォーマット音楽情報（オーディオ、記号、楽譜、メタデータ等）の合理的な統合法、実世界における問い合わせの分析と問い合わせタスクの選定等である。

Draft #2: Simplified Schematic of Internationally Accessible Yet Secure Music Information Retrieval Development and Testing Database

27 October 2002  
J. Stephen Downie

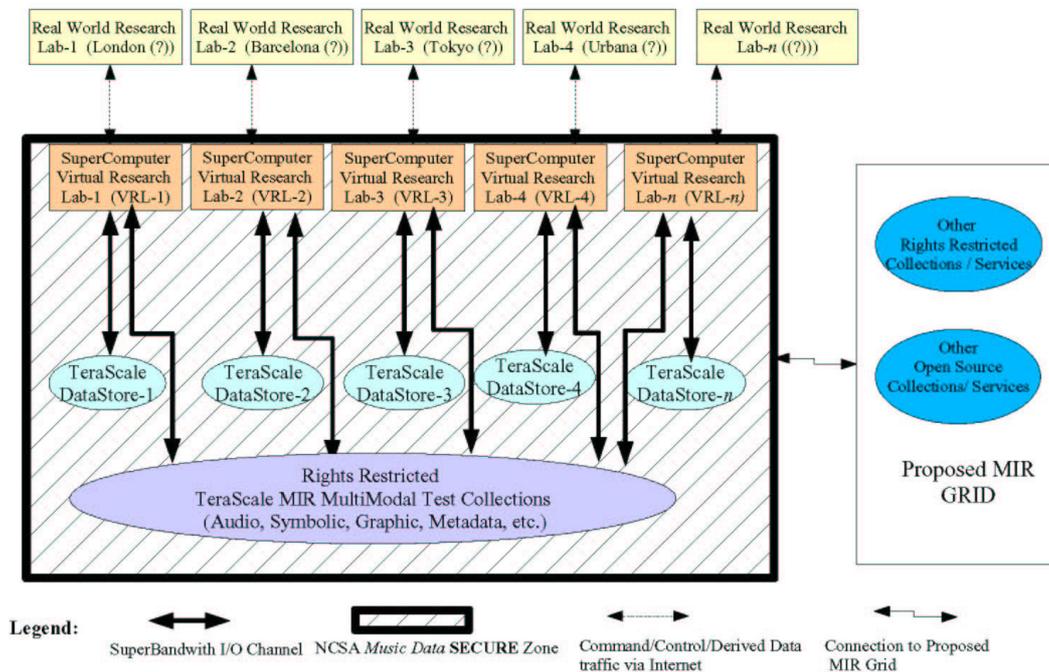


図2 音楽情報テストデータベースの利用イメージ

## 3.10.2.3 その他

NSF は ISMIR (International Conference on Music Information Retrieval) の開催を要請し、マサチューセッツ大学の提案を採用した (第 1 回 ISMIR 2000)。

プロジェクト成果物に関する情報が得られなかったので、調査を続行したい。

直接的後継プロジェクトかどうかは不明だが、Digital Music Research Network <http://www.elec.qmul.ac.uk/dmrn/> というプロジェクトが 2002 年 12 月発足している。

## 3.10.3 CUIDADO (Content-based Unified Interfaces and Descriptors for Audio/music Databases available Online)

表 2 CUIDADO プロジェクトの概要

研究費	ESPRIT <sup>2</sup> 28793, FP5
研究組織	CUIDADO Working Group: IRCAM, SONY CSL Paris, Univ. of Pompeu Fabra (Spain), Ben Gurion University (Israel), Oracle Spain, CreamWare (Germany), ArtsPages (Norway)
研究代表者	Huges Vinet (IRCAM)
プロジェクト期間	ワーキンググループ発足 2000 年 (ESPRIT 28793)。 FP5 は 2001 年 1 月からの 3 年間
参考資料	<a href="http://www.cuidado.mu/">http://www.cuidado.mu/</a>

## 3.10.3.1 プロジェクト概要

次の 3 つの技術開発を目標としている。(i) オンライン音楽データベースのためのコンテンツ (オーディオ) に基づく統一的なインタフェースとそのデスクリプタ抽出法。(ii) プロ音楽家やスタジオのための楽曲特徴を利用したオーサリングシステム及び音楽検索システム。(iii) レコード会社や著作権保有者が情報管理する際に必要となる技術。

対象とする音楽データは、低次のオーディオデータ (サンプル) から高次のデータ (曲名、スタイル、リズム等) まで幅広い。MPEG 7 と MP3 規格を採用することで汎用的な技術開発を目指す。

OMRAS と比較すると、プロジェクトとしては小規模であるが、これは目標を絞り込ん

<sup>2</sup> EU IT Programme の意味。情報関連産業を振興するには、タイミング良く研究開発を行い市場に投入していく必要がある。そのために EC/EU における IT 関連研究、テレコミュニケーション (ACTS) 研究、情報サービスやインフラの充実 (Telematics) 等に支出する研究開発予算を Esprit という予算枠に一本化した。FP5 以降は Society Technology program (IST) と呼ばれるようになった。

でいることの裏返しである。つまり、ユーザは専門家を想定し、欧州音楽産業との連携を強く意識し、扱うデータは MPEG 7、MP3 に限定している。

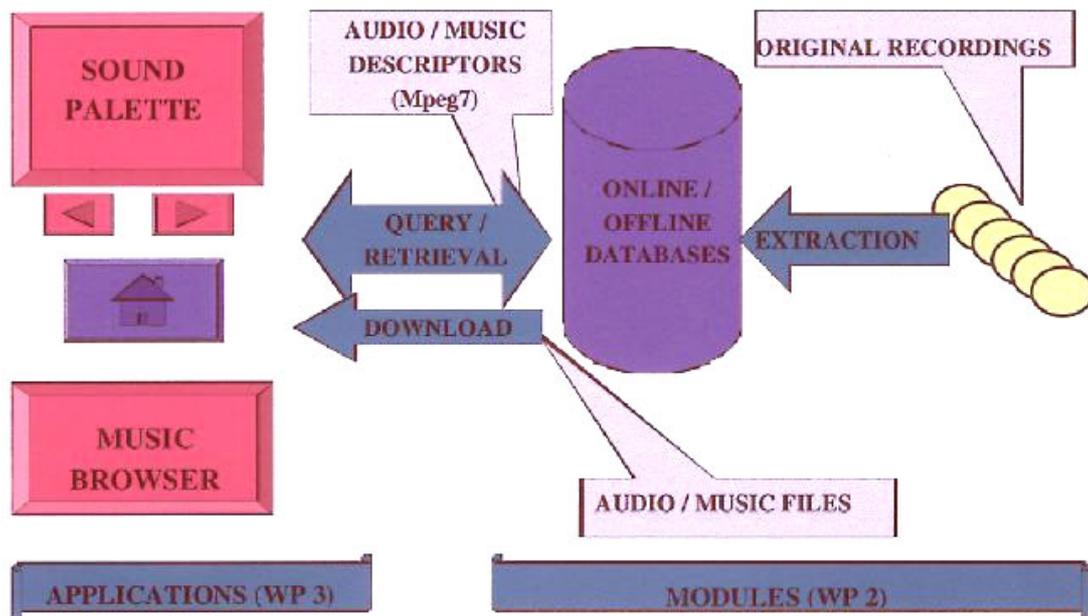


図 3 CUIDADO システムの全体構成

### 3.10.3.2 研究テーマ

#### (1) 研究対象とする 3 プロセスと研究内容

- a. 音楽分析プロセス: デスクリプタの抽出。最初は MPEG 7 の Instrument Timbre Description を参考にし、次に Multimedia Description Scheme を参考にする
- b. ナビゲーション・プロセス: 最新のデータベースシステム上に web と相互運用性を持った音楽検索法とインタフェースを実装する。Java Beans と MPEG 7 DDL あるいは XML を用いて実装。相互運用性は CORBA で実現する。
- c. 制作プロセス: 聴取環境とオーサリング・ツールの提供

#### (2) 従来の Web 上の音楽システムに対する不満

次のような不満がよく聞かれる: 情報過多なのに有用な情報が見つけない、類似したオーディオを検索できない、オーディオデータの高速なブラウジングができない、音楽をインターネット配信する際に貧弱なサービスしか付加できない、検索のためのキーワードは提供側しか指定できない。

**(3) 試作予定のアプリ**

ユーザの嗜好に従う音楽類似検索 (Music Browser) と自動編集 (Sound Palette) を実現する。

**a. The Music Browser**

Web 上の音楽ファイルをモニタする機能を持ち、Web 上の音楽販売とカスタマイズされたソフトウェアラジオとしても機能する。またユーザの個人性を考慮した学習機能 (制約を利用した playlist 生成、個人プロファイラ) もある。

**b. The Sound Palette**

音楽制作者のためのオーサリング・ツール (検索、編集、処理)。オーディオデータから抽出された特徴量を用いて、既存のオーディオ処理環境において稼働する。オンライン版 (web 上でもスケーラブルに動作。

<http://www.ircam.fr/produits/technologies/multimedia/cuidado-e.html>) とオフライン版 (既存のオーサリング・ツール上に構築したコンテンツに基づく編集・処理環境) が提供されている。

これら 2 つの試作プロトタイプは、以下の再利用可能なモジュールを含んでいる: オーディオ特徴量抽出器、統計的インデックス生成器、データベース管理機構、ネットワーク対応の制約に基づくナビゲータ。

**3.10.4 WEDELMUSIC (Web Delivering of Music)**

表 3 WEDELMUSIC プロジェクトの概要

研究費	FP5/IST
研究組織	学術系 Universita' di Firenze, Institute for Language and Speech Processing (Greece), Fondazione Scuola di Musica di Fiesole (Italy), IRCAM, Fraunhofer Gesellschaft (Germany), Stichting Studieen Vakbibliotheek voor Visueel en Anderzins Gehandicaptten (SVB) (Netherlands). 企業系 Artec Group (Belgium), Sugarmusic (Italy), CESVIT (Italy), Bmg Ricordi (Italy).
研究代表者	Paolo Nesi (Universita' di Firenze)
プロジェクト期間	2000 年 1 月より 2 年 4 ヶ月間
参考資料	<a href="http://www.wedelmusic.org/">http://www.wedelmusic.org/</a>

### 3.10.4.1 プロジェクト概要

WEDELMUSIC が目指す音楽環境とは、インターネットを介して出版社/作曲者からエンドユーザに直接音楽情報（楽譜記述等）が配信され、エンドユーザ側では配信された音楽情報を処理（編集、印刷）して演奏、器楽練習、学術等に利用するというものである。音楽をデジタル情報としてやりとりすることで、大きな新規マーケット開拓につながる事が期待される。

他プロジェクトと比較すると、独自の標準規格を目指している点が大きく異なる。例えば、楽曲に関するあらゆる情報が記述できるような言語の提案や、クライアント側のオールインワンのフロントエンド・ソフトウェア（機能的にはブラウザ、シェル、GUI、ツール群を兼ねる）の提案など。ユーザとしては、やはり音楽産業（出版社、レコード会社）が中心だが、大学、図書館、一般ユーザも含んでいる。

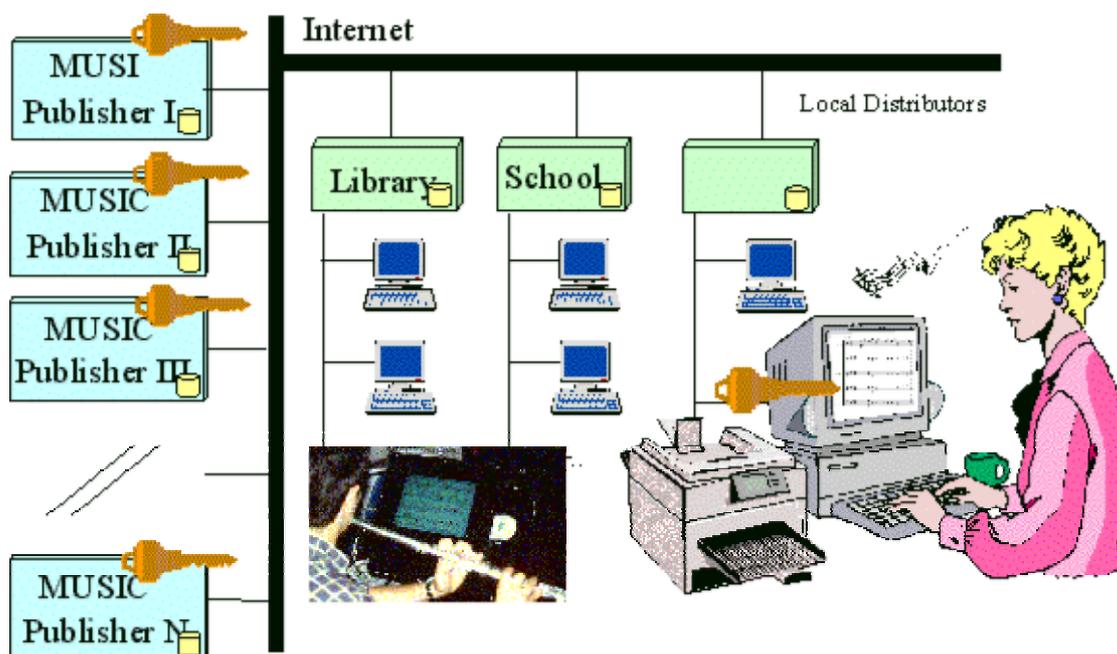


図 4 WEDELMUSIC システムの全体概念

### 3.10.4.2 研究テーマ

#### (1) 具体的な成果物

次の 3 点である：

- a. XML に基づくオーディオ、音楽記号、イメージ、ドキュメントを統一的に記述するためのフォーマット
- b. 信頼性の高い著作権保護機構

- c. インターネット上で音楽を創作、交換、格納、配布するためのツール (WEDELMUSIC editor) 現在 WEDELMUSIC editor v8.0 を配布中 (<http://www.wedelmusic.org/dld/dld.html>)。

## (2) WEDELMUSIC editor

以下にその特徴を挙げる: 多種類の楽譜が記述できる (記述力高い)、柔軟にカスタマイズ (フォント、ビュー等)、転調/ピアノ符への変換/奏法符生成等の便利なコマンド群、点字楽譜対応、歌詞記述機能、楽譜分析機能 (比較、検索)、種々のフォーマットに対応 (WEDELMUSIC XML Notation, SCORE, MIDI, MOODS, etc.)。

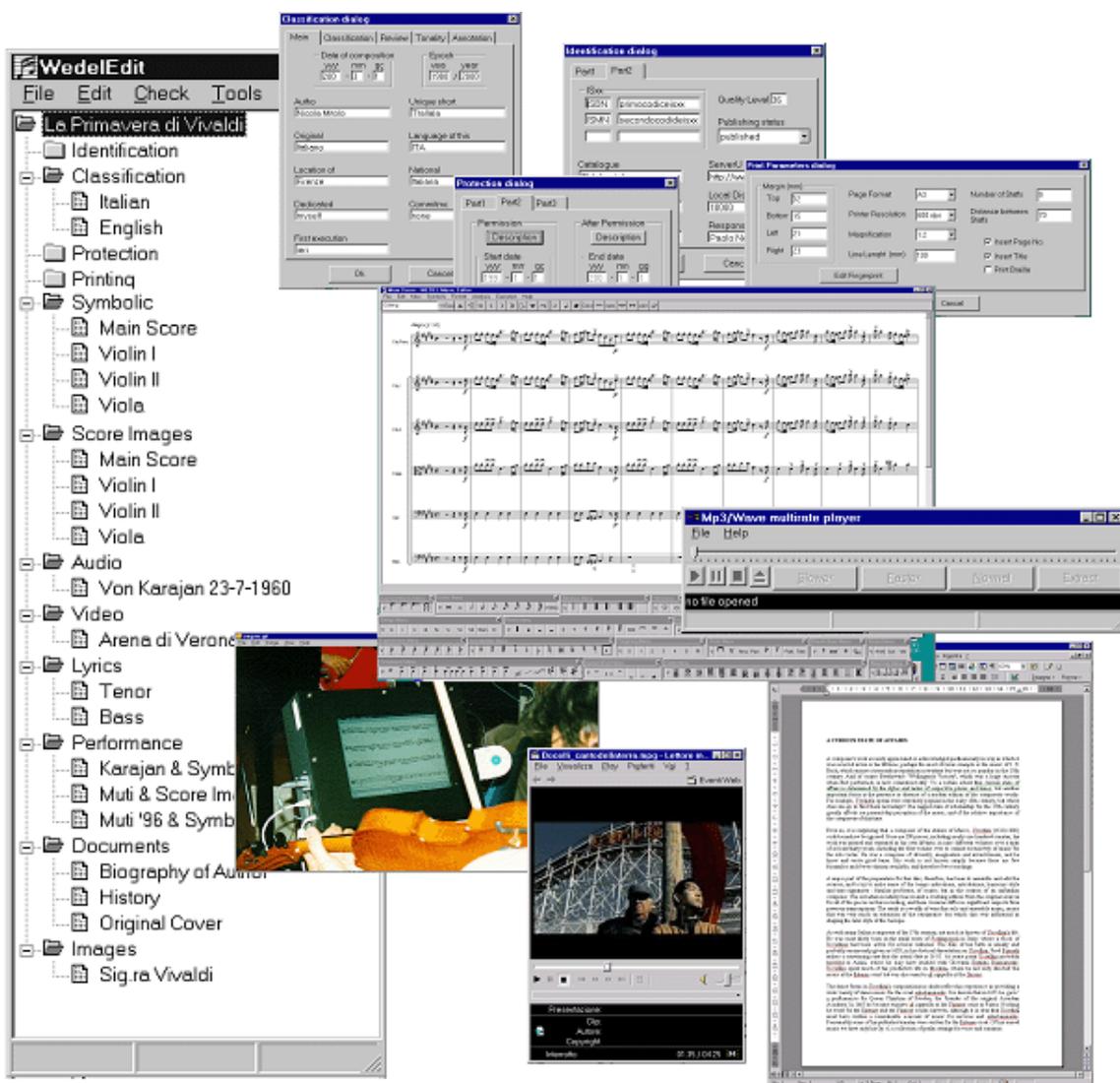


図5 WEDELMUSIC editor の作業イメージ

### 3.10.5 MUSICNETWORK

表 4 MUSICNETWORK プロジェクトの概要

研究費	FP5/IST, FP6/IST
研究組織	学術系: Universita' di Firenze, Institute for Language and Speech Processing (Greece), Fraunhofer Gesellschaft (Germany), Musik Informations Centrum Austria (Austria), University of Leeds (UK), IRCAM. 企業系: Central Research Laboratories (UK), Amazing Music World (Denmark), Interactive Labs (Italy), Exitech (Italy), Notissimo (France), Stichting FNB (Netherlands), Arca Progetti (Italy), Rigel Engineering (Italy).
研究代表者	Paolo Nesi (Universita' di Firenze)
プロジェクト期間	2002 年 8 月より 3 年間 (FP6)
参考資料	<a href="http://www.interactivemusicnetwork.org/">http://www.interactivemusicnetwork.org/</a>

#### 3.10.5.1 プロジェクト概要

楽譜記述の技術を確立し欧州の音楽産業やコンテンツ提供者と共にインタラクティブなマルチメディア産業を振興することを目標とし、音楽情報に関する、標準化、表現法、配信、著作権保護、セキュリティ、可視化等の技術開発を行う。コンテンツに関連した最新の音楽分析技法の開発も行う。ユーザとしては、個人消費者、出版社、マルチメディア配信者、音楽ショップ、教育機関、ソフトウェア開発者等を想定している。

さらに産業界や社会・文化へ波及効果を与えることもプロジェクトの目標に含まれる点が特徴的である。例えば、音楽コンテンツ提供者/産業が研究成果により、アクセスし易い協調的な環境の創出、マルチメディア音楽のための技術に関する講習会・技術移転の支援、著作権保護を遵守する一方でマルチメディアコンテンツの制作・配布のための積極的なコーディング規格標準化活動等である。これより結果として、(a) マルチメディア音楽配信のマーケットにおける欧州企業の競争力増強、(b) コンテンツ保持者/提供者の協調と競争を促し、新しい技術やモデル化技法を開発し、音楽アーカイブをデジタル化する際の障壁を取り除くこと、(c) FP6 期間中に発生するであろう新ニーズへの準備、(d) e-publishing、e-book、電子図書館、エンタテイメント・コンピューティング、e-commerce 等にならない、新しい音楽制作と消費市場の開拓等が期待されている。

MUSICNETWORK は WEDELMUSIC の枠組みを踏襲し、より押し進め具体化したものと言える。先行プロジェクトは WEDELMUSIC の他に HARMONICA、MIRACLE、CANTATE が挙げられる。



図6 MUSICNETWORK のコンセプト概略

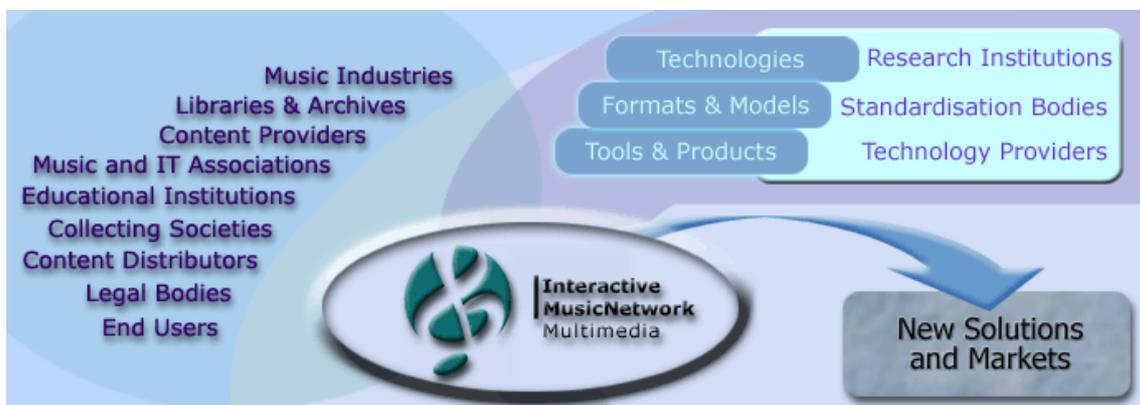


図7 MUSICNETWORK の貢献が期待される分野

### 3.10.5.2 研究テーマ

7つのワーキンググループでは次の研究開発活動を行う。

- (1) マルチメディア音楽記述方式 (WEDELMUSIC エディタをベースに)、
- (2) 音楽ライブラリ (複数ドメインを横断、メタデータ、情報検索等)、
- (3) 音楽記述の標準化 (MPEG 7/21 等を検討、インターネット対応、メディア統合)、
- (4) 音楽配信 (B2B、B2C、P2P、モバイル端末、Web-TV、トランザクションモデル等)、
- (5) 音楽保護 (セキュリティ、暗号化)、
- (6) アクセス容易性 (UI、盲人用点字楽譜等)、
- (7) 楽譜イメージ処理 (浄書、手書楽譜読取、古典楽譜等)。

### 3.10.5.3 その他

研究成果は、随時、関連する国際会議やイベント等で発表していく。

2002年12月11～12日 Darmstadt にて第1回 MUSICNETWORK Open Workshop が開催された(WEDELMUSIC 2002 に併設)。

### 3.10.6 SALIERI/GUIDO

表5 SALIERI/GUIDO プロジェクトの概要

研究費	Darmstadt University of Technology (TU Darmstadt、ドイツ)
研究組織	TU Darmstadt
研究代表者	Holger H. Hoos (現 Univ. of British Columbia, カナダ (UBC) & TU Darmstadt)
プロジェクト期間	1992年開始。1998年より GUIDO の開発に主軸が移る
参考資料	SALIERI: <a href="http://www.informatik.tu-darmstadt.de/AFS/CM/SALIERI/">http://www.informatik.tu-darmstadt.de/AFS/CM/SALIERI/</a> GUIDO: <a href="http://www.salieri.org/guido/">http://www.salieri.org/guido/</a>

#### 3.10.6.1 プロジェクト概要

SALIERI の研究目標は、音楽のアルゴリズム的な側面に焦点をあてて、(1) 作曲、編曲、楽曲分析のモデル化、(2) 音楽表現及び処理、(3) 応用システムや言語の設計と実装の3つである。SALIERI システムは、SALIERI 言語インタプリタとコマンドライン UI あるいは GUI から成る。

SALIERI は Holger Hoos の個人的な長期プロジェクトであり、研究費は大学のみから得ているようだ。一方ではこのような個人的なプロジェクトを継続しつつ、同時に前述のような大規模な共同プロジェクトに参加するというスタイルの研究者は多いようだ。本節では音楽記述に関連したプロジェクトという理由で GUIDO を取り上げる。

#### 3.10.6.2 研究テーマ

##### (1) GUIDO 楽譜記述言語

GUIDO は楽譜レベルで音楽を表現する形式言語である。伝統的な五線譜に含まれる情報だけでなく、ユーザが処理に必要とするような情報を定義し記述することができる。記述対象を柔軟に拡張することができるので、単純なモチーフから複雑なオーケストラ楽譜まで容易に記述できる。

**(2) GUIDO ソフトウェア**

現在提供されている GUIDO ソフトには次のものがある: NoteAbility (楽譜エディタ)、GUIDO パーザ (C 言語)、GUIDO ファイル操作ツール、GUIDO フォーマット処理関数。

また GUIDO HP (<http://tempo.iti.informatik.tu-darmstadt.de/salieri/nview/noteserver.html>) では GUIDO NoteServer という無料オンラインサービスが提供されており、ユーザが入力した GUIDO 記法 (GMN) の楽譜を五線譜の image ファイル (GIF, JPEG, PS) に変換できる。

The image displays a musical score for the hymn "O Lamm Gottes". It consists of eight staves of music in G major (one sharp) and 12/8 time. The lyrics are in German and are aligned with the notes. The score is as follows:

O Lamm Got- tes un- schul- dig,  
 Am Stamm des Kreu- zes gæhlach- tet.  
 all- zeit er- fund'n ge- dul- dig,  
 wie- wohl du wa- rest veracht- et  
 All Sünd hast du ge- tra- - - gen  
 sonst müß- ten wir ver- za- - gen;  
 er- barm' dich\_ un- ser, O Je-  
 su, O Je- su

© by GUIDO Noteserver 0.8 (final) (Nov 21 2001) www.noteserver.org

図 8 GUIDO NoteServer の出力例

### 3.10.6.3 その他

SALIERI プロジェクトは Holger Hoos と Thomas Helbich (TU Darmstadt) によって 1992 年に始められた。1992 年に試作された SALIERI プロトタイプシステム用に楽譜記述言語が設計され、それが GUIDO 音楽記述言語の前身となる。楽譜に現れる音楽的な構造のレベルで表現、操作するというコンセプトで設計された（個々の音符のレベルではなく）。1996 年 H. Hoos による SALIERI の仕様の見直しが行われ、GUIDO 楽譜記述言語が誕生する。同時に sic!system (Salieri Interactive Computing Inc.) が設立され SALIERI システムの開発を担当する。1997 年以来 Keith A. Hamel (UBC) が GUIDO 言語や処理系の改良や実装を担当している。

## 3.11 次世代音声認識の手法と実用化の動向

児島 宏明 委員

### 3.11.1 はじめに

現在の音声認識は、HMM<sup>1</sup>（隠れマルコフモデル）と統計的言語モデルに基づく手法が主流となっており、その枠組み内では技術的にはほぼ確立されている。本章ではまず、このような枠組みの限界を検討し、その改善を目指す試みを紹介する。そのような手法のうち、ここでは主として音響モデルを中心とした認識アルゴリズムについて概観し、位置づけと方向性をまとめる。次に、このような音声処理技術の応用に関して、実用化の展望を探る。そのなかでも、音声認識技術の応用については、これまでに、認識対象語彙の拡大や認識精度の向上などの改良が進められてきたにもかかわらず、現在のところ、広く普及し利用されるには至っていない。ここでは、そのような広範囲の普及の鍵となるような要因として、小型化と低価格化を中心に検討し、その事例を紹介する。

### 3.11.2 次世代音声認識手法の動向

#### 3.11.2.1 HMM の限界

まず、現在確立されている音声認識手法の概要とその問題点を述べる。現在主流となっている統計的音声認識の枠組みは、以下のように定式化される。

$$\hat{w} = \underset{w}{\operatorname{argmax}} P(w|x) = \underset{w}{\operatorname{argmax}} P(x|w) \cdot P(w)$$

このとき、 $P(x|w)$ を音響モデルに、 $P(w)$ を言語モデル対応させることにより、認識スコアは対数尤度を用いて「音響モデルの尤度 + 言語モデルの尤度」で表されることになる。通常、音響モデルとしては、音素を単位とする HMM を単語辞書に基づいて連結したものが用いられる。言語モデルとしては、範囲の限定された小さいタスクでは、構文的文法やネットワーク文法などが用いられるが、ディクテーションなどの大語彙タスクでは、そのような文法を記述することが実際上不可能であるため、連鎖統計モデルが用いられる。連鎖統計モデルは、n-gram と呼ばれ、n 個の単語の並びが出現する確率を統計的に求めたものである。

このような枠組みの限界を考える際にも、言語モデルの限界と、音響モデルの限界に分けて議論することができる。音響モデルの改善による約 5%の認識率の向上と、言語モデルの改善による約 50%のパープレキシティ（平均分岐数）の減少とが、ほぼ同等の効果に

---

<sup>1</sup> Hidden Markov Model

相当するという報告もあり[1]、ここでは、主として音響モデルについて議論を進める。ただし、米国 DARPA のプロジェクトで採用された SWITCHBOARD コーパスのような、自由発話文に対しては、未だに 50%程度の認識率しか達成されておらず、このようなタスクに対しては、言語モデルの本質的な改良が必要と考えられる。

一般に、HMM の限界として次のような要因が挙げられる。

- 1) 継続時間長のモデル化が弱い
- 2) 各時点（フレーム）での出力確率が、単一の状態（または2状態間のリンク）にのみ依存する（それ以外は独立）と仮定している。

1)は状態遷移を単純マルコフ過程で近似したこと、2)は出力確率の依存条件を単一のリンクで近似したことに起因する限界である。

これら各々の弱点を回避するための手法として一般的に、1)に対しては、明示的な継続時間長モデルの導入が、2)に対しては、回帰や導関数に基づく動的特徴量が広く用いられている。2)に対しては、これ以外にも、条件付き Gaussian HMM[2]や Segmental HMM[3]の適用が検討され、有効性が確認されている。また、継続時間長以外にも単純マルコフ近似による限界として、音韻的環境以外にも考慮すべき様々に異なる状況（例えば、状態ごとの確率分布で表現しきれないような話者による変動や、韻律環境など）に対するモデルの依存性が表現できない、という問題がある。その対応策としては、Bayesian Net[4]や Markov Random Field[5]などが研究されている。また、上記のような近似による悪影響を直接的に回避する手法として、変分ベイズ法[6]なども検討されている。また、モデルの分布が事象に正しく適合していない場合は、上記のような最尤の枠組みの前提がくずれることになるため、上記の式に基づく決定が必ずしも最適とはならない。このため、直接的に識別誤りを最小にする手法として、識別誤り最小化（MCE<sup>2</sup>）規準に基づく手法が研究され、有効性が報告されている[7]。

以下の各節では、このような問題点の解消を目指して現在の手法を拡張していく方向性を、4種類の流れに大別して紹介することにより、今後の音声認識手法の動向を位置づける。

### 3.11.2.2 セグメントモデルに基づく HMM の拡張

通常の HMM が各状態（または各リンク）から 1 つの観測ベクトルを出力するのに対し、M. Ostendorf らは、ひとつの状態（セグメント）から可変長の観測ベクトル列を出力するような統計的セグメントモデルを定式化した[8]。出力を可変長のベクトルとして扱うこ

---

<sup>2</sup> Minimum Classification Error

とにより、モデルの出力を区分定数としてでなく、滑らかな軌道 (trajectory) として表現することができ、表現能力が拡張される。また、セグメントモデルに確率的動的システムを導入することにより、通常の HMM および、その改良としての継続時間長制御 HMM や条件付き Gaussian HMM などは、セグメントモデルの特殊な場合として位置づけることができる。出力される trajectory の決定方法や、セグメント化の方法などに関して、多くの可能性が考えられ、HMM を発展させる枠組みの指針として期待される。

### 3.11.2.3 複数モデルによる HMM の拡張

HMM を含む統計的パターン認識手法に共通する限界として、認識対象が学習用サンプルに適合する場合は有効に機能するが、学習用サンプルとして想定外の入力に対しては、精度が大幅に低下するという問題がある。例えば、マイクの特長や雑音環境や残響などが、想定したものと異なれば、認識精度は大きく低下する。これに対処するための最も一般的な方法は、学習サンプルを対象に合せて拡張するとともに、モデルのパラメータや状態数を増やして、HMM の表現能力や精度を向上させることである。これまでにも、そのような方向性としては、出力確率分布の混合数を増やしたり、音素単位のモデルを triphone と呼ばれるような音素環境依存型のモデルにしてモデル数を増やす、などの拡張が行われてきた。このような方向は、CPU の演算速度など処理能力の向上と、学習用のサンプルなどの音声データベースの大容量化が前提となるが、近年のハードウェア技術の急速な進歩により、適用可能な規模は急激に拡大している。

処理能力が拡大したとしても、単純にモデルの状態数や学習サンプルの種類を増やすだけでは、学習が局所最適解に陥る可能性が高く、有効に機能しない。これに対し、異なる環境のサンプルで学習したモデルを、単一の HMM の中の複数のパスとして統合するマルチパス音響モデルの研究が、近年活発に進められている [9]。これは、学習サンプルの状況を予めヒューリスティクスにより整理することにより、学習が有効に進むようにした手法として捕えることもできる。

このような方向をさらに進めれば、統計的な一般化を行わず、サンプルをそのまま標準パターンとして記憶するような、事例に基づく手法に発展する可能性がある。このような手法は、現在の統計的手法が主流になる前に、DP<sup>3</sup>マッチングや k-NN<sup>4</sup>などの手法とともに広く使われてきたが、学習サンプル数やモデルの精度の問題もあり、統計的手法に対抗できなかった。しかし、近年は人工知能や自然言語処理などの分野において、事例に基づく手法が注目されてきており、今後さらに計算機の処理能力が拡大すれば、パターン処理

<sup>3</sup> Dynamic Programming

<sup>4</sup> k - Nearest Neighborhood

においても、大きな流れとなる可能性があると考えられる。

#### 3.11.2.4 大局的特徴の利用

現在の手法では、ほとんどの場合、HMM を音素単位でモデル化しており、文脈や状況に関しては、前後の音素程度の環境しか考慮していない。これに対して、大局的状況に関する特徴を利用することにより、性能の改善が期待される。

大局的状況としては、収録環境なども含めて様々な要因が考えられるが、ここでは主としてコミュニケーションに関する状況について述べる。そのような状況は大局的特徴としては主に韻律として現れる。韻律的特徴には、アクセントによる単語の意味の違いや構文的な区切りといった言語的状況以外にも、話し手の自信や態度といったパラ言語的状況や、感情などの非言語的状況なども反映される。韻律的特徴としては、基本周波数、パワー、発話速度などがあるが、これ以外にも母音の韻質と呼ばれるような、フォルマントやスペクトル的特徴としても現れる。従って、韻律的状況によって音響モデルを変えたり制御したりすることにより、性能の改善が期待できる。それ以外にも、韻律から構文的情報を取り出し、これを利用して、認識候補の探索空間を最適化する手法などが研究されている[10]。従来から、韻律的特徴の重要性は認識されていたものの、認識に有効に利用することは困難とされてきたが、音声認識の対象が自由発話や話し言葉に移行するにつれて、韻律の重要性は高まっており、今後の研究の進展が期待される。

韻律的特徴に限らず、スペクトル的特徴空間における大局的特徴を音韻的特徴と並行して抽出し、認識に利用する手法は、今後研究が期待される分野である。そのひとつとして、大局的特徴をストリームとしてトラッキングしながら抽出する手法の研究が注目される[11]。複数話者の音声の分離だけでなく、話者ごとの大局的特徴を抽出して正規化に利用したり、認識の単位を音素だけでなく文節や文などの長い単位で扱うような場合の特徴抽出にも、応用が期待される。

#### 3.11.2.5 認識単位・記号系の検討

音声認識は、一般的に捉えれば、音響的信号を言語的記号に変換するシステムである。このようなアナログ情報（パターン情報）から記号情報への変換を、どの時点（レベル）で行うかということに関して、様々な可能性が考えられ、手法のバリエーションとなる。記号化により、計算機の処理の効率が向上するが、いくつかのパターンをひとつの記号に統合する際に、細かい情報が捨象され、一般的には精度が低下する。従って、認識対象に適したモデル化と記号化を行うことが、システム設計の上で重要である。音声信号を直接的に記号に変換する方式は、音声符号化と呼ばれ、簡易な手法としては、クラスタリング

に基づいて記号化を行うベクトル量子化 (VQ<sup>5</sup>) が広く用いられる。これは HMM に基づく音声認識にも適用され、VQ に基づいて出力確率を定義する離散 HMM や、VQ 符号に連続的確率分布を付加した Semi-Continuous HMM などがある。また、音響モデルの単位は、音声認識が出力する記号の基本単位であり、この設計は、認識手法や性能に大きな影響を及ぼす。以前には、簡易型の単語認識システムとして、単語単位の音響モデルが用いられることもあったが、現在では、ほとんどの場合、音素を単位とした音響モデルが用いられている。これに対して、音節、半音節、半音素、音素片など様々な単位が検討され、性能の改善も報告されている。また、このような単位を、人間が言語学的知見に基づいて予め与えるのではなく、音声サンプルからボトムアップに自動生成する手法も重要になると考えられる[12]。さらに、前節でも触れたように、状況に応じた可変長の認識単位を利用する手法なども期待される。

音素は、言語体系に応じて定義される単位であるため、多言語に対応した音声認識システムを構築する際に問題となる。言語ごとに認識システムを設計して、その結果を統合する手法も考えられるが、多言語をシステム中で統一的に扱う場合には、音素体系の共通化を考慮する必要がある。このような体系として、音声学の分野で用いられている国際音声記号 (IPA<sup>6</sup>) がある。これは、発音記号に近い単位であり、時間方向では音素と同程度の長さの単位となる。同じ音声記号で表される音声でも、異なる言語においては実際の物理的特徴にかなり違いがあると考えられるため、田中らは、VQ のレベルと IPA のレベルの中間的な共通記号系を設計し、音声認識や音声検索に適用する研究を行っている[13]。放送音声データベースなどを音声発話のキーワードで検索するような音声検索の手法としては、主として次の2通りが考えられる。

- 1) 音響的特徴や VQ のレベルでマッチングを行う。
- 2) 音声データベースをいったん音声認識によってテキストに変換し、同じくテキストに変換したキーワードと、テキストレベルで検索する。

1)は、マッチングの計算量が膨大になる問題が、2)は、認識誤りの問題や、多言語やノンネイティブなどの話者に応じた認識システムを構成する必要があるなどの問題がある。中間的共通記号系は、これらを最適化するような記号化レベルを探る試みである。

---

<sup>5</sup> Vector Quantization

<sup>6</sup> International Phonetic Alphabet

### 3.11.3 音声処理技術の実用化の動向

#### 3.11.3.1 音声処理技術の実用化の鍵

前節では、現在の音声認識アルゴリズムの限界と、それを乗り越えようとする試みについて述べた。これは、音声認識の認識精度自体を向上させようとする方向である。しかし、このような限界があるとはいえ、雑音の少ない環境で、マイクからの距離が近く、書き言葉的な文章をある程度正確に発声する、などの条件が満たされれば、現在の技術水準でも、数万語の対象語彙を含む連続音声に対して、90%以上の認識精度が得られる。つまり、この水準でも、使い方によっては、応用範囲を拡大していく可能性は多く残されていると考えられる。本節では、音声認識の応用が生活の中で広く普及するようになるための、認識精度以外の要因について述べる。

これまでに、本調査研究の報告において、音声認識技術の実用化を進めるための課題として、

- 1) 認識対象の拡大と認識精度の向上
- 2) ユーザーインターフェースの改良
- 3) 意味や意図など人間並みの知的能力の実現

などを挙げた[14]。また、現在の技術水準の制約のもとで適用可能な応用分野として、

- 1) 高度な意味処理を必要としないもの（ディクテーションや音声検索など）
- 2) 認識対象となる状況が限定されるもの（パソコンの OS やアプリケーションのメニューの代替、秘書システム、電話応答案内システムなど）、
- 3) 必ずしも正確な応答を必要としないもの（ゲームや玩具やペットロボットなど）

を挙げた[15]。本報告では、システムの実装上の観点を中心として、実際の製品化の上では、上記のような情報処理技術に関する要因よりも、むしろ小型化と低価格化といった実装上の要素技術に関する要因が重要である、という観点で述べる。

#### 3.11.3.2 小型化と低価格化

音声認識の研究は、通常ワークステーション（高速 CPU の PC や PC クラスタなどを含む）上で行われるため、その時点での最先端の性能を期待して製品化しようとする、サーバ型の大規模システムが必要になる。電話による音声応答(CTI<sup>7</sup> / IVR<sup>8</sup>)など、このような大規模システムでも実現可能な応用分野もあるが、通常の家電製品などを想定した場合は、小規模の組込みシステムとして実装する必要がある。従来は、組込みシステムにおけ

---

<sup>7</sup> Computer Telephony Integration

<sup>8</sup> Interactive Voice Response

る音声処理には、DSP<sup>9</sup>が用いられることが多かったが、近年は、ほとんどの場合、組み込み用の小型マイクロプロセッサ（MPU<sup>10</sup>）が用いられるようになってきた。これは、MPUの処理速度の向上と、低価格化によるものであり、現在は、数十から数百 MIPS の組み込み用低消費電力型 MPU が、数百円から数千円で利用できるようになってきている。すなわち、ここ数年で、組み込み用 MPU の処理性能は約 10 倍に向上し、同一性能あたりの価格は約 1/10 に低下した。

さらに、部品や実装上のコストだけでなく、音声認識技術自体のコストも低下している。HMM などの技術が確立され、開発用ミドルウェアやソフトウェアツールや普及することにより、専門家でなくても、比較的容易にプログラムの一部として音声処理機能を組み込むことが可能になってきている。また、MPU を採用することにより、DSP や専用チップを用いる場合と比べて、汎用のプログラミング言語を利用することができ、ソフトウェア上でも開発コストを低減できる。また、汎用 MPU は、他の機能を実現するためにも、いずれはほとんどの機器に組み込まれることになると予想されるため、そこに音声機能を追加する形であれば、コストはさらに低くなる。

このような状況の変化は、音声処理技術の応用のされかたにも変化をもたらす。これまでは、音声認識機能を組み込むことが、その製品のコストにおいて大きな負担となるため、その機能が大きな効果をもたらさなければ、採用する意味を見出せなかった。しかし今では、コストに占める割合を、無視できる程度に抑えることができるため、すべての家電製品に、いわばオマケの機能として組み込むことも、不可能ではなくなっている。例えば、駅の切符販売機やビデオレコーダの予約などに音声認識機能を使うような場合に、従来は、その機能の有効性を前面に押し出す必要があったため、少しの認識誤りがあっても「使えない」システムとして切り捨てられてしまう傾向があったが、通常のボタン操作などによるインターフェースと併用する形で、オマケとして音声認識機能を提供するようになれば、多少認識精度や使い勝手に問題があっても、身体的障害などにより通常のインターフェースの利用が困難なユーザや、手荷物などの状況によっては、利便性が向上につながるし、そうでない場合には従来のインターフェースを使える。このようにインターフェースを選択的に利用できるような製品を提供していくことにより、無理なく普及が進むと期待できる。これが進めば、将来的には、どのような製品にも音声認識機能が組み込まれ、どこに居てもボタンなどの入力装置を意識することなく機器を利用できるというような、所謂ユビキタスな環境の実現に近づくと期待される。

---

<sup>9</sup> Digital Signal Processor

<sup>10</sup> Micro Processing Unit

### 3.11.3.3 事例

上で述べたような、小型 MPU を用いた音声処理装置の製品の事例を紹介する。

写真 1 は、組み込み用音声認識装置の例である（ケンウッド社製）。約 20MIPS 程度の組み込み用 MPU を用い、数百から数千語程度の不特定話者音声認識を 20cm 角程度のワンボードで実現している。開発用ボードのため、プログラム書き換え用のフラッシュメモリなどが含まれているが、製品としてプログラムを動作させるだけであれば、MPU とメモリのみで構成して、3 から 4 チップ程度で実装できる。他に、このボードでは、下段のアナログ回路により、雑音対策などを実現している。

写真 2 は、1 DIN サイズの車載用 PC の試作機である（ケンウッド社製）。音声処理専用ではなく汎用装置の例として紹介する。これは、MPU としてクロック 800MHz の Crusoe (Transmeta 社製) を採用した、PC/AT 互換パソコンで、通常の Windows2000/XP や Linux などの汎用 OS が動作する。USB や PC カード (PCMCIA) や Bluetooth などの一般的な外部入出力インターフェースを備えているため、USB カメラや GPS などのセンサからの情報の取り込みや、無線 LAN によるインターネット接続なども可能である。近い将来において、道路上の通信インフラの整備が進み、これと連携した電子商取引システムが発展すれば、道路情報の把握やドライブスルーなどのサービスの利用が拡がり、ITS やテレマティークにおける中心的デバイスとなると期待される。また、自動車の制御情報も含めた車載電子機器の電子的なインターフェースの統一化が進めば、自動車の安全な運行にも寄与すると考えられる。このような車載 PC 上で、通常の汎用 PC 用の音声認識・合成ソフトウェアを問題なく動作させることができるため、自動車の電子化においても、音声処理技術の導入が急速に進むと予想される。



写真 1 組み込み用音声認識装置



写真 2 車載用 PC

## 3.11.4 おわりに

次世代音声認識の動向として、認識手法の発展の方向性を、セグメントモデル、複数モデル、大局的特徴、認識単位の4通りの潮流として紹介した。また、今後の音声認識技術の普及のためには小型化と低価格化が鍵となる、という観点で、実用化の動向について述べた。音声認識システムの課題としては、これ以外にも、複数の話者が同時に発話するような状況への対応や、話者がマイクから離れた場所で移動するような状況、雑音やBGMなどへの対応など多くの問題が残されている。現在は、多くのユーザが、音声認識システムを実際に使ってみると、期待に反して「使えない」と評価することが多いが、「使える」と評価に近い将来逆転することを、最後に希望的予測として記したい。

## [参考文献]

- [1] S. Nakagawa, "Relationship among perplexity word accuracy and phoneme accuracy, and drawback and modification of perplexity," Proc. First Int. Workshop East Asian Language Resources and Evaluation (1998).
- [2] C. J. Wellekens, "Explicit time correlation in hidden Markov models for speech recognition," in Proc. ICASSP '87, pp. 384-386 (1987).
- [3] M. Gales and S. Young, "Segmental HMMs for speech recognition," in Proc. Eurospeech '93, pp. 1579-1582 (1993).
- [4] G. Zweig and S. Russell, "Probabilistic modeling with Bayesian Networks for automatic speech recognition," in Proc. ICSLP '98, pp.3010-3013 (1998).
- [5] Y. Zhao, et. al, "Application of the Gibbs distribution to hidden Markov modeling in speaker independent isolated word recognition," IEEE Trans. SP, Vol.39, No.6, pp.1291-1298 (1991).
- [6] 渡辺、他, 「ベイズ的アプローチに基づく状態共有型 HMM 構造の学習」, 信学技報, SP2002-14, pp.43-48 (2002).
- [7] S. Katagiri, et. al, "Pattern recognition using a family of design algorithms based upon the generalized probabilistic descent method," IEEE Proceedings, Vol.86, No.11, pp.2345-2372 (1998).
- [8] M. Ostendorf, V. Digalakis, and O. Kimball, "From HMMs to segment models: a unified view of stochastic modeling for speech recognition," IEEE Trans. ASSP, Vol. 4, pp.360—378 (1996).
- [9] 李、他, 「認識エンジン Julius におけるマルチパス音韻モデルの実装」, 音響学会春季講演論文集, No.1, pp.89-90 (2002).

- [10] Shi-wook Lee, et. al, "Efficient Search Strategy in Large Vocabulary Continuous Speech Recognition Using Prosodic Boundary Information," in Proc. ICSLP-2000, Vol.IV, pp.274-277 (2000).
- [11] T. Nakatani and H. G. Okuno. "Sound Ontology for Computational Auditory Scene Analysis," In Proc. AAAI-98, pp.1004-1010 (1998).
- [12] H. Kojima and K. Tanaka, "Formation of phonological concept structures from spoken word samples," in Proc. ICSLP-92, pp. 269-272 (1992).
- [13] K. Tanaka, et. al, " Speech Date Retrieval System Constructed on a Universal Phonetic Code Domain," in Proc. IEEE Workshop ASRU2001, pp.1-4, (2001).
- [14] 「人間主体の知的情報処理技術に関する調査研究 III」 AITEC (2000).
- [15] 「人間主体の知的情報処理技術に関する調査研究 IV」 AITEC (2001).

## 3.12 アクティブオーディションを利用した ヒューマンロボットインタラクションの高度化

中臺 一博 講師

### 3.12.1 はじめに

近年、ヒューマノイドを代表とするロボットは AI のテストベッドなど研究目的としての利用にとどまらず、人間と知的なソーシャルインタラクションを行い、将来ロボットが“人間のパートナー”となることが期待されている。このような人間とのソーシャルインタラクションでは、人間のコミュニケーションの多くが音声に依存していることから明らかなように、聴覚は最も重要な機能の一つである。

そうした状況では、ロボットは複数のイベントを同時に聞き分け、他の音声や自分自身が作り出すモータノイズを抑制して、混合音をうまく扱えるように音源分離機能を備える必要がある。この機能は、音源分離をフロントエンド処理とした音声認識にも有効である。

音源分離問題については音環境理解 (*Computational Auditory Scene Analysis*, CASA) の分野で、様々なアプローチが取り組まれてきた。しかし、心理学的知見を利用した音源分離[11,17]、マイクロホンアレーを用いたビームフォーミング[3,18]、独立成分分析 (*Independent Component Analysis*, ICA)[19,9]、視聴覚統合による音源分離[16] など、そのほとんどはシミュレーション環境で行われており、実環境・実時間処理への配慮があまりされていない。実時間処理に関しては、2本のマイク間の強度差と位相差を利用した音声強調[2]が報告されているが、環境が既知であること、マイクや音源が静止していることが前提であり、ロボットへの適用は難しい。

また、ロボットでの音声認識については、ロボット自身が動作時に発するノイズ問題から、動作中の音声認識は難しく、Sony AIBO のように“*stop-perceive-act*”原理に従わざるを得ないか、部屋やロボット自身の影響による音声の伝達歪みや他の音源からのノイズの影響で、ロボットではなく話者の口元にマイクを設置している。

アクティブオーディションはこれらの問題に対する解決の糸口を与える[14]。これは、人間や動物を見習って、アクティブな動作を積極的に利用してロボット聴覚を向上させるもので、ロボット動作時に問題となるノイズをキャンセルすることにより、これまで、実時間・実環境での音源定位・追跡、分離を報告してきた[13,14,15]。

本報告では、音源定位・分離だけでなく、分離音の認識までを対象として、アクティブオーディションを利用したロボット聴覚システムを提案する。システムは、視聴覚を統合し正確に複数人物を定位・追跡する実時間人物追跡システム、フィルタの通過帯域を音源方向に応じてアクティブに制御し、分離精度を向上できるアクティブ方向通過型フィルタ

(*Active Direction-Pass Filter*, ADPF)、分離音声に対し、複数の音響モデルを用いた音声認識結果を統合して認識率を向上できる音声認識システムからなっている。また、実際にシステムを2本のマイクを搭載したロボットに実装し、その評価を行う。

以降、3.12.2章ではロボット聴覚システムについて述べる。3.12.3、3.12.4、3.12.5章では、ロボット聴覚システムを構成する実時間複数人物追跡システムによる音源定位・追跡、アクティブ方向通過型フィルタによる音源分離、および複数の音響モデルを利用した音声認識をそれぞれ説明する。3.12.6章でシステムを評価し、3.12.7章でまとめる。

### 3.12.2 ロボット聴覚システム

アクティブオーディションを利用したロボット聴覚システムを図1に示す。システムは、ロボット(SIG)のカメラ、マイク入力から、音源が複数ありかつ動作している場合でも、ロボット自身のアクティブな動作、視聴覚の統合により、これらを定位・分離・認識することが可能である。システムは大きく3つのサブシステム「視聴覚を統合した実時間複数人物追跡」、「アクティブ方向通過型フィルタ(以後、ADPF)による音源分離」、「複数の音響モデルを使用した音声認識」からなる。以下にヒューマノイド SIG を紹介し、次章以降で各サブシステムを説明する。

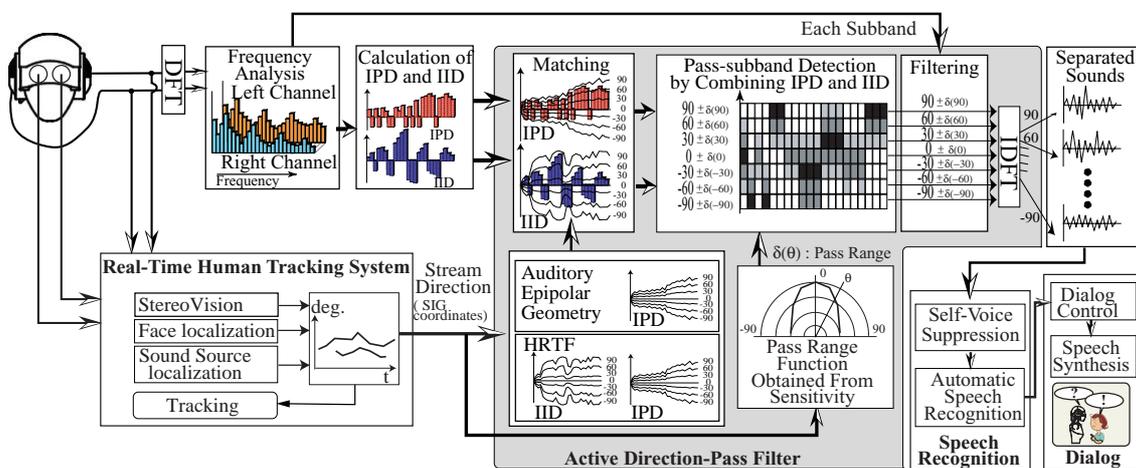


図1 ロボット聴覚システムのアーキテクチャ

#### 3.12.2.1 ヒューマノイドプラットフォーム:SIG

研究のテストベッドとして、上半身のヒューマノイド SIG を使用している。SIG は4自由度を有し、各モータには、ポテンシオメータによって、位置制御、速度制御が可能なDCモータを用いている。また、音響的にロボットの内外を区別できるよう設計されたFRP製の外装を備えている。カメラには、左右の目の位置に一組のCCDカメラ (Sony

### 3.12 アクティブオーディションを利用したヒューマンロボットインタラクションの高度化

EVI-G20) を、マイクには、計 4 本の無指向性マイク (Sony ECM-77S) を使用している。4 本のマイクは、外装を挟んで一組ずつ取り付けられており、内部に設置されている一組は、主にロボット自身のモータによって発生する内部ノイズをキャンセルするために使用している[14]。音源定位・分離には、外界からの音響信号を收音するよう SIG の左右の耳部に設置されているもう一組のマイクを利用している。

#### 3.12.3 視聴覚統合による実時間複数人物追跡

図 1 の実時間複数人物追跡システムの詳細な構成を図 2 に示す。このシステムは、SIG のカメラやマイクなどから得られるセンサ情報を統合して、複数の人物の位置を把握し、これを追跡することができる。本報告では、このシステムで認識される正確でロバストな音源方向情報をアクティブ方向通過型フィルタへの入力としている。

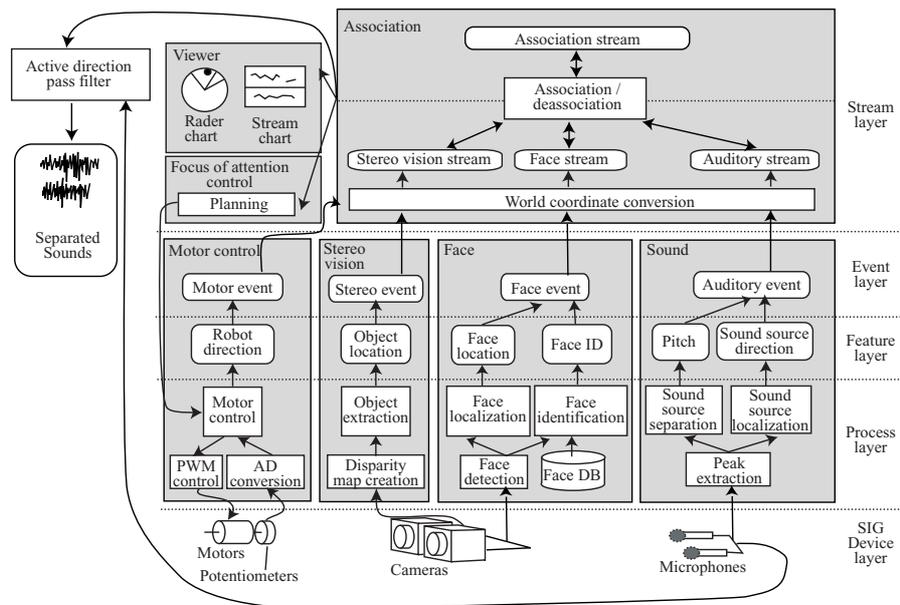


図 2 実時間追跡システムの階層的なアーキテクチャ

システムは、音源定位、顔認識・定位、ステレオビジョン、アソシエーション、アテンション制御、モータ制御、ビューワ の 7 モジュールから構成されている。なお、以降ではモジュール名はボード体で表す。

モジュール内のサブモジュールや情報は、5 階層に分けられる。SIG のカメラ、マイク、モータシステムなどのセンサデバイスは SIG デバイス層に属している。

プロセス層、特徴層、イベント層では、SIG デバイス層から得られるセンサ情報から、位置、名前情報といった特徴を抽出し、これらに時間情報を付加して、特徴の種類に拠ら

ない透過的な表現であるイベントに変換し、ストリーム層に出力する。各センサ情報の観測タイミングは非同期であるため、イベントも非同期に発生する。これらの処理は、**音源定位、顔認識・定位、ステレオビジョン、モータ制御**といったイベント抽出モジュールで行われる。

具体的には、**音源定位**では、マイク入力から、ピッチ(音高)を抽出し、倍音の調波構造を利用したグルーピングを行い、複数の音源が同時に存在する場合でも、それらを定位し、音イベントを生成する。

**顔認識・定位**では、顔画像の認識と定位を行い、顔の ID と位置情報を含んだ顔イベントを生成する。顔認識では、まず、肌色抽出と相関演算に基づくパターンマッチングにより、顔領域抽出を行う[7]。次に、抽出顔領域に対しオンデマンド更新が可能で最適な判別空間を生成できるオンライン判別分析(LDA)[8] を適用して顔認識を行う。顔定位は、抽出した顔の大きさに一定の仮定をおくことで、顔の 3 次元定位を行う。

**ステレオビジョン**は、高速な視差マップ生成法[10]を用いて、人物のように縦に長い物体を抽出・定位し、人物位置情報を含んだステレオイベントを生成する。横を向くなど顔が見えない場合にも人物の位置情報を得ることができるため、システムのロバスト性を向上することができる。

**モータ制御**では、モータのポテンシオメータから得られるロボットの姿勢情報をもとに、モータイベントを生成する。

ストリーム層では、イベントを種類ごとに時間方向に接続し、ストリームを形成する。ストリームとイベントの接続には、**Kalman** フィルタを用いて、観測誤差・処理誤差にロバストな処理を実現している[15]。

このようにして生成されたストリームのうち、複数のストリームが同じ人物に由来すると判断すると、これらを一つに束ねアソシエーションストリームを生成することにより、センサ情報の統合を行っている。これらの処理は、**アソシエーション**で行われる。アクティブ方向通過型フィルタの入力は、音ストリーム、および音情報を含んでいるアソシエーションストリームの方向情報である。送出されるストリーム方向情報には、時間情報およびストリーム ID が含まれているため、データの同期および複数音源分離が可能である。

また、**注意制御**はストリームの状態に応じて SIG の動作を決定し、**ビューフ**はストリームの状態を、レーダチャート、ストリームチャートとして可視化するためのモジュールである。**モータ制御**は、**注意制御**からの信号をもとに、PWM(Pulse Width Modulation) 信号を生成し、DC モータを駆動するためにも使われている。

実装上は、これらのモジュールをギガビットイーサ、ファストイーサの 2 つのインタフェースを備えた 5 台の Linux ノード(Pentium III 1GHz)に分散させている。ギガビット

イーサは、トラフィックが多く、通信量も大きいモジュール間通信用に、ファストイーサは、同期信号などの通信用に使い分けている。結果として、200msのレイテンシ、および100  $\mu s$ 以下の精度のノード間同期を実現している。レイテンシについては、ストリーム生成で用いたKalmanフィルタを予測に利用することで補っており、これによりリアルタイム動作を可能としている。

#### 3.12.4 アクティブ方向通過型フィルタによる音源分離

図1の網掛け部分がアクティブ方向通過型フィルタの構成に対応する。アクティブ方向通過型フィルタへの入力は4つあり、入力のスペクトル、入力スペクトルから計算されるIPDとIID、および、実時間人物追跡システムから得られる音源方向情報である。出力は、入力方向に対する分離音響信号である。

アクティブ方向通過型フィルタでは、方向通過型フィルタに対し、聴覚中心窩に基づくアクティブな通過帯域制御とロボットの伝達関数を利用した仮説生成により、実環境での高速な音源抽出を可能にしている。ここで、ロボットの伝達関数は、部屋の伝達関数、ロボット頭部による音の歪みなどを考慮して、特定方向のIPDおよびIIDを推定するための関数である。以下では、アクティブ方向通過型フィルタのアルゴリズムの詳細について説明する。

##### 3.12.4.1 アクティブ方向通過型フィルタのアルゴリズム

アクティブ方向通過型フィルタのアルゴリズムは以下の6ステップで構成される。

1. ある時刻に左右のマイク入力信号から得られたスペクトルから、各サブバンドのIPD  $\Delta\phi'$  と IID  $\rho'$  を計算する。
2. 抽出すべき音源の方向を  $\theta_s$  とする。 $\theta_s$  は3節で述べる実時間人物追跡システムから、ロボット座標系での水平角として得られる。
3. 通過帯域関数に従って、 $\theta_s$  に対応するアクティブ方向通過型フィルタの通過帯域  $\delta(\theta_s)$  が選択される。通過帯域関数は、聴覚中心窩に基づき、ロボットの正面方向で最小となり、周辺部で大きな値をとる関数である。詳細は4.2.2節で述べる。選択された通過帯域  $\delta(\theta_s)$  を用いて、 $\theta_l = \theta_s - \delta(\theta_s)$ 、 $\theta_h = \theta_s + \delta(\theta_s)$  と定義し、 $\theta_l$  から  $\theta_h$  の範囲にある音響信号を抽出するのがアクティブ方向通過型フィルタの基本的な動作である。
4.  $\theta_l$  と  $\theta_h$  に対するIPD、IIDを推定する。これらの推定には、ロボットの伝達関数を利用する。
5. 音源方向  $\theta$  に対して、ロボットの伝達関数を利用して、入力スペクトルから以

下の条件を満たすサブバンドを選択する。

$$f < f_{th} : \Delta\varphi_E(\theta_l) \leq \Delta\varphi' \leq \Delta\rho_E(\theta_h)$$

$$f \geq f_{th} : \Delta\rho_H(\theta_l) \leq \Delta\rho' \leq \Delta\rho_H(\theta_h)$$

6. IPD  $\Delta\varphi_E(\theta)$  と IID  $\Delta\rho_H(\theta)$  は、それぞれロボットの伝達関数から推定される IPD、IID である。 $f_{th}$  は、フィルタリングの判断基準に IPD と IID のどちらを用いるかを定める閾値である。一般に、低周波数域では IPD、高周波数域では IID が大きく影響し、この閾値はマイク間距離に依存する。我々のロボットでは、理論的にも、実験的にも  $f_{th}$  として 1500Hz が妥当であることが報告されている [15]。
7. 選択されたサブバンドから、音響信号を再合成し、該当範囲にある音響信号を抽出する。

実際には、音源方向  $\theta_s$  は時間  $t$  の関数であるため、特定音源を抽出し続ける際には、時間方向の連続性を考慮する必要がある。本報告では、3 節に述べる実時間人物追跡システムから音源方向を得ることでこれを解決している。実時間人物追跡システムでは、すべての情報をストリームという時間的な流れを考慮した表現を用いて表しているため、同時に複数の音源が存在したり、音源や自分自身が移動する場合でも、一つのストリームに注目することによって、特定音源からの方向情報を連続的に得ることができる。また、ストリームは視聴覚情報を統合するためにも使用しており、これにより、視覚情報による音源定位精度向上を実現している。

### 3.12.4.2 聴覚中心窩による通過帯域制御

#### 3.12.4.2.1 聴覚中心窩とは

霊長類の視覚は、中心窩と呼ばれる解像度が高い部分が中心部に存在し、周辺部では解像度が低くなる代わりに、広範囲な視野を得ている。このような構造を用いれば、対象物を中心窩で捕らえることにより、高解像度の情報を取得することができる。つまり、広い視野と高い解像度を併せ持ち、かつ脳の情報処理量を劇的に削減できる効率的な構造を有している。

ロボットでも、同様の構造により計算量を削減できることから、中心窩を利用した視覚処理はアクティブビジョン(Active Vision)[1]の典型的な例として、しばしば利用されている [12,22]。

人間の聴覚においても、水平方向の音源定位の精度は正面方向で最も高く、周辺部に行くに従って低くなることは、古くから知られている [5]。耳に 2 つのマイクを備えたロボットによる音源定位でも、人間と同様の傾向が見られる。図 3 は、3 節で説明した 実時間人物

追跡システム[13]における3つの定位モジュール音源定位、顔定位、ステレオ物体定位による定位結果の平均値、図4は、音源定位による定位結果の分布を表している。

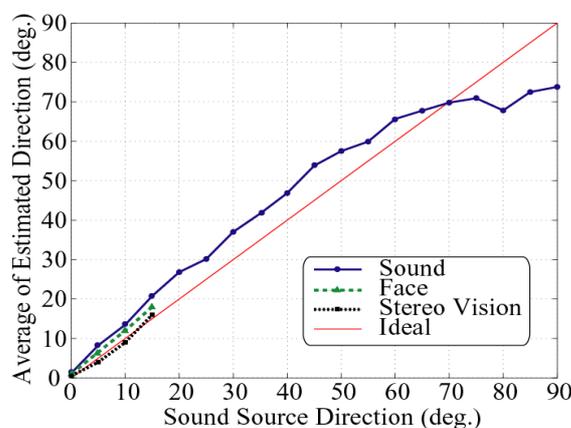


図3 音源定位結果の平均

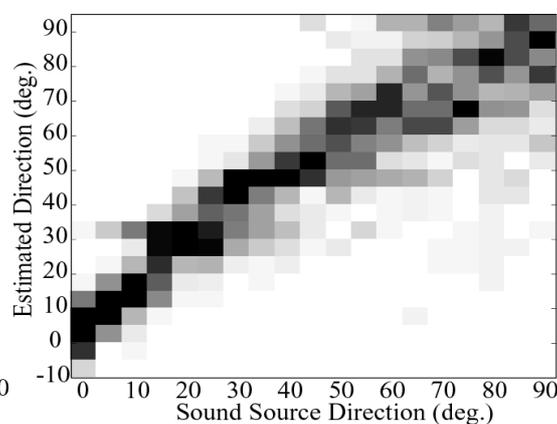


図4 音源定位結果の分布

図3から、音源定位による定位誤差は、正面方向から20°付近まで増加した後、70°付近までは6°程度で一定だが、それ以降は大きく悪化し、90°では、15°以上になる。また、図4から、正面方向のばらつきは少なく、正面から離れるにつれ、ばらつきが目立ち、分散が大きくなる。このように定位結果の平均、分散は、ともに正面方向で音源定位の精度が高くなることを示しているため、本現象をロボットにおける聴覚中心窩と呼ぶ。

なお、神経行動学(neuroethology)では、ドップラー効果によるエコー音の周波数変化を抽出するため、キクガシラコウモリの蝸牛殻で特定の周波数に対する感度が高くなっている部分を聴覚中心窩と呼んでいる[23]。選択的注意という広義の意味では、両者は似ているが、本報告では、ロボット頭部の正面方向で感度が高いという意味で聴覚中心窩という言葉を使用する。

図3では、ステレオビジョンによる定位誤差は1°、顔定位による誤差は2°程度と、聴覚処理よりも正確であることがわかる。これは、音源方向が正面に近く、視覚情報が利用できる場合には、高精度の視覚情報によって、聴覚の精度不足を補うことが可能であることを示している。

これらから、音源定位では、視覚の中心窩と同様に、聴覚中心窩を利用して音源に正対するようなアクティブな動作を行えばシステムの精度の向上が期待できる。さらに正面方向で視覚情報が利用できれば、視聴覚統合によりシステムのロバスト性を向上できると考えられる。

### 3.12.4.2.2 通過帯域制御

方向情報を利用した音源の分離抽出を考えた場合、正面方向の音源であれば、正確な音源方向を利用することができるが、音源方向が正面から離れるにつれ、方向情報に精度を期待できなくなるため、音源方向によってフィルタの通過帯域を制御する必要がある。

従来の方向通過型フィルタ[20]はスペクトルの各サブバンドで、両耳間位相差 (*Interaural Phase Difference*, IPD) と 両耳間強度差 (*Interaural Intensity Difference*, IID) に対する仮説推論を行うことによって特定方向の音を抽出するものであるが、フィルタの通過帯域が音源方向によらず一定であることが、十分な精度が得られない一因であった。

そこで、アクティブ方向通過型フィルタでは最適な通過帯域を求めるために、音源数 1 の場合に音源方向や通過帯域を様々に変化させて、抽出精度の違いを調べた。

音源には、スピーカから出力される音声信号を用いた。スピーカとロボットの距離は 1m とし、スピーカの水平方向を、ロボットの正面から、 $0^{\circ}$   $90^{\circ}$ まで  $10^{\circ}$  おきに変化させた。また、音源を抽出する際には、スピーカ方向は既知であるものとし、方向通過フィルタの通過帯域を  $\pm 5^{\circ}$   $\pm 90^{\circ}$  まで  $\pm 5^{\circ}$  単位で変化させて音源を抽出し、S/N 比による比較を行った。

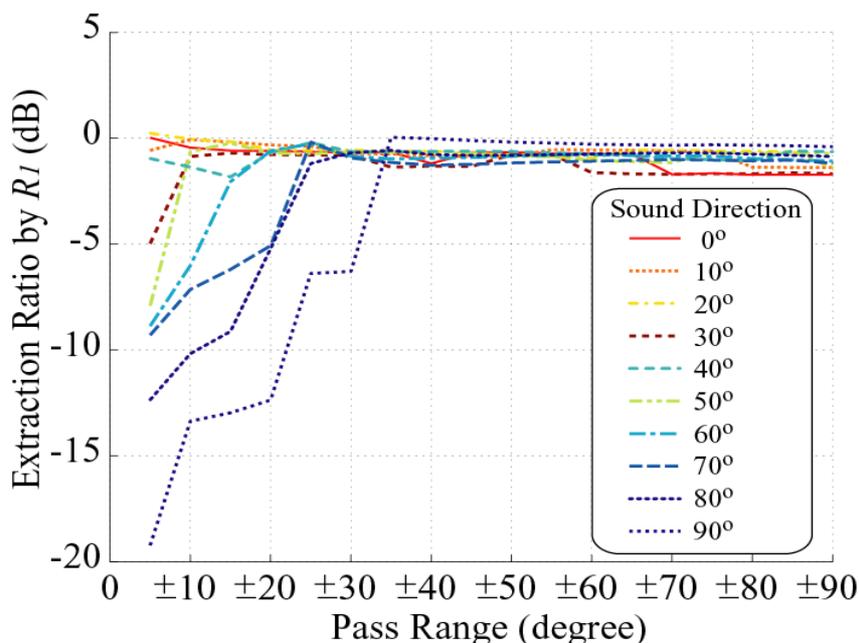


図5 単一音源の抽出

図5に結果を示す。実験では、背景雑音は無視できる程度に小さかったため、音源数が

### 3.12 アクティブオーディションを利用したヒューマンロボットインタラクションの高度化

1 の場合は、S/N 比が 0dB となった時に、元波形が完全に抽出できたと解釈する。音源方向が  $0^{\circ}$   $30^{\circ}$  と正面方向に近い場合には、通過帯域が  $\pm 10^{\circ}$  程度で元波形を抽出できているが、音源方向が正面から離れるに従い、元の波形に含まれるパワーを抽出するために、広い通過帯域を必要とし、音源方向が  $90^{\circ}$  の場合には、最低でも  $\pm 35^{\circ}$  程度の通過帯域が必要である。

音源数が 1 の場合には、通過帯域が広ければ広いほど、S/N 比の高い信号を抽出することができるが、実環境では、背景雑音を含め、複数の音源を考慮する必要があるため、なるべく通過帯域を狭くとることが望ましい。そこで、図 5 から、ほぼ元波形が抽出でき、かつ極力狭い通過帯域を音源方向ごとに抽出し、図 6 のように通過帯域関数を導出した。通過帯域は正面方向では狭く、周辺部では広がっていることがわかる。これは、音源定位と同様に、音源分離でも聴覚中心窩を利用することが可能であることを示している。

アクティブ方向通過型フィルタでは、このような通過帯域制御を行って、正面方向では S/N 比の高い音響信号を抽出し、正面方向から離れた音源に対しては帯域を広く取り、背景雑音の混入により S/N 比は多少落ちるものの、必要な情報をできるだけ抑制せずに、特定の音源の強調を行う。正面方向から離れた音源を精度よく抽出する必要がある場合は、聴覚中心窩を利用できるように、音源方向を向くような制御を行う。

実際の利用では、他の音源の音を極力抽出したくない場合、単なる音響信号の強調として利用したい場合など、状況に応じたチューニングが必要な場合もあると考えられるが、以後の実験では、図 6 に示された通過帯域関数を利用するものとする。

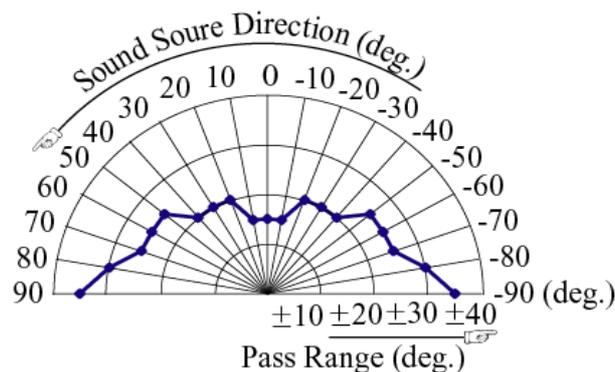


図 6 通過帯域関数

#### 3.12.4.3 ロボットの伝達関数

一般に、ロボットの伝達関数は計測によって求める。本報告では、無響室で水平方向について  $5^{\circ}$  刻みでインパルス応答を計測して得られた計測伝達関数を用いている。

しかし、部屋の音響環境の変化に動的に対応するためには、音響環境が変わるたびに部屋の伝達関数、もしくはロボットの伝達関数の再計測が必要となるなど残響や動的な音響環境の変化に追従させることが難しい。また、各方向からの測定が必要であるため、測定にも時間がかかるといった欠点を抱えている。

そこで、IPD については、水平角から計算的に IPD を推定する手法である聴覚用エピポーラ幾何(*Auditory Epipolar Geometry*)[14] を利用している。これは、ステレオビジョンで利用されるエピポーラ幾何[6]と同様の概念を 2 本のマイクによる定位に当てはめたものである。音源とロボット間の距離が 50cm 以上では無限遠の音源を仮定できるので[15]、頭部形状による影響を考慮すると、式(3)として表すことができる。

$$\Delta\phi = \frac{2\pi f}{v} \times r(\theta + \sin\theta) \quad (3)$$

ここで、 $\Delta\phi$ 、 $\theta$ 、 $f$ 、 $v$  は、それぞれ IPD、音源方向、周波数、音速を示す。また、 $r$  はロボット頭部を球形とみなした場合の半径である。

最終的に、ロボットの伝達関数として IID については、計測した伝達関数、IPD については、式(3)を用いている。IID について計測によらない手法が望まれるが、これは今後の課題である。

### 3.12.5 分離音の音声認識

音声認識の分野では、マルチコンディショニングやミッシングデータなどノイズにロバストな音声認識へのアプローチが行われている[4,21]。

しかし、これらは S/N 比が小さい場合は有効ではない。このような場合には音声認識のフロントエンドとして音源分離が必要である。また、S/N 比が大きい場合も有効である。フロントエンドとして ADPF を使用し、複数の音響モデルを使った音声認識を提案する。

#### 3.12.5.1 音響モデル

音声認識エンジンには、京大で開発された Julian[24]を利用している。本報告では、音声データは、男性 2 名、女性 1 名の計 3 名の発話による色、数字、食べ物といった 150 語を使用している。

音響モデル用の音声データとして、まず、3m×3m の部屋で、SIG から 1m の距離にスピーカを置き、その音を SIG のマイクで録音した。スピーカは、SIG から 0、±60°の位置におき、それぞれの方向について、すべてのデータを録音した。また、0、±60°の 2 箇所から同時に音声を出力する場合、3 箇所から同時に音声を出力する場合についてもすべての組合せについて録音を行った。次に、音源方向を既知として ADPF による音声抽出を行

### 3.12 アクティブオーディションを利用したヒューマンロボットインタラクションの高度化

った。抽出した音声を話者、発話方向ごとに整理し、音響モデルのトレーニングセットとした。音響モデルにはトライフォンを用い、トレーニングセットごとに、Hidden Markov Model Toolkit (HTK)を用いて作成した。したがって、本報告では、3 話者、3 方向の組合せで 9 種類の音響モデルを使用している。

#### 3.12.5.2 複数の音響モデルを利用した音声認識

音声認識では、並列に 9 つの音声認識プロセスが実行される。各音声認識は、図 7 に示されるようにそれぞれ異なる音響モデルを用いる。セレクトタはすべての音声認識結果を統合し、最も信頼性が高いと判断される結果を出力する。

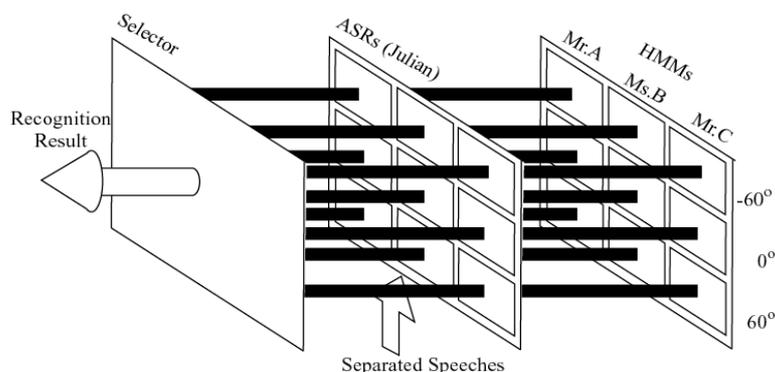


図 7 複数の音響モデルを統合した音声認識

統合のアルゴリズムを定義するために、特定話者の音響モデルに対する単語認識率を調べた。図 8 に示した結果から、話者よりも方向の違いによる認識率の低下が少ないことがわかる。また、話者も方向もあっている場合は 80% 上の認識率であることがわかる。この結果を踏まえ、音声認識の際には、音源方向は既知であることを利用し、セレクトタは式(4)に示すコスト関数を統合のために使用している。

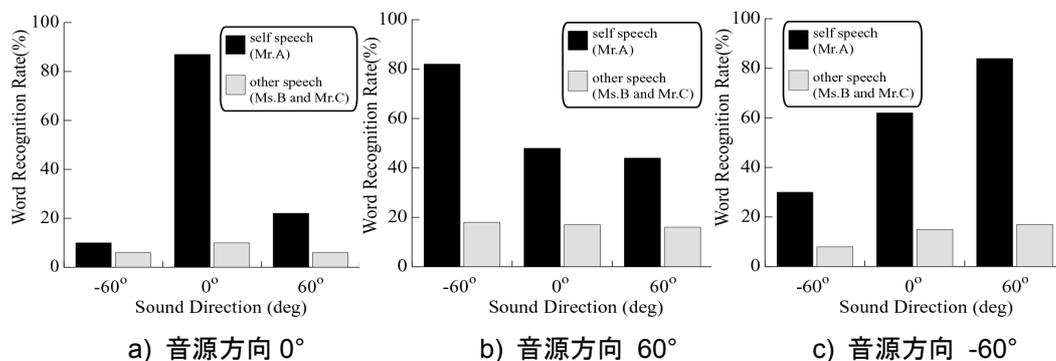


図 8 A 氏の音響モデルによる音声認識結果

$$V(p_e) = \left( \sum_d r(p_e, d) \cdot v(p_e, d) + \sum_p r(p, d_e) \cdot v(p, d_e) - r(p_e, d_e) \right) \cdot P_v(p_e) \quad (4)$$

$$v(p, d) = \begin{cases} 1 & \text{if } \text{Res}(p, d) = \text{Res}(p_e, d_e) \\ 0 & \text{if } \text{Res}(p, d) \neq \text{Res}(p_e, d_e) \end{cases}$$

ここで $r(p, d)$ 、 $\text{Res}(p, d)$ は、話者 $p$ 、方向 $d$ の音響モデルを使用した場合の単語認識率と入力音声に対する認識結果を示している。また、 $d_e$ は実時間人物追跡システムから得られた音源方向であり、 $p_e$ は、評価対象の人物である。 $P_v(p_e)$ は顔認識モジュールで生成される確率であり、顔認識ができない場合は、常に1.0となる。最終的に、セクタは最も大きな $V(p_e)$ を持つ人物 $p_e$ と認識結果 $\text{Res}(p_e, d_e)$ を出力する。

$V(p_e)$ の最大値が1.0以下もしくは、2番目に大きい値と近い場合は、SIGは認識が失敗もしくは、一つの候補に絞りきれなかったと判断して、音源方向を向き該当の人物に再度尋ねなおす。このように、複数の音響モデルを利用して、分離音と話者の認識を行う。また、顔認識が利用可能であれば、人物名がわかるためロバスト性を向上できる。

### 3.12.6 実験と評価

同時3話者発話のシナリオを通じてロボット聴覚システムを評価した。シナリオの内容を以下に示す。

1. ロボットから1mの距離に60度間隔(SIGから見て、 $0^\circ$ 、 $\pm 60^\circ$ )で3名の人間が並んでいる。
2. SIGは、3名に質問をする。
3. 各話者は3人同時に質問に対する回答を行う。
4. SIGは3話者の混合音声の定位・分離・認識を行う。
5. 最終的に、SIGは各話者に向きながら、向いた方向の人が誰で何を言ったかを答えていく。
6. 音声認識に失敗したと判断した場合は、該当話者の方向を向いた時に再び尋ねなおす。



図 9 3 話者同時発話の認識結果例

本報告では、実際の人間の代わりにスピーカとその前面に貼られた写真を用いている。スピーカは音響モデル作成時に使用したスピーカと同じものである。各スピーカから流れる音声は、そのスピーカに貼られた写真と同じ人物のものである。以下にこのシナリオの典型的な結果を 2 例示す。

1. SIG が好きな数字に関する質問をする(図 9a)。
2. 各スピーカから 1 から 10 までの互いに異なる数字が同時に流れる。ただし、数字の組合せはトレーニングセットに含まれる組合せと同じものである(図 9b)。

3. SIG は各音声から、実時間人物追跡システムを利用して定位する。定位情報を利用して ADPF がその方向の音声を抽出する。各分離音に対し 9 つの音声認識プロセスが同時に実行され、結果を統合し、最も適合のよい話者名、認識結果を求める。

4. SIG は各話者に向きながら、求めた話者名、認識結果を答える(図 9c-e)。

結果では、3 話者の認識がすべて成功しており、同時発話の場合でもロボット自身のマイクを使った音声の定位・分離・認識を行うロボット聴覚システムの有効性を示すことができた。しかし、図 8 に見られるように、各分離音声の認識率は高々 80% 程度である。音声認識に失敗する場合は、対象音源の方向を向き、聴覚中心窩をうまく利用し、曖昧性を解消するように訊き返すようなアクティブオーディションを用いて解決することができる。また、事前に顔認識によって顔の名前がわかっているときには、音声認識で使用する音響モデルの数を削減することができるので、高速で正確な認識が可能であるという結果も得られている。

### 3.12.7 結論

本報告では、ロボットへ搭載してアクティブオーディションに基づき音源の定位・分離・認識を行うロボット聴覚システムを提案・評価し、その有効性を示した。

システムは、聴覚中心窩に基づき音源方向に応じたアクティブな通過帯域制御と音源方向を向くというアクティブな動作を行い、高速で高精度な音源定位・分離・認識を実現した。これは、アクティブオーディションの有効性を示しており、ロボット聴覚では、アクティブな動作による知覚向上が本質的であることを示している。音源方向を向くという動作は、ロボット聴覚の向上だけでなく、人間とのフレンドリーなインタラクションを実現したり、トレイグジスタンスによる会議では、相手の注意を向けさせたりという意味でも重要であろう。

しかし、より制約の少ない環境での利用に耐えうる音声分離・認識には、多くの課題がある。これには、よりロバストで高精度な音源分離も必要であろうが、音声認識エンジンにも、*missing data* や *missing feature* など分離データの性質を考慮した改良[4,21]が必要であろう。

### [謝辞]

豊橋科学技術大学の中川聖一教授と京都大学の河原達也助教授の助言に感謝する。また、北野共生システムプロジェクトのメンバに感謝する。

## [参考文献]

- [1] Y. Aloimonos, I. Weiss, and A. Bandyopadhyay. Active vision. *International Journal of Computer Vision*, 1987.
- [2] M. Aoki, M. Okamoto, S. Aoki, H. Matusi, T. Sakurai, and Y. Kaneda. Sound source segregation based on estimation of incident angle of each frequency component of input signals acquired by multiple microphones. *Acoust. Sci. and Tech.*, 22(2):149-157, 2001.
- [3] F. Asano, M. Goto, K. Itou, and H. Asoh. Real-time sound source localization and separation system and its application to automatic speech recognition. *Eurospeech 2001*, pp.1013-1016. ESCA, Sep. 2001.
- [4] J. Barker, M. Cooke, and P. Green. Robust asr based on clean speech models: An evaluation of missing data techniques for connected digit recognition in noise. *EUROSPEECH 2001*, volume 1, pp. 213-216. ESCA, 2001.
- [5] J. Blauert, *Spatial Hearing*. The MIT Press, 1999.
- [6] O. D. Faugeras. *Three Dimensional Computer Vision: A Geometric Viewpoint*. The MIT Press, MA., 1993.
- [7] K. Hidai, H. Mizoguchi, K. Hiraoka, M. Tanaka, T. Shigehara, and T. Mishima. Robust face detection against brightness fluctuation and size variation. *IROS-2000*, pp. 1397-1384. IEEE, 2000.
- [8] K. Hiraoka, S. Yoshizawa, K. Hidai, M. Hamahira, H. Mizoguchi, and T. Mishima. Convergence analysis of online linear discriminant analysis. *IJCNN 2000*, pp. III-387-391. IEEE, 2000.
- [9] M. Z. Ikram and D. R. Morgan. A multiresolution approach to blind separation of speech signals in a reverberant environment. *ICASSP-2001*, pp. 2757-2760. IEEE, 2001.
- [10] S. Kagami, K. Okada, M. Inaba, and H. Inoue. Real-time 3d optical flow generation system. *MFI'99*, pp. 237-242, 1999.
- [11] K. Kashino, K. Nakadai, T. Kinoshita, and H. Tanaka. Application of bayesian probability network to music scene analysis. In *Working Notes of the IJCAI-95 Computational Auditory Scene Analysis Workshop*, pp. 52-59. AAAI, 1995.
- [12] W.N. Klarquist and A.C. Bovik. Fovea: A foveated vergent active stereo vision system for dynamic 3-dimensional scene recovery. *RA*, 14(5):755-770, October

- 1998.
- [13] K. Nakadai, K. Hidai, H. Mizoguchi, H. G. Okuno, and H. Kitano. Real-time auditory and visual multiple-object tracking for robots. IJCAI-01, pp. 1424-1432. MIT Press, 2001.
  - [14] K. Nakadai, T. Lourens, H. G. Okuno, and H. Kitano. Active audition for humanoid. AAI-2000, pp. 832-839. AAI, 2000.
  - [15] Nakadai02-AAAI, K. Nakadai, H. G. Okuno, and H. Kitano. Exploiting auditory fovea in humanoid-human interaction. AAI-2002, pp. 431-438. AAI, 2002.
  - [16] Y. Nakagawa, H. G. Okuno, and H. Kitano. Using vision to improve sound source separation. AAI-1999, pp. 768-775. AAI, 1999.
  - [17] T. Nakatani and H. G. Okuno. Harmonic sound stream segregation using localization and its application to speech stream segregation. *Speech Communication*, 27(3-4):209-222, 1999.
  - [18] T. Nishiura, M. Nakamura, A. Lee, H. Saruwatari, and K. Shikano. Talker tracking display on autonomous mobile robot with a moving microphone array. ICAD 2002.
  - [19] H.G. Okuno, S. Ikeda, and T. Nakatani. Combining independent component analysis and sound stream segregation. In *Proc. of IJCAI-99 Workshop on Computational Auditory Scene Analysis (CASA'99)*, pp. 92-98. IJCAI, 1999.
  - [20] H.G. Okuno, K. Nakadai, T. Lourens, and H. Kitano. Separating three simultaneous speeches with two microphones by integrating auditory and visual processing. *Eurospeech 2001*. ESCA, 2001.
  - [21] P. Renevey, Rolf Vetter, and Jens Kraus. Robust speech recognition using missing feature theory and vector quantization. *EUROSPEECH 2001*, volume 2, pp. 1107-1110. ESCA, 2001.
  - [22] S. Rougeaux and Y. Kuniyoshi. Robust real-time tracking on an active vision head. *IROS-1997*, pp. 873-879. IEEE, 1997.
  - [23] G. Schuller and G. Pollak. Disproportionate frequency representation in the inferior colliculus of horseshoe bats: evidence for an "acoustic fovea". In *J. Comp. Physiol. A*, volume 132, pp. 47-54, 1979.
  - [24] 鹿野 清宏, 伊藤 克亘, 河原 達也, 武田 一哉, and 山本 幹雄. 音声認識システム. オーム社, 2001

## 付属資料 海外調査報告

加藤 俊一 委員

### 1. はじめに

データベース・エキスパートシステムとその応用に関する国際会議(International Conference on Database and Expert Systems Applications - DEXA2002) は、ヨーロッパを主としたデータベース・知識ベース分野の国際会議の一つである。本年度の当 WG の海外調査では、DEXA2002 に参加し、データベース・知識ベース分野における人間中心の情報技術の研究開発の動向を調査した。

### 2. DEXA2002 の概要

第 13 回の DEXA2002 は、フランスのアクサンプロバンス(Aix en Provence)のアクスマルセイユ工科大学(Universite Aix - Marseille II, Institut Universitaire de Technologie)で、2002 年 9 月 2 日～6 日にかけて、開催された。

DEXA は、一連の DEXA イベントの形で、次の 4 つの国際会議の集合として開催・運営されている。

- (a) 第 13 回データベース・エキスパートシステムとその応用に関する国際会議(13th International Conference on Database and Expert Systems Applications - DEXA 2002) (30 セッション、92 件) (LNCS 2453 に収録)
- (b) 第 4 回データウェアハウスと知識発見に関する国際会議(4th International Conference on Data Warehousing and Knowledge Discovery - DaWaK 2002) (9 セッション、32 件) (LNCS 2454 に収録)
- (c) 第 3 回電子商取引と Web 技術に関する国際会議(3rd International Conference on Electronic Commerce and Web Technologies - EC-Web 2002) (15 セッション、42 件) (LNCS 2455 に収録)
- (d) 第 1 回電子政府に関する国際会議(1st International Conference on Electronic Government - EGOV 2002 -) (13 セッション、78 件) (LNCS 2456 に収録)

また、今回の DEXA には計 13 ものワークショップも含まれた。

- (a) Fifth International Workshop on Network-Based Information Systems - NBIS'2002 (12 件)
- (b) First International Workshop on Web Semantics - WebS 2002 (9 件)
- (c) Third International Workshop on Theory and Applications of Knowledge Management - TAKMA 2002 (11 件)

- (d) Third International Workshop on Natural Language and Information Systems  
- NLIS 2002 (11 件)
- (e) Second International Workshop on Web Based Collaboration - WBC 2002 (9 件)
- (f) Third International Workshop on Management of Information on the Web –  
MIW 2002 (11 件)
- (g) International Workshop on Trust and Privacy in Digital Business - TrustBus  
2002 (20 件)
- (h) International Workshop on Presenting and Exploring Heritage on the Web -  
PEH'02 (9 件)
- (i) Third International Workshop on Industrial Applications of Holonic and  
Multi-Agent Systems - HoloMAS 2002 (13 件)
- (k) Third International Workshop on Negotiations in electronic markets - beyond  
price discovery - e-Negotiations 2002 (6 件)
- (l) Fifth International Workshop on Mobility in Databases and Distributed  
Systems - MDDS 2002 (12 件)
- (m) International Workshop on Very Large Data Warehouses  
- VLDWH 2002 (11 件)
- (n) Second International Workshop on Electronic Business Hubs  
- WEBH 2002 (8 件)

論文発表は全体で 386 件、参加者数は 400 名程度の規模であった。データベース・知識ベース分野が主体の国際会議であるが、会議・ワークショップの名称からもわかるように、web 上でのビジネス・電子商取引、および、電子政府関係の研究発表がかなり多かった。会議の規模としてはこじんまりとした印象であったが、電子商取引関係、電子政府関係のセッションは、データベース・知識ベースの会議に匹敵するくらい（あるいはそれ以上）の熱気にあふれていた。

DEXA は、1990 年を第 1 回に今回で 13 回目を迎える。筆者は、たまたま第 1 回会議に参加し「画像の感性検索システム」について発表した経験があるので、会議の規模の拡大、データベース・知識ベース分野の応用の広がり、人間中心の情報技術との関わりなど、様々な点で興味深かった。

### 3. DEXA2002

データベース・知識ベースのシステム開発的な論文が多く、データベース・知識ベースに軸足を置きながら、web 応用や XML 技術、情報検索・情報提供サービス関係の話題も

多くカバーしていた。話題別の発表件数は以下の通り。

Web 応用 .....	6 件
XML .....	4 件
ワークフロー .....	4 件
データウェアハウス、データマイニング .....	7 件
知識工学 .....	7 件
分散システム .....	6 件
アドバンストデータベース .....	6 件
質問処理 .....	13 件
応用 .....	14 件
情報検索 .....	9 件
索引構成法 .....	7 件
その他 .....	9 件

冒頭の Prof. S. Abiteboul (INRIA & Xyleme)による招待講演 "Issues in Monitoring Web Data" では、今日、最も普及した分散データベースシステムの形態である Web データベースについて、電子商取引やマイニングに有効利用するために、どのような観点からモニタリングを行えばよいか、統計的な手法に対する批評も含めて、示唆にとんだ意見が紹介された。

人間中心の情報環境の観点からは、以下のような発表が興味深かった。

**(1) "MWM: retrieval and Automatic Presentation of Manual Data for Mobile Terminals"**

M. Shikata, et al. (神戸大学)

ビデオ映像を含む操作マニュアルを、携帯電話上から検索・表示できるようにしたシステムの試作評価を行っている。有限オートマト的に質問内容の推移をモデル化し、いわゆるベビーフェイスのモニタ、少ないボタン数という貧弱なユーザインタフェースでも、操作性と検索しやすさを両立させることを図っている。

**(2) "Web Information Retrieval Based on the Localness Degree"**

C. Matsumoto, et al. (神戸大学)

Web ページに記述された内容に基づいた web ページ間の関係の構造化、検索技術の試みである。web ページに記述している内容と特定の地域・場所との関わりに注目して、web ページを分類したり、ページ間の関係を抽象化して、他の地域を話題にした検索にも利用するなど、今後の展開が期待されるアイデアが示されていた。

### (3) "Temporal Pattern Mining of Moving Objects for Location-Based Service"

J. D. Chung (Chungbuk National University, Korea)

モバイル端末を持つ利用者に、その位置情報に基づいた情報提供サービスを行うものである。位置に対して固定的な情報を提供するのではなく、利用者の移動するパターンなどを検出して、過去の行動履歴から利用者の意図などを推定し、それにマッチした情報提供サービスを行うものである。

### (4) "A System for Retrieval and Digest Creation of Video Data Based on Geographic Objects"

T. Ueda (奈良先端大)

身に付けたカメラからキャプチャーした画像をキーとして、利用者に情報提供サービスを行うシステムの試作評価を行っている。GPS とジャイロを併用することにより、およその位置と方向を計測し、その方向に見えるであろうオブジェクトとのマッチングを行って対象を認識する点がミソである。

### (5) "Similarity Image Retrieval System Using Hierarchical Classification"

M. Tada, et al. (中央大学)

筆者らのグループの発表。局所的な高次のコントラスト特徴を採用し、風景写真やテキストなど多様な画像を例示画により類似検索するシステムの試作評価。類似尺度を、サンプル画像の段階的・階層的な分類による教示学習で、個人ごと、コンテンツごとに動的に構成できる。

## 4. ECWeb2002

Web 技術の典型的でもっとも「熱い」応用例として、Web 上でのショッピングや旅行代理店関連の情報提供サービス、オークションなどのエージェント技術との融合、セキュリティとプライバシー保護など、DEXA 本体の会議よりも活発な質疑応答も見られた。話題別の発表件数は以下の通り。

オークション技術	4 件
応用	3 件
アーキテクチャ	3 件
エージェント技術	4 件
Web 質問処理	3 件
Web データ統合	2 件
セキュリティ、プライバシー	8 件
レコメンデーション	3 件

ビジネスモデル .....	2 件
電子決済 .....	3 件
その他 .....	7 件

人間中心の情報環境の観点からは、以下のような発表が興味深かった。

**(1) "Strategies and Behaviours of Agents in Multi-phased Negotiations"**

S. Akine (Universite Paris 6, France)

MNP-ネゴシエーション問題 (m人の買い手、n人の売り手、p個の製品) を2相の複合的ネゴシエーション過程としてモデル化し、売り手・買い手のエージェントにより実現するアルゴリズムを提案している。売り手・買い手の戦略の立て方などで、カスタマイズが出来れば、ある種のキャラクターを持ったネゴシエータを実現できるように思われる。

**(2) "Extending Decision Making in Tourism Information System"**

F. Puhretmair, et al. (University Linz, Austria)

Tourism Information System (TIS)は、今回の ECWeb で最も注目を集めた発表の一つである。旅行代理店業務での顧客管理から旅行案内、チケット予約までを、一体化したシステムを試作評価している。顧客からの質問を分析し、どのような旅行を希望しているのかを推定し、また、事例推論により、そのような旅行に必要な種々のサービス (飛行機や列車のチケット、ホテル予約など) を統合的に企画し、グラフィック表示や仮想現実空間により利用者に旅行先の案内を行うなどができる。利用者の希望の推定や事例推論の仕組みなどは、まだまだ固定的でカスタマイズ機能は実現されていないが、非常に興味深い試みである。

**(3) "Homogenous EDI between Heterogenous Web-Based Tourism Information Systems"**

W. Woss (University Linz, Austria)

上述の TIS グループのシステムのアーキテクチャの紹介。旅行代理店業務での (場合によっては別の会社の) 種々の業務を Web 上で統合する仕組みについて説明している。

**(4) "User Preference Mining through Collaborative Filtering and Content based Filtering in Recommender System"**

S. Ko (Inha University, Korea)

個々の利用者がよくアクセスする Web ページ中の言葉の出現頻度などの統計量から、その利用者が関心を持つ話題・商品などの推定が出来るようなモデル化を行う。同じようなユーザモデルの利用者群の購買履歴なども利用 (協調フィルタリング) して、個々の利用者に適切な商品のレコメンデーションを行う。この際、ユーザをどのようにグループ分けして、いわばステレオタイプなユーザのモデルを作るかがポイントである。

## **(5) "An Improved Recommendation Algorithm in Collaborative Filtering"**

T-H Kim (Yonsei University, Korea)

協調フィルタリングを効果的に行うためには、ある消費者と嗜好の良く似た（ということは、検索・購買履歴などのよく似た）別の消費者を発見する必要がある。グラフ上でのクラスタリングにより、このような消費者の集合を発見するアルゴリズムを提案している。

## **(6) "Series of Dynamic Targeted Recommendation"**

N. Modani (IBM India)

Web ページの広告領域にどのような広告を表示すべきかを、そのセッションで利用者がアクセスしている web ページ上の言葉などの関連性から動的に求める仕組みを提案している。特に面白いのは、特定利用者の長期的なアクセスログをあえて使わず、毎回のセッションの中で、毎回の web ページの移動のたびに、どのような話題に興味を持っているかの推定を計算しなおす点である。Modani らの意見では、広告の内容はすぐに古くなるし、クーポンなどのサービスも 1 回きりであるなど、長期的なモデルを構築して利用する必要も無いし、個人の特定の必要も無いと割り切って考えている。

## **5. まとめ**

筆者が 12 年前に参加した第 1 回 DEXA90 の当時と比べて、情報通信環境は革命的に変化した。当時は、インターネット上でのアプリケーションサービスの実現自身が珍しく、また、インターネット上のコンテンツがそのままデータベースとして利用されるようになるという意識は希薄であった。今日では、インターネット・モバイルネット・ユビキタスネットが融合しはじめる時期にさしかかっており、「消費者」としての利用者が情報通信環境に対して期待するサービスも大きく変化してきている。

そのような状況にあって、「個」としての消費者をどのようにとらえ、サービスを提供していくのかは、今後の高度情報化社会（＝複合的な情報通信環境）において IT 系企業にとって生き残りをかけた取り組みになると言えよう。

そのような観点から、今回の出張で見聞した様々な研究開発を見ると、データベース技術は、まだまだ「マス」としての消費者のパラダイムから脱却していないように思われる。データベース技術の研究開発は、データ構造を標準化し、利用者を標準化し、利用のプロトコルを標準化することで、大きな発展を遂げてきた。しかし今や、個別化・流動化・その場限りなど、対極的なシステム技術が必要となってきたのではないかと。基盤技術としてのデータベース技術の重要性は不変なので、感性情報処理技術など「個」を指向した情報技術との効果的な融合化を進めていくことが、情報技術にとってもビジネスにとっても、重要課題の一つとして認識されるべきであろう。





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VI

© 平成 15 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 2 丁目大門ビルディング 4 階

TEL(03)3456-2511