

昭和63年度

情報機器のパーソナル化の技術動向に関する
調査研究資料

平成元年3月

財団法人 日本情報処理開発協会
財団法人 パーソナル情報環境協会



技術動向調査専門委員会 名簿

委員長	田村浩一郎	電子技術総合研究所 情報科学部長
委員	石川 真澄	電子技術総合研究所 情報科学部主任研究官
委員	守谷 伸行	沖電気工業㈱ 技術開発本部総合システム研究所研究主任
委員	福田 譲治	ソニー㈱ 総合研究所 AV 技術研究部門グループ長
委員	戸田 明	大日本印刷㈱ 研究開発推進本部副課長
委員	菊池 紀芳	㈱東芝 情報通信システム技術研究所開発主務
委員	塚原 祐輔	凸版印刷㈱ 総合研究所基盤研究センタ副主任研究員
委員	加藤 至朗	日本アイ・ビー・エム㈱ 大和研究所副主幹開発計画担当部員
委員	後藤 富雄	日本電気㈱ パーソナルコンピュータ開発本部商品企画部課長
委員	森 秀司	㈱日立製作所 情報事業本部第1マーケティング部主任技師
委員	上田 道夫	富士ゼロックス㈱ 技術企画部担当係長
委員	牧村 信之	富士通㈱ 情報システム事業本部ソフトウェア開発推進統括部 課長
委員	吉村 哲也	松下電器産業㈱ 東京研究所
委員	南部 元	三菱電器㈱ 産業システム研究所画像システム開発部 次長
事務局	瀬野 健治	㈱パーソナル情報環境協会 技術部長
事務局	宮田 達夫	㈱パーソナル情報環境協会 調査課長

目 次

1. 情報機器のパーソナル化とインタフェース	1
1.1 情報と情報機器	1
1.2 現世代のヒューマンインタフェース	2
1.3 次世代ヒューマンインタフェースにたいする要請と技術的シーズ	3
2. ヒューマンインタフェース・アーキテクチャ	6
2.1 はじめに	6
2.2 パイプライン・モデル	7
2.3 PDP モデル	9
2.4 黒板モデル	10
2.5 おわりに	12
3. ヒューマンインタフェースにおける物理的入出力	13
3.1 入力手段	13
3.1.1 キーボード	13
3.1.2 マウス	17
3.1.3 その他のポインティング装置	20
3.2 インタフェース媒体	23
3.2.1 コマンド言語	23
3.2.2 メニュー	23
3.2.3 アイコン	24
3.2.4 ウィンドウ	25
3.3 出力（メッセージ、音、音声、画像）	27
3.3.1 メッセージ、音、音声	27
3.3.1.1 はじめに	27
3.3.1.2 メッセージによるコンピュータインタフェースの研究	27
3.3.1.3 音あるいは音声によるコンピュータインタフェースの研究	32
3.3.1.4 ヒューマンインタフェースの評価ガイドライン	34
3.3.1.5 ヒューマンインタフェースの設計ガイドライン	36
3.3.1.6 おわりに	37
3.3.2 画像	38
3.4 マニュアル（印刷物、オンライン）	46
3.4.1 マニュアルの効用	46
3.4.2 日本におけるマニュアルの歴史	46
3.4.3 編集の新しい価値	47
3.4.4 デスクトップ・パブリッシング（DTP）	48

3.4.5	ハイパーメディアの発展	50
3.4.6	日本におけるハイパーメディアの開発状況	53
3.4.7	米国におけるハイパーメディアの開発状況	53
3.4.8	ユーザー・インタフェース	54
3.4.9	ユーザー・インタフェースの統一の例	56
3.4.10	オンライン・インフォメーションの要件	56
3.4.11	将来のオンライン・インフォメーションの要件	57
3.4.12	本当に必要なもの	57
4.	ヒューマンインタフェースにおけるユーザモデルと認知心理学的アプローチ	59
4.1	はじめに	59
4.2	ユーザモデルの分類・特徴	61
4.3	情報処理モデル	62
4.4	ルールベース形モデル	64
4.4.1	ACT*モデル	64
4.4.2	三段論法における思考のモデル	64
4.5	コネクショニスト形モデル	66
4.5.1	タイピングのモデル	66
4.5.2	タイプミス訂正モデル	66
4.5.3	コマンド系列中の特徴抽出モデル	67
4.6	認知心理学的アプローチ	69
4.7	おわりに	70
5.	ヒューマンインタフェースにおけるメンタルモデル	72
5.1	はじめに - 要約に替えて -	72
5.2	メンタルモデルに関する6つの論文	74
5.2.1	認知科学におけるメンタルモデル	74
5.2.2	Some Observation on Mental Models	75
5.2.3	The black box inside the glass box : presenting computing concepts to novices	76
5.2.4	Mental Models : Aggregation, Abstraction, and Analogy	76
5.2.5	Instruction manipulation of users' mental models for electronic calculators	81
5.2.6	The machine inside the mashine : users' models of pocket calculators	83
6.	ヒューマンインタフェースにおける問題解決支援	86
6.1	個人レベルでの問題解決支援	86
6.1.1	システムを理解するためのヒューマンインタフェース	86
6.1.2	気くばりのヒューマンインタフェース	90
6.1.3	知的ヒューマンインタフェース	92
6.2	集団レベルでの問題解決支援	95

6.2.1	はじめに	95
6.2.2	システム事例	95
6.2.2.1	会議での問題解決支援システム：黒板を越えて	95
6.2.2.2	電子会議システム	98
6.2.2.3	「半構造メッセージ」は、連絡調整のために大変役立つ	101
6.2.3	ヒューマンインタフェースの問題	105
6.2.3.1	なぜ、CSCW システムは失敗するのか？	105
6.2.3.2	組織に関する研究上の視点と人間工学	106
6.2.4	今後の予感	108
6.2.4.1	グループ通信：社会システムへの影響と課題	108
6.2.5	おわりに	109
7.	ヒューマンインタフェースにおける関連研究	111
7.1.	音声入出力技術の動向	111
7.1.1	はじめに	111
7.1.2	音声認識の基本技術	111
7.1.2.1	音響分析	112
7.1.2.2	認識	114
7.1.3	音声理解システム	117
7.1.3.1	CMU と ATR のシステム	117
7.1.3.2	SPHINX システム	119
7.1.4	音声合成	123
7.1.4.1	DECTalk	123
7.1.4.2	NTT の日本語テキスト音声合成システム	125
7.2.	画像処理技術	128
7.2.1	Recent Progress in Recognition of Objects from Range Data	129
7.2.2	Chinese Character Recognition : A Twenty-Five-Year Retrospective	130
7.2.3	顔面像計測に基づく視線関知法	130
7.2.4	Physically Based Modeling for Vision and Graphics	131
7.2.5	Flocks, Herds, and Schools : A Distributed Behavioral Model	132
7.2.6	Using Dynamic Analysis for Realistic Animation of Articulated Bodies	133
7.2.7	自然言語を解読し動画を自動生成するアニメーション・システム	134
7.2.8	画像データベース向き、2階層符号化方式	135
7.2.9	顔面像の分析合成符号化方式の動向	136
7.2.10	画像と音声の知的インタラクティブ符号化の一検討	137
7.2.11	ハイパーメディアシステム構成のためのデータ構造とその応用 に関する一考察	138
7.3.	自然言語	140
7.3.1	はじめに	140
7.3.2	自然言語理解の仕組み	140

7.3.3	文法理解	141
7.3.4	意味論	145
7.3.5	曖昧性の解決	148
7.3.6	あとがき	150

1. 情報機器のパーソナル化とヒューマンインタフェース

1.1. 情報と情報機器

情報とは、観測者から見たとき、解釈するものと解釈されるものとの不即不離の組である[1]。普通、解釈されるもののみを指して情報と呼ぶことが多いが、解釈者を同時に考慮して初めて、情報の存在意義が生じる。解釈者が異なれば、同じ記号が全く異なる意味を持つ。解釈者を想定しない信号は単なる物理現象に過ぎない。また、解釈されるものを考慮せずに解釈者を特定して議論しても意味がない。しかも、解釈者は必然的に社会的存在、即ち、おかれた環境にさらされ、影響される存在である。即ち、自分が解釈者とされる情報によって、内的変化を起こす。ここでいう解釈者は解釈されるものを解釈するであろうと想定されるものであって、実際に物理的に情報を扱うものである必要はない。しばしば物理的に情報を受け取るものは、情報の観測者である。

情報機器は、何かしらの形で情報を扱う機器である。情報は作られ、処理され、配付され、蓄積される。情報は必ず物理的な媒体を必要とする。情報機器はその物理的な媒体を処理する機械である。一番古い人工の情報媒体は貨幣であろう。一円玉は、解釈者が適切な場合、一円分のリソース（ものであれ、サービスであれ）を自由にできるという情報を運ぶ。かくて自動販売機は情報機器の側面を持つ。

コンピュータの出現は、処理対象としての情報の存在を強く社会に印象づけることになった。情報の処理とは、それを受けた解釈者が別の解釈者へある目的をもって変換することである。情報を作るということも実はその一種であるに過ぎない（例えば、編集作業）。ともかく、今や、コンピュータはあらゆる機器に忍び込み、従って、あらゆる機器が情報機器になりつつある。情報機器のヒューマンインタフェースを考えると、従って、情報を扱う側面から見たあらゆる機器のヒューマンインタフェースを考える、ということになる。即ち、人が機器を使用するとき、情報のやりとりによって、機器が本来果たすべき機能を実現させるという過程を支援することが、「情報機器のヒューマンインタフェース」を考えるということである。人は機器を解釈者として情報を送る。機器は人を解釈者として情報を送り返す。人にとってこのやりとりはまさに知的活動である。この活動を円滑に進めるための機械側の工夫がメンタルなヒューマンインタフェースである。コンピュータのヒューマンインタフェースはその典型ではあるが一例に過ぎない。以下ではメンタルなヒューマンインタフェースのみを扱うことにする。

コンピュータの歴史の中で、ヒューマンインタフェースがまるでなかった時代ではなく、それなりの変遷をとがっている。確実にいえることは、全体のリソース（ハードウェアとソフトウェア）の中で占める割合を高めてきていることである。リソースが増えるとその多くがヒューマンインタフェースに費やされるのである。また、当然ながら、時代のハードウェア技術によってその基本概念そのものが影響される。さらには、技術が社会に影響すると同時に、影響された社会から技術が逆に影響され、その意味でヒューマンインタフェースの役割が大きく変化してきている。

1.2 現世代のヒューマンインタフェース

フレンド21のプロジェクトは情報機器の次世代のヒューマンインタフェースを目指すものである。次世代を語るためには現世代を位置づける必要がある。技術の内容から見て、現世代を1980年代後半から1990年代にかけてポピュラーになっているものとして切り出すのが適切であろう。すると、次世代はそれから質的な転換をとげて2001年前後以降に一般化するものであると想定することになる。しかし、世代の交代はすべてを完全に一新するわけではなく、旧世代の中にみられる萌芽が大きく成長して新世代の主潮流となるものである。

現世代として定着しつつある“Look and Feel”を特徴付けると以下のようになる。

まず、歴史的な流れを極くおおざっぱに見ておく。多くの人が認めるように、その源はXEROX社のパロアルト研究所（PARC）の1970年代における活動にある。片やSRIから来たEngelbartのグループがあり、片や大学院を卒業したばかりのAllan Kayのグループがあり、PARCのヒューマンインタフェースにおける成果は両者の合体によるものとされる[2]。前者は、コンピュータのエキスパートがいかに効率良くコンピュータを使用できるかという観点で、後者は逆に子供が使えるものを目指して、ヒューマンインタフェースを追及した。前者はマルチウィンドウ、マウス、ハイパーテキストを生み、後者はオブジェクト指向、アイコン、マルチメディア、メタファを生んだ。両者に共通するハードウェアに対する想定は、グラフィックディスプレイの持つ二次元性の重視であり、パーソナルな使用である。即ち、テレタイプなどの一次元文字列入出力を使うTSSからの脱皮である。XEROX社そのものは商品化に成功したとはいえなかったが、Apple社のMacintoshにその技術の大半が流れ込み、広く一般に普及することになった。これらの技術はワークステーション、パーソナルコンピュータのみならず、スーパーコンピュータやメインフレームにさえも導入されつつある。特に、現在提案されている各種のUNIX標準シェルのほとんどすべてが同様のLook and Feelを持っている。

これらの技術の背景にある概念として重要なものが「直接操作感」である。人が機器を使うのは、機器の持つ機能を利用したいからである。即ち、ユーザが機器に発する「情報」は、解釈者を、「機能の世界」とするのが自然である。ところがしばしば手段であるべき機器の世界が解釈者として想定される。しかし、人はたとえば帳簿を付けたいためにコンピュータを使用するのであって、OSコマンドを覚えたり特別にプログラミングをしなければならないのは多くの人にとって苦痛でしかない。余計な介在物なしに直接帳簿付けの仕事をしたいのである。少なくともその気分になりたい。そこで、メタファが利用され、メタファを提示するアイコンなどの表現が工夫され、さらにマウスなど、ポインティングと操作の道具が使われる。しかも、操作の感覚を補強するため、操作の結果を即座に提示する。これらは、ユーザに、機能の世界（例えば、帳簿の世界）を直接操作している感覚を与える。ユーザは自分の操作の結果を予想しやすくまた確認しやすいので、操作に確認を持つことができる。このとき間を介在するコンピュータは自分自身を隠して「透明に」なる。ユーザから発せられる情報は機器の本来の機能の世界を受取側（解釈者）と想定して作られるのであって、機器そのものではない。これが使いやすさを本質的に生むものになる。

こういった技術の多くは世代が変わっても引き継がれることになるだろう。従って、次世代ヒューマンインタフェースを述べるときにはこれらの技術を含むことを前提とし、わざわざ言及しない。

1.3 次世代ヒューマンインタフェースに対する要請と技術的シース

メンタルなヒューマンインタフェースが重視される最大の要因は、コンピュータを初めとする情報機器のユーザ数の増大である。互いに他方の原因となって、ヒューマンインタフェースが向上し、ユーザ数が増大する。旧世代においては、百人に一人以下がコンピュータを利用していたが、現世代では十人に一人、次世代では、十人のほとんどすべてが利用することになる。多くの人にとってそれだけ機器とのメンタルな接触が多くなるということである。

ここでいうコンピュータには注釈が必要である。ファミコンもコンピュータならば、家電製品に組み込まれたマイクロコンピュータもコンピュータである。そうだとすれば、すでに、誰もが何十台もコンピュータをつかっていることになる。しかし、ここでいうのは、情報を先の意味で扱い、一定の汎用性を持つものとしている。現在のパーソナルコンピュータをやや拡大した感覚でつかっている。

あらゆる機器が本来パーソナルなもの、即ち個人専用の持ち物になることを指向する。そうならないのはそうできない経済的、物理的事情があるからである。次世代の時代にはコンピュータは確実にパーソナルなものとなる。十人のすべてがコンピュータを利用するということは、十人すべてが自分専用のものを持つということである。その時コンピュータは、誕生の時からひそかに仕込まれていた役割、即ち人の知的活動の増幅という役割を初めて十分に発揮できる段階にはいる。ヒューマンインタフェースはこの役割に奉仕すべきものである。

以上の観点から、次世代ヒューマンインタフェースに対する要請を見ていくことにしよう。

(1) パーソナル化

ユーザが発する情報はまずヒューマンインタフェースが受け取り、解釈する。解釈するには基礎データが必要である。パーソナル化が進むということは、個人の使用環境の履歴をエピソード的な記憶として保持することを可能とする。このデータを活かす方法は未知である。これは極端としても、ユーザの個人情報をふんだんに蓄積し、それに基づいて適切な解釈を行う方法が開拓されよう。初歩的なところでは、個人のスケジュール、住所録、それにかな漢字変換の辞書などすでにそのきざしとしてあらわれている。

個人環境の蓄積は、機械の個人適応を可能とする。ユーザから送られた情報の解釈がより適切にユーザの意図に添ったものになっていくということである。進化すると言ってもよい。例えば、音声認識や手書き文字認識にはそのための手法が工夫されてきている。これに限らず、一般的に、ユーザにメンタルになじんでいく、といった感覚をもたらすものになるだろう。

どんな設計者もユーザのあらゆる要望や使い方をあらかじめ予測することはできない。したがって、ユーザの意思表示をもっと直接的に受け入れる方法を用意することも重要である。現在のコンピュータにおけるUNIXのシェルやパソコンの環境パラメータの設定などがその初歩的形態になっている。ヒューマンインタフェースの構造がよりわかりやすくなることによって、ユーザがその組み合わせを用意に変更できるようになれば、より自由に自分好みのヒューマンインタフェースを持つことができる。

ヒューマンインタフェースがこのように個人化するということは、衣服のファッションのように、多様化することを意味する。ヒューマンインタフェースの標準化は困難である。ここで個人間の情報の流通やソフトウェアとハードウェアの市場流通性をどう確保するかという問題が生じる。当然、内部構造における整合性を保持しなければならない。ここにヒューマンインタフェースのアーキテクチャの問題が生じる。

(2) 能動性

現世代のヒューマンインタフェースは、旧世代までのものが余りにも使いにくかったという反省の上に立っている。ここまでできれば、使いやすさはもう十分だという意見も強いが、まだ不足だという見方もある。いずれにせよ、使いにくいものを使いやすくするという受け身の工夫が基本となっている。次の段階を目指すとするれば、より積極的に人と機械の情報のやりとりに介入し支援するものとしてのヒューマンインタフェースを考えてみる必要がある。これが能動的ヒューマンインタフェースである。一種のインテリジェント性ということもできよう [3]。

インテリジェント性を発揮するためには相手にしているユーザについてのデータを知的に利用する仕組みが必要である。何をもって知的とするかは大問題であるが、とりあえず、人工知能などの研究によって出てきた各種技法を利用できよう。

ヒューマンインタフェースをひとつのモジュールとしたとき、それは内部のサブモジュールで構成される。当報告書では、エージェントと呼ばれるものがそれに相当する。それらもまた、何かしらの形でインテリジェントでなければ全体としてのインテリジェント性を発揮できない。こうして、インテリジェント性を真に実現するにはフラクタル性を持つことになる。それによって、柔軟なアーキテクチャもまた実現できることになる。また、各モジュールはヒューマンインタフェースの役割を実現するという「機能」を持つと同時に、ユーザに対してわかりやすくするというヒューマンインタフェースの側面を持ち、この意味でもフラクタル性が必要となる。このフラクタル性を具現化するためには、モジュール構成、即ちアーキテクチャの在り方に対する深い考察が必要となる。極めて困難な課題である。

(3) 可能性

以上の要請をみたそうとすれば必然的に可搬性が問われる。個人環境のすべてを持つためにはユーザは絶えずそれを持ち歩けなければならない。また、ユーザの発する情報に知的に解釈する装置はユーザの多くの活動が絶えず必要とするものになる。そこまで行かずとも、アポイントを決めたり、電話をかけたりするために、すでにポータビリティは必須条件となりつつある。

ハードウェアの進歩から見て、紙の薄さ、重さ、大きさ、分解能を持つ情報機器の実現性がある。少なくとも、多くの技術がそれを指向しだしている。ルーズリーフのようにそれらを綴じて使用することも可能であろう。

逆に、この可搬性を前提として現世代を越える次世代ヒューマンインタフェースの基本概念を導出することもできよう。現世代はどちらかといえばデスクトップを前提としているからである。

ディスプレイパネル、不揮発性メモリ、電源としての電池、入力装置など、難問が多い。にもかかわらず、強いニーズがあり、それが技術を引き上げていくものと思われる。不揮発性メモリには強誘電体 [4]、入力装置にはデータグローブなどが注目される。

(4) 機能からの分離、独立

一人のユーザから見たとき、ヒューマンインタフェースは各種機能にたいして共通であって欲しい、例えば、同じコンピュータをつかっているのに、応用ソフトが違えば毎にヒューマンインタフェースが異なるのは極めて使いにくい。また、各種機能を実現するときそれぞれでヒューマンインタフェースを工夫し盛り込むのは経済的でない。その意味でも共通化が望ましい。共通化することは、ヒューマンインタフェースの機能から分離、独立を意味する。

機能は文明であり、汎用性を持ち、効率が重視される。一方、ヒューマンインタフェースは文化であり、個別性が強く、効率よりも効果、即ち余裕が問われる。

単に分離、独立するだけでなく、各種機能を統合する機能をヒューマンインタフェースに求めたい。即ち、ユーザが各種機能を工夫しながら使い分けるのではなく、必要なときに必要な機能が引き出されて動いて欲しい。全体の有機的な脈絡をヒューマンインタフェースが管理するのである。ハイパーカードを応用ソフトのプラットフォームとしてみたとき、素朴ながらその機能を持つことに気がつく。

この分離、独立性を進むと、ヒューマンインタフェースそのものの機械ができることになる。この機械を仮にフレンドマシンと名付けよう。フレンドマシンは、あらゆる機器にたいするヒューマンインタフェースとなる。ユーザはフレンドマシンを通してあらゆる機器にアクセスする。ユーザと他の機器との間にあって会話を円滑に進めることだけがフレンドマシンの役目である。ユーザ個人の生活環境に関するあらゆる情報が蓄積され、それが機器との会話に役立たせられる。フレンドマシンは一種のメタ機械である。こうなるためには、逆に、可搬性、能動性、パーソナル化の要請が満たされることが必須となる。

注釈と文献

- [1] この議論は、主体と客体との不即不離を主張するヴィトゲンシュタイン、オルガ、ハイデッガなどの思想に倣っている。この観点に立ってコンピュータ技術を見るものとして、
Terry Winograd and Fernand Flores: "Understanding Computers and Cognition-A New Foundation for Design", Ablex Publishing Co. 1986
が参考になる。また、認知心理学の立場からは、
Donald Norman: "Psychology of Everyday Thing", 1987
が同様の観点から、優れた、しかもわかりやすいヒューマンインタフェース論を展開している。
- [2] Stuart K. Card (XEROX PARC) による。
- [3] 所謂 intelligent interface と呼ばれるものである。最近の研究動向については、
Architectures for Intelligent Interfaces: Elements and Prototypes.
という1988年3月に開かれたシンポジウムに詳しい。Proceeding が発刊されている。
- [4] Bernard C. Cole: "How the U. S. is leading the way in strategic nonvolatile technology", Electronics, March 1989, p. 80.

2. ヒューマンインタフェース・アーキテクチャ

2.1 はじめに

過去のコンピュータ開発において、ヒューマンインタフェースは数々の改良を加えられ、その結果として不十分な点が多いとはいえコンピュータとの対話が飛躍的に容易になった。しかしこれらのヒューマンインタフェースは、各コンピュータ開発の過程で最も好ましいと思われるのがアドホック的に設計されたものであり、ヒューマンインタフェースとはどうあるべきかの議論を通して基本的なアーキテクチャを明確にし、その上で設計されたものではなかった。従って移植性やユーザから見た使い勝手といった観点からは極めて不十分であり、その基本アーキテクチャの部分で十分な検討がなされていないといえる。

ここでは、ユーザとコンピュータとの対話において、今後のパーソナル・コンピュータを設計する際にも共通的に使用できるヒューマンインタフェースの基本アーキテクチャを構築するために参考となるような研究を調査した。

調査した結果、いくつかのアーキテクチャの報告が見い出せたが、それらは必ずしもヒューマンインタフェースのアーキテクチャそのものを目指した研究ではなく、主として AI における知識情報の取扱いに関するアーキテクチャの研究に焦点を当てたものであったが、今後のヒューマンインタフェースのアーキテクチャ研究に応用できると考えられるので紹介する。

取り上げたアーキテクチャ・モデルは、パイプライン・モデル、PDP モデル、黒板モデルの 3 つであり、以下にその研究概要を述べる。

2.2 パイプライン・モデル

ここで述べるパイプライン・モデルは、Sheehein Workshop on User Interface Management Systems (Eurographics and IFIPS 主催) において開発されたユーザインタフェースのためのモデルであり、Sheehein model [1]) と呼ばれる。

図 2.2-1 に示すように、パイプライン・モデルは、アプリケーションとユーザの間に 3 つのコンポーネントを含む。これらは、互いに「トークン」を送ったり、受け取ったりすることでコミュニケーションを行う。

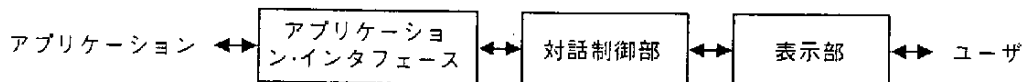


図 2.2-1 パイプライン・モデルの基本アーキテクチャ [1]

3 つのコンポーネントは以下の通りに説明される。

(1) アプリケーション・インタフェース部 (application interface component)

直接ユーザ・インタフェースに関係するアプリケーションのデータ構造とルーチン群を表現する。

(2) 対話制御部 (dialogue control component)

ユーザ・アプリケーションにおける対話を監視する。ここでは、アプリケーション・インタフェース部によって生成されたトークンを表示部用のトークンへ変換したり、また、その逆を行ったりする。変換のためのスキーマは、遷移ネットワークや文脈依存文法、アプリケーション・インタフェース部や表示部からのトークンの到着によって到着されるイベント・ハンドラー等で記述される。

(3) 表示部 (presentation component)

ディスプレイやデバイスのインタラクションを制御したり、受け取ったトークンから、どのようにディスプレイに表示するのか等を決定する。ここでは、メディアやデバイスに依存した部分のみを含む。

なお、Sheehein のパイプライン・モデル自身は各コンポーネントがどのように構築されるのか、どのように実装されるのかについての詳細は規定していない。

S. Feiner[1] は、この Sheehein model を基本に、知識ベースとの対話や複数のメディアを表示する際の問題等を説明するための知識ベースグラフィカル・インタフェースのアーキテクチャ・コンセプトを発表している。

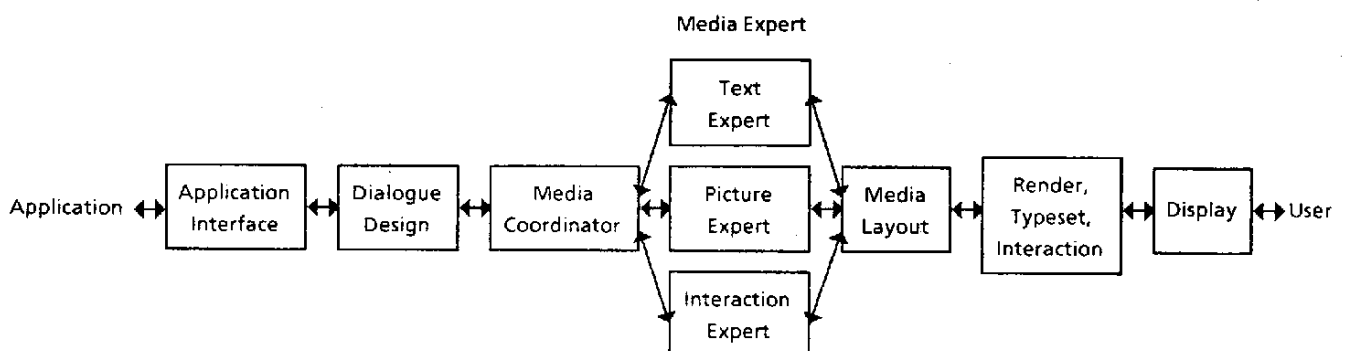


図 2.2-2 パイプライン・モデルのグラフィカル・インタフェースへの応用例 [1]

ここでは、各コンポーネントはパイプラインにおける各ステージと言うよりも、コミュニケーション・エージェントとして考え方が適切であると思われる。

なお、図 2.2-2 では描かれていないが、多くの知識ソースがあり各インタフェース・コンポーネントに対してアクセスできる。これらの知識ソースはセッションの履歴の他、ユーザモデルや現在の状況、ハードウェア、アプリケーション特有の知識等を含んでいる。

2.3 PDP モデル [1],[2],[4]

「脳の神経回路網」の研究をベースとする情報処理に関するアーキテクチャの検討は以前から盛んにおこなわれているが、ここで述べる PDP (Paralled Distributed Processing) モデルも、UCSD の D. E. Rumelhart, J. L. McClelland 等 [2] によって人間の認識モデルとして提案されたものである。これはコネクショニストモデル (Connectionist Model) と呼ばれ、並列コンピュータ、分散協調型問題解決の基本アーキテクチャにも応用されている。

PDP モデルのコンセプトは図 2.3-1、図 2.3-2 に示すように非常にシンプルで、「ユニット」と呼ばれる多数のノードがあり、それらは互に重み付けされた「リンク」によって結合されたネットワークを構成している。

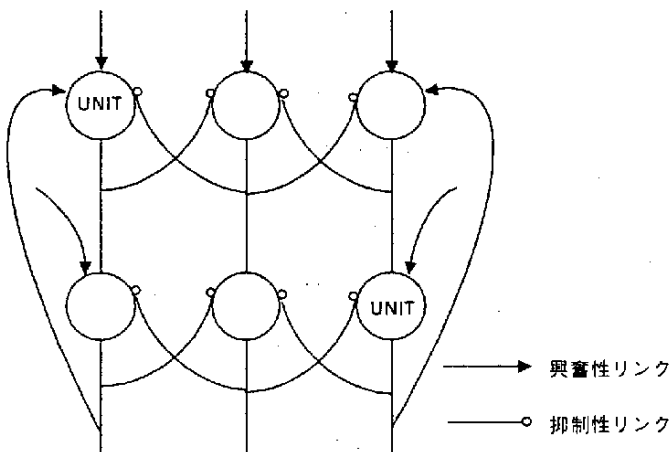


図 2.3-1 PDP モデルの概要図

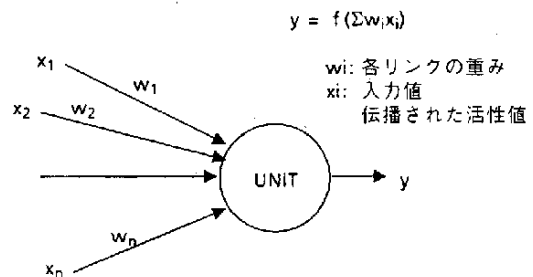


図 2.3-2 ユニットにおける活性化関数

「ユニット」と「リンク」は以下のとおりに説明されている。

(1) ユニット

シンプルな処理エレメントで、「活性化値 (activation value)」と呼ばれる内部状態を持つ。

(2) リンク

ユニット間を結合するもので、以下の2種類がある。

- | | |
|----------------------|---------------------|
| 興奮性 (excitatory) リンク | 活性化値を高めるように働く興奮性の入力 |
| 抑制性 (inhibitory) リンク | 活性化値を低めるように働く抑制性の入力 |

各ユニットでは、外部からの入力、またはリンクを通して伝播された値から「活性化関数 (activation function)」によって活性化値を計算し、その値は連結されたユニットへと伝播される。

ネットワークの機能は、ネットワークの結合パターン、リンクの重み付けのパターンによって決定される。また、各ユニットの活性化値のパターンによって、様々な概念が表現される。

現在の研究では、ネットワークは正しく結合されているという前提で、リンクの重み付けのパターンがバック・プロパゲーションやボルツマン・マシン等のアルゴリズムによって計算され、学習されるというアプローチが一般的である。

PDP モデルの特徴は、十分な情報が与えられていない、あいまいな状況に対処する能力や状況の変化に適応できる柔軟性 (学習能力)、大規模な分散性・並列性にある。

また、実際の応用分野としては音声認識や画像認識などにおける研究が盛んなようである。

2.4 黒板モデル [5],[6],[7],[8]

黒板モデルは、人間のパターン認知モデルを模した問題解決のアーキテクチャであり、カーネギー・メロン大学の L. D. Erman ら [5] によって開発された音声理解システムである HEARSAY-II によって提案された。

黒板のモデルの基本アーキテクチャは図 2.4-1 に示す通りであるが、モデルは以下のような 3 つのコンポーネントから構成される。

(1) 黒板 (BlackBoard)

階層化されたデータベースで、ここには入力情報や問題空間、仮設等が書き込まれる。

(2) 知識源 (Knowledge Source)

問題解決における知識の集合。黒板を常に監視し、黒板の状態に応じて起動される。知識源は互いに独立しており、黒板を通じて互いにコミュニケーションを行う。

(3) スケジューラ

黒板を監視し、競合する仮設の評価選択を行ったり、その仮設を処理する知識源に対して処理能力を分配することで全体の問題解決の調整を行う。

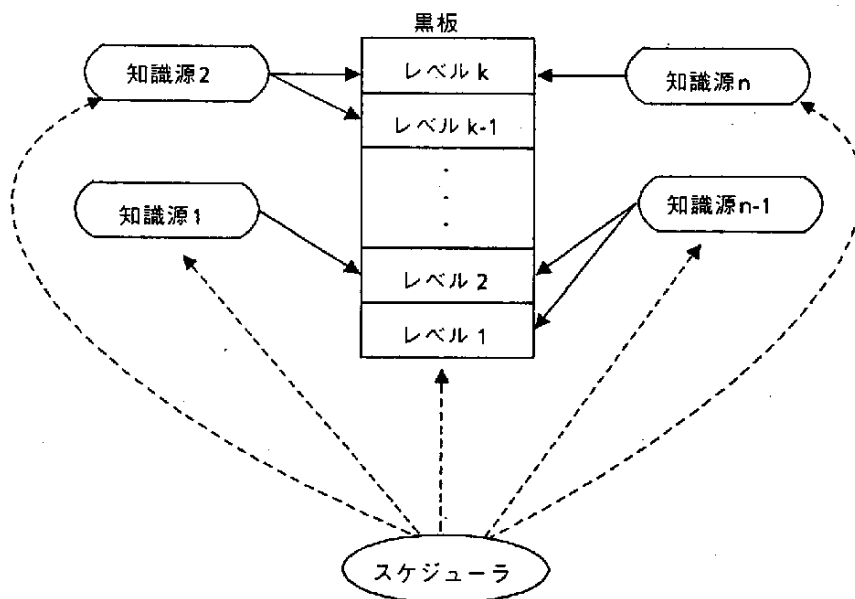


図 2.4-1 黒板モデルの基本アーキテクチャ [5]

新しい仮設が黒板に書き込まれると、それを監視していた知識源がその変化に反応するように起動され、新しい仮設を生成する。これらの仮設の中には、さまざまなレベルの仮設が含まれる。また、相反する仮設が含まれている可能性もある。これらの仮設はスケジューラによってそのつど評価され、もっとも妥当性の高い、処理効率のよい（GOAL に近そうな）仮設が新しい仮設として提案され、黒板に書き込まれる。これが問題空間を網羅し、かなり信頼性の高い仮設が導き出されるまで続けられる。

具体例として、HEARSAY-II が対象としている音声理解システムにおける黒板の階層レベルと知識源を図 2.4-2 に示す。

発話が行われると、聞き手はその音素の解析を行いつつも、概念レベルの文章の意味の予測を行う。これらは平行して処理される。このように異なったレベルの処理が同時に進められることも人間の認知モデルの

特徴である。

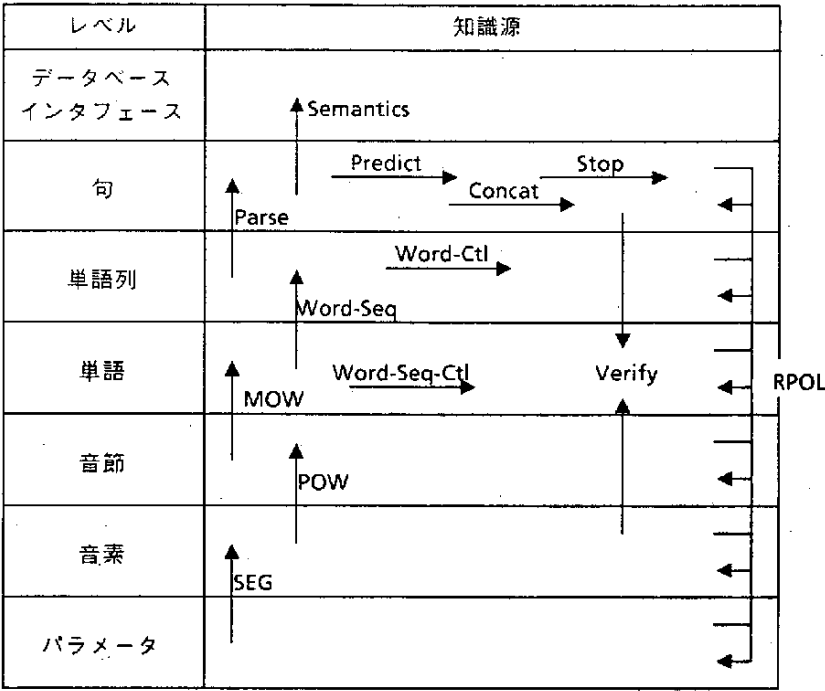


図 2.4-2 HEARSAY-II のレベルと認識源

以上のように、黑板モデルは、複数の独立した知識源による「協調型問題解決」におけるモデルを提供する。

言い換えれば、黑板モデルは各知識源が身分が分析できる情報を捜しながら黑板をじっと見ていればよいし、他の知識源が分析するために自分の分析結果を黑板に書き加えてゆけばよい。このルールを守っている限りは、例えば処理の方法が「データ駆動型処理（ボトム・アップ処理）」であろうと「概念駆動型処理（トップ・ダウン処理）」であろうとも、その区別を全く意識しなくても処理が自動的に進んでゆくことになり、より柔軟な推論機構を提供できることを意味している。

2.5 おわりに

冒頭に述べたように、現在まではヒューマンインタフェースに狙いを絞ったアーキテクチャの研究が少なく、他の分野での研究成果をヒューマンインタフェース・アーキテクチャに応用しようとするものが多かった。従ってアーキテクチャの評価についても、特定のアプリケーションに対して実装し、その最適化を図った上でのものとなっている。

しかし今後は、コンピュータのパーソナル化の進展に伴い、いかなるアプリケーションに対しても（基本的に）同様なヒューマンインタフェースを提供するための新しいアーキテクチャの研究が必要となってくるものと考えられる。

参考文献

- [1] Feiner, S.: An Architecture for Knowledge-Based Graphical Interfaces, Architectures for Intelligent Interfaces: Elements and Prototypes, pp. 129-140 (1988).
- [2] McClelland, J. L. and Rumelhart, D. E.: An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings, Psychological Review, Vol. 88, No. 5, September, pp. 375-407 (1981).
- [3] Feldman, J. A. and Ballard, D. H.: Connectionist Models and Their Properties, Cognitive Science, Vol. 6, pp. 205-254 (1982).
- [4] Rumelhart, D. E. and Zipser, D.: Feature Discovery by Competitive Learning, Cognitive Science, Vol. 9, pp. 75-112 (1985).
- [5] ERMAN, L. D., HAYES-ROTH, F., LESSER, V. R., and REDDY, D. R.: The Hearsay-II Speech-Understanding System: Integrating Knowledge to Resolve Uncertainty, Webber, B. L. and Nilson, N. L. eds, Readings in Artificial Intelligence, pp. 349-389 (1981).
- [6] Hayes-Roth, F., Waterman, D. A. and Lenat, D. B.: Building Expert Systems, Addison-Wesley (1983).
(AIUEO 訳: エキスパート・システム, 産業図書)
- [7] Lindsay, P. H. and Norman, D. A.: Human Information Processing: An Introduction to Psychology 2nd Edition, Academic Press (1977).
- [8] Rumelhart, D. E.: An Introduction to Human Information Processing, New York: Wiley (1977).
(御領謙訳: 人間の情報処理—新しい認知心理学へのいざない—, サイエンス社)

3. ヒューマンインタフェースにおける物理的入出力

3.1 入力手段

3.1.1 キーボード

(1) アメリカにおけるキーボードの技術動向

アメリカに於けるキーボードについてのいろいろな議論は：

Shneiderman, 東（監訳）：ユーザ・インタフェースの設計，日経マグローヒル社，pp. 192-199, 1987
によくまとめられている。ポイントは次のようである。

データ入力的主要な方法は、今なおキーボードである。この装置はとかく批判されることが多いものの、見事な成功例といえるであろう。何億もの人々がうまく使いこなしており、入力速度は最大で毎秒 15 キー（およそ毎分 150 単語）に達する。もっとも、初心者を入力速度は毎秒 1 キーに満たず、平均的なオフィス作業員では毎秒 5 キー（およそ毎分 50 単語）である。

キーボードの配列に関しては、1879 年代、Sholes の設計したタイプライタは、優れたメカニズムと、うまく文字を配列してタイプ速度をわざと遅くし、キーの詰まりをめったに起こさないようにしたことによって、成功を収めつつあった。この QWERTY 配列では、使用頻度の高い文字の組み合わせをたがいに離して配列し、指の移動距離を長くしている。Sholes の成功は広い範囲にわたって標準化され、およそ 1 世紀後には英語キーボードのほとんどすべてが QWERTY 配列を採用していた。（QWERTY とはクワータイト発音し、キーボードの左上にその順番でキーが並べられていることが名前のゆえんである。）

電氣的キーボードの開発により機械的な問題が解決されたため、指の移動距離が少なくするような、別の配列を提案した [Montgomery, 1982]。1920 年代に開発された Dvorak 配列は、指の移動距離を少なくとも 1 桁短くしたといわれ、熟練タイピストのタイプ速度を毎分およそ 150 単語から 200 単語以上に向上させるとともに、エラーも減少させている [Kroemer, 1972; Martin, 1972]。しかし、Dvorak の設計はなかなか受け入れられなかった。キーボードを Dvorak 方式に簡単に交換でき、約 1 週間の通常のタイプ作業を行えば切り換えられることを実証して報告したが、ほとんどのユーザは進んで努力しようとはしなかった。これは、たとえ改善できることが実証されていても、変えることによって得られる利益がそれに費やす努力を明らかに上回っていると実感されなければ、その改良は広まりにくいということのよい例である。

3 番目の興味深いキー配列に、ABCDE スタイルがある。これは 26 文字の英字をアルファベット順に並べたものであるが、QWERTY キーボードがどこでも手にはいるので、タイプ技能は誰にでもできる一般的なものとなり、ABCDE スタイルの重要性はなくなってしまった。

数字パッドの配列についても、電話機は 1-2-3 のキーを最上列に、電卓は 7-8-9 のキーを最上列に置いている。電話機の配列の方がわずかに有利であることが示されているが、ほとんどのコンピュータのキーボードでは電卓の配列を採用している。

キーそのものについては、キーの間隔、指先のあたりのへこみ、反射とすべりにくさのためのつや消し、力と変位の値、間隔や音によるフィードバック、コントロールキーの大きさ、ホームポジションキーの深いへこみや突起についてのべている。

ファンクションキーについて、その利点はキー入力数と誤りを減らせるという利点にあるが、1つ1つのファンクションキーの用途がユーザによっては分かりにくい場合があることや、指をホームポジションから離さなければならなくなるなどの欠点もある。初心者には、対応する機能を思い出させるような画面表示やテンプレートが必要である、また、ファンクションキーの使用が一貫していないシステムも多く、ユーザを混乱させている。ユーザによっては、指をホームポジションから離すよりも、6～8文字をタイプする方を好む場合もある。

ファンクションキーの特別な種類であるカーソル移動キーについては、通常、上・下・左・右を示す4つのキーからなる。カーソル移動キーの配置は、早く、誤りなく操作するためには重要である。

古い論文ではあるが、

Alden, D.G., "keyboard Design and Operation: A Review of the Major Issues", Human Factors, Vol.14, No.4, pp. 275-293, 1972

は、

各指の圧力

キーの入力速度とエラー率（スピードが一番重要）

練習は意味のない文で行ってはいけない

キーの形状に関する標準化

キーを打ったときのフィードバック

キーの配列に関する標準化、QWERTY、Dvorak との比較

いくつかのタイプのテンキー

マルチプレス・キーボード

キーボードの形状

Interlock Systems（同時に打ってもロックしない機構）

等につき、多くの論文をサーベイしている。

また、

山田、小笹（訳）：タイプライタとその入力方法の歴史的考察——日本語タイプライタの開発動向への視点は、bitのVol.13 No.7-11, 13 1981と6回に渡って連載されたもので、約60頁に及ぶ力作であり。第一回のまえがきを原文のまま引用しておく。

最近のエレクトロニクス技術の進歩にともない、日本語文章の効率的処理が妥当な価格によって近いうちに提供されうる見通しがたってきた。入力段階での最適化だけが現在残された課題である。手で入力する方法は近い将来に活用される唯一の方法であり、コンピュータの助けを借りた各種のキーボードによる入力方法が現在研究され実用に供されつつある。しかし多くの諸技術の分野で起こったように、ここでもごく一部の方法が時の試練に耐えて残ることになるだろう。

この小論文は欧米のタイプライタのたどった過程で何が起こったのかを検証し、またその時点で何が盲点となり落とし穴となったかを指摘し、日本語タイプライタを開発していくうえでの試行錯誤期間を短縮することを心がけたものである。少なくともこうした知識が研究開発の諸局面において、生起するであろう諸々の問題を回避する武器となることができれば幸いである。

なお、本稿においてはタイプライタが社会に受け入れられるために必要な基本的で重要な特色について注意を喚起することに意を用いた。意外に多くの人々が“タイプする方法”についての認識に欠けており、それゆえにキーボードの配列に誤りを犯し、今日においてさえもそのことによって苦しんでいるのである。

(2) 日本におけるキーボードの技術動向

次に、国内の技術動向を調べるために、電子情報通信学会と情報処理学会の論文誌と研究会の資料を調査したところ、キーボードに関する論文は最近の1年間の論文誌には見当たらなかったが、研究会には多くあった。特に、情報処理学会の文書処理とヒューマンインタフェース研究会に多くあった。そこで、同研究会の発足当時からの研究会資料を調べた。

文書処理とヒューマンインタフェース研究会は

1981-1984年度、日本語入力方式研究会

1985-1986年度、日本語文書処理研究会

を前身としている。当初の論文は研究会の名前の示すとうり、キーボードに関するものの他に、かな漢字変換などの日本語の入力に関するものも多かった。

論文の内容を大別すると：

- (1) 日本語の入力方式に関するもの
- (2) キーボードの配列や形状に関するもの
- (3) 2ストローク入力やタッチタイプに関するもの
- (4) 人間側の情報処理に関するもの
- (5) 学習と入力速度のかかわりあいについて

となる。

山田尚勇：「日本文入力法研究委員会」からの8年の歩み、12-4(1987-5)

において、同研究会のその時点までの研究の概要を述べている。また、各発表の概要の部分のコピーがあるので、同研究会の動向を調べるには良い報告である。

同研究会におけるキーボードに関する論文は次のようである。

《日本語入力方式研究会》

細川、1-1：英文キーボードによる日本文の入力について

伊藤、1-2：邦文タイプライタ配列をとり入れた日本語 WP 用ペンタッチキーボード

中山、1-3：2ウェイキーボード日本文入力方式

亀山、2-1：かなキーボードにおけるかな配列方式

平賀、2-3：タッチタイプによる日本語入力方式

黒須、3-2：かなキーボードの文字配列の評価

小川、4-2：配列対応方式による日本文入力法

高嶋、4-3：タッチタイプによる日本語入力の一方式とその練習法

磯道、5-3：字形検字方式による漢字入力の一方法

高木、6-1：日本語ワードプロセッサにおける多段シフト（漢字ストローク）方式の入力習熟度

小田、6-3：2ストローク入力と教育

山田、6-4：タッチタイピングにかかわりのある実験心理学的および大脳神経学的現象について

古川、7-3：印刷屋の見た日本語入力

石井、8-1：音訓コードによる日本文の入力実験

松為、8-2：漢字入力作業における人間側の情報処理について

木越、10-1：英文キーボードによる日本文入力について PART II

竜岡, 10-3: かな漢字変換方式のワード・プロセッサのキー配列
山内, 11-3: 文字種区分方式2ストローク日本文入力法の打鍵特性
渡辺, 11-4: カナタイピストにおける指の運動特性について
黒須, 12-3: 3ストローク英字コードによる日本文入力法の検討
中山, 13-1: 日本語入力速度予測モデルの検討
森田, 13-2: 新しい日本文入力方式
宮沢, 13-3: 二角拼音方式による日本文入力システム: CHAMPS
竜岡, 15-2: HS20 配列について (変換方式のワード・プロセッサの新しいキー配列)
伊藤, 16-3: かな漢字変換入力と一般印刷用原稿
高嶋, 17-2: あいうえお順配列キーボードの使い勝手について
中山, 17-4: 仮名文字鍵盤配列の評価と設計
白鳥, 18-1: 日本語入力用ローマ字キー配列の最適化
坂下, 18-2: カナ文字鍵盤の文字配列
岡留, 19-3: タイプ入力作業の構成要素間に起こる干渉

《日本語文書処理研究会》

谷口, 2-2: 日本文入力方法についての一考察
岡留, 2-3: 脳におけるタイプ作業処理過程のバイブラインモデル
大野, 3-1: ローマ字入力方式のモードに関する検討
木村, 6-4: 日本語ワープロ向け新打鍵レベルモデルの検証と応用
鄭, 7-1: ハングル (韓国語) 入力用キーボード「正音」
坂村, 7-2: BTRON における入力方式 - TRON キーボードの設計 -
森田, 8-2: 完全ブラインド・タッチ型キーボードの一提案
富樫, 9-2: 超多段シフト和文キーボード
渡辺, 9-4: カナタイピストにおける指の運動特性について (続報)
神阪, 10-1: 5指コードによる日本語入力5指コードの汎用性
大島, 11-1: 標準時間法による入力方式別スピード比較実験
増田, 11-4: キーボードのブラインド・タッチ短時間練習法

《文書処理とヒューマンインタフェース研究会》

森田, 12-3: 各種の日本文入力方式の性能の定量的比較
山田, 12-4: 「日本文入力法研究委員会」からの8年の歩み
白鳥, 15-4: 左右対象形キーボードの操作性評価
小野, 16-1: 全かな相対母音方式日本語入力法
森川, 16-2: キーボードエミュレーションを行うとき望まれる BIOS の機能について
竜岡, 16-3: キー配列の自由化環境の確立のために (mykey 配列の提案を通して)
喜多, 16-4: 2ストローク入力用仮名漢字変換システム
平井, 19-1: パーソナル・キーボード

最近の発表 16-3 と 19-1 はパーソナルなキーボードについてのヒントを与えている。

次に、電子情報通信学会: 技術研究報告「オフィスシステム」のうちキーボードに関係あるものを調べたが、86年からのものは次のように、情報処理学会に比べて多くはなかった。

大岩, OS-86-31: OA のためのキーボード教育

高嶋, OS-86-32: ヒューマン・コンピュータ・インタフェースとしてのキーボード

大岩, OS-87-41: 3時間でマスターできるブラインド入力技術

次に, 大きな流れを示していると思われる論文:

山田: 専任タイピストによる日本語入力方式に関する基礎的研究課題, 電子通信学会誌 pp. 939-945(9/'83)
につき, 概要をまとめると, 次のようになる。

専任タイピストが高速に日本語文を入力することに関して述べている。英文タイピストの場合は脳の右半球の操作空間領域の機能に頼っているが, かな漢字変換によるかなの入力では, 漢字を読むために左半球を使用し, 変換のための作業も干渉している。(大腦についていうと, 反射的な行動は右半球で, 思考的な行為は左半球で行っている。したがって, 単純なタイプ入力は右半球だけで済むが, かな漢字変換のように, 一度漢字を読み, かなに直すという思考的なものが入るため, まず右半球を用い, 次に左半球を用いて入力することになる。このため, 左右半球の干渉が生じ, 作業は複雑になる。)

文字パターンをキーボード上の指の動きに直結して打つ2ストローク方式の「反射式」がよいとしている。しかし, 熟練者が長時間作業したとき作業能率と疲労度にどれだけの開きができるかは明らかではない。

「もし一生のあいだ, 文字を書くのが楽で, しかも数倍の速さになれるとすれば, そのために数百時間のタイプの練習を上積みすることは十分に引き合う投資であろう。」というところは印象的であった。

アメリカにおけるキーボードの歴史とヨーロッパにおける八の字の形のキーボードについてもふれている。操作性について, 従来専任タイピストに最も好まれたIBM セレクトリック・タイプライタには:

- (a) 押し下げの途中でキーイングの終了を指に感じさせるタクタイル・フィードバックの良さ
- (b) 押し下げが終ったところでの指先きへのクッション
- (c) 高速操作をしても引っかかりや文字落ち, チャタリングなどの出ない機械的性能
- (d) 多段のロールオーバーがきいて, 連続キーイングが重なっても文字落ちがないこと
- (e) 指の強弱に合わせてキーごとに調整された反力調整

などが実施されていた。エレクトロニック・キーボードになってからは, コストダウンの結果, 上記の(a), (b)や(e)を欠くものが普及し出し, 中には(c)や(d)さえも満足していないものが出回っている。

3.1.2 マウス

情報処理学会と電子情報通信処理学会の研究会の資料を調べたところでは, マウスに関するものは見付からなかった。キーボードについては多くあったことを考えると, キーボードに関する研究の方が圧倒的に多いといえる。また, 87/1 から 89/1 のINSPEC2を検索したところ, 次の4件のみであった。

Burgess: Helping Users use UNIX, 10th Annual for Computers and Industrial Engineering, 23-25 March 1988 P.244-8 1988

UNIXを使用する新しいユーザに対して, 構造化されたヘルプの絵をポイントすることでのヘルプシステムについて述べている。

竹村: ソフトウェア手法によるポイント効率の改善について, 電子情報通信処理学会論文誌 D Vol. J70-D No.12 pp. 2402-2409 1987年12月

マウスカーソルの移動をマウス本体の移動速度によって変化させることで, マウスを用いたポイントに必

要な机上の作業領域を削減する手法を提案し、実験し結果について考察している。マウスの作業領域の減少に有効であり、ポイント時間も従来とほぼ同等としている。また、オフディスプレイ型のタッチパネルのエラー率改善についても論じている。

Abernthy: Ergonomically Determined Pointing Device (Mouse) Design, Ergodesign 86- The Evolution of the Electronic Workspace. 21-24 Oct. 1986 P. 311-14 July-Sept. 1987
使いやすいマウスの形状のデザインに関する論文である。

Jackson: What's behind the Windows, Informations.(GB) Vol. 8, No. 1 P. 22-4 Jan. 1987
ウィンドウ環境の歴史と特徴について述べている。また、商品についても調べている。

最近のマウス（一部ポインティングデバイス）に関する論文として、次のものがあった。
池本：マウスを用いたユーザ・パフォーマンスの研究，情報処理学会第36回（昭和63年前期）全国大会，
7N-6

Card らによれば，人間の認知情報処理モデルは独立して働く，
知覚システム（Perceptual System），
認知システム（Cognitive System），
運動システム（Motor System）

から構成され，連続的な入力などは，1入力ごとに知覚－認知－運動システムの1系列の処理を行い，この直列的処理を繰返し行っているとしている。

マウス操作によりカーソルを目標まで移動し打鍵する課題において，明らかになったことは，

- (1) 訓練効果はマウスを移動する方向には関係なく一定である。
- (2) 課題の遂行の時間としてのパフォーマンスは人間の身体的制約に依存した方向性がある。（右利きの場合，左および下方向は右および上方向と比較して遂行時間を要する）
- (3) ユーザの情報処理過程における知覚，認知，運動処理の段階のうち，技能の習得により最も効果の表われるのは，運動処理の段階である，
としている。

西中：座標入力のためのポイント手法について，電子情報通信学会論文誌 D Vol. J71-D No. 12 pp. 2604-26
12 1988年12月

細かい目標の選択において，エラー発生率を低減するためのポイント手法としての，マウスのためのブレーキタイプの考案である。

マウス（機械式・工学式），タッチパネル，カーソルキーにより，メニューやアイコン，テキスト中の語の選択，グラフィック画面での座標の入力について測定，比較を行った。結果として，マウスの机上の作業領域を低減させる手法として，マウス本体の移動速度に応じてカーソルの移動量を非直線的に変化させる非リニアの手法が有効である，としている。位置確立のためにマウスボタンを押すときには，移動が遅いので，カーソルの動きを制限する（ブレーキタイプ方式）のでずれない。

西中：画像解像度の違いに対するポイント効率の評価について，ソフトウェア工学 64-20 (1989. 2.3)

計算機の性能向上にともなって今後進むであろうディスプレイの高解像度化を想定して、マウスを用いたポイント手法について、画面解像度の違いがポイント効果に与える影響について調べている。その結果、解像度の高いディスプレイを用いる場合には、マウス本体の移動速度に応じてカーソルの移動量を非線形に変化させる非リニアの手法が、机上の作業領域を削減し、かつポイントに要する時間も短縮することが明らかとなった、としている。

富樫：人はねずみと共存できるか、情報処理学会・夏のシンポジウム「ヒューマン・フレンドリなシステム」, pp. 121-131, 7月1986年

カーソルキーやトラックボール、ジョイスティックに比べてマウスでは、移動方向だけでなく移動量そのものも実移動量に比例してあたえることができる。しかし、マウスの致命的な欠点は、キーボードとの積極的な併用を真面目に考えていないところにある、としている。キーボードとマウスの間の手の移動がわずらわしい。

キーボードとマウスの間の矛盾を両者の交配であるキーマウスで解決した、としている。マウスを両手で、手の平で包み込む。親指を多用し、親指には4×2の主に機能鍵群を、他の4指には6×4の文字鍵群を割り当てる。外観を図3.1-1に示す。

また、キーマウスの有用性を実証するためのテキストエディタについてと、足で操作するフットスイッチについても論じている。

久野：窓はねずみ無しでも操れるか？、情報処理学会・第30回プログラミングシンポジウム, pp. 93-104, 1月1989年

ウィンドウの制御では、マウスとキーボードの間を手が往復するオーバーヘッドはかなり大きく、マウスを使わない方法の方が優れている可能性がある。そこで、マウスの代わりにキーボードのみを使用してウィンドウを操作する機構をX-Window (ver. 11) 上に作成し、比較分析した。

マウスに手を伸ばした方が速いということの無いように、キー操作を短くした場合には、学習の負担が大きくなるが、熟練者が使用した場合の評価を想定した。

ShiftキーとControlキーを同時に押しながらあるキーを押すと、ウィンドウマネージャへの入力となる。

当初は無意識のうちに手がマウスに動いていることがあったが、やがて慣れて来るとキーボードのみで快適に操作できるようになり、マウスに手を移すのがうっとうしくなるように思われた、と述べている。

マウスとキーボードによるウィンドウの操作を

1. 新しく必要になった窓を作り出す
2. 複数の窓を切り替えながら作業する
3. 窓の配置を好みに合わせるため、移動する
4. 窓の大きさを作業しやすさのため変更する
5. 窓をアイコンにしたり、戻したりする

の5通りにつき実験を行い比較して、その結果キーの方が優れており、「窓はねずみ無しでも操れる」、と結論している。

吉田：マウス操作に関する基礎研究, 2nd Symposium on Human Interface, Oct. 29-30, 1986 Tokyo, 1211

「人がいかにしてマウスを使いこなしているか」という観点に立ち、マウス操作におけるアナログ的な動作をターゲット付近まで素早くカーソルを移動させるプログラム動作と、微調整を行うフィードバック動作に分けて解析し、2つの動作のかかわりについて述べている。

結論は、プログラム動作はターゲットの間の長さ、ターゲットの大きさおよび表示感度に殆ど関係なく一定時間（約0.3秒）行われ、フィードバック動作はターゲットの手前約2割の地点あたりで開始され、ターゲットまでの長さやターゲットの大きさによってその動作時間が異なる、としている。

武藤：マウス感度の切換法の提案と選択時間の解析, 2nd Symposium on Human Interface, Oct. 29-30, 1986 Tokyo, 1212

マウスの感度とは「画面上のカーソルの移動距離と机上でのマウスの移動量の比」である。

マウスの感度を移動操作中に動的に変化させることにより、移動距離を減少させる感度切換え法を検討し、数種の切換え法において選択時間等を比較した、としている。

Pearson: Exploratory Evaluation on a Planar Footoperated Cursor-positioning device, ACM CHI '88

カーソルのポジショニングのために、手でなく足を使用することについてである。素人が1/8インチの四角の目標を選択することは、容易に学習でき、マウスの場合のキーボードのホームポジションへ手を戻すことまで計算に入れば、時間は等しい、としている。

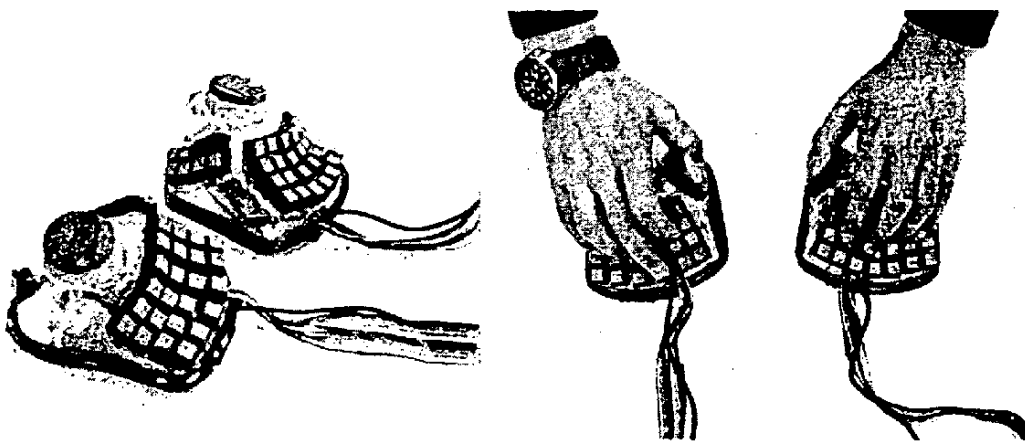


図 3.1-1 キーマウスの外観

富樫：人はねずみと共存できるか。情報処理学会夏のシンポジウム「ヒューマンフレンドリーなシステム」1986年7月 p. 124

3.1.3 その他のポインティング装置

Shneiderman, 東（監訳）：ユーザ・インタフェースの設計, 日経マグローヒル社, pp. 199-210, 1987
の論文より、様々なポインティング装置の比較をしている部分を中心にまとめる。

ライトペンは、画面を直接指すので、腕が疲れてしまうことがある。キーボードから手を離さねばならない。取り上げるために手を伸ばさねばならないといった問題もある。

タッチスクリーンは、指の太さに依存した不正確さがあるので、正確な位置付けが必要ない場合に有効である。画面が汚れるという問題もある。また、ライトペンは近付けて押せるが、タッチスクリーンはスイッチが一段なので、スクリーンへの接触に即座に反応してしまい、正しいかどうかを確認する機会がユーザに与えられない。ただし、赤外線を利用して、画面に近付けたことを検知できるものもある。ライトペンとの

差はあまりない。

マウスは、拾い上げなければならないこと、机上の一定の空間が占領されること、ケーブルに気をつけなければならないこと、移動距離が長い場合は、一度持ち上げて置き直す動作が必要になること、使いこなすにはある程度の練習が必要であることが問題としてあげられる。ジョイスティックよりも速いことが認められた。画素単位で正確に位置付ける必要があるときに有効である。

トラックボールは、マウスをひっくり返したもので、ボールを強く打って急激に回転させることもでき、ビデオゲームには好まれる装置である。ジョイスティックよりも速くかつ正確である。マウスと同様に、画素単位で正確に位置付ける必要があるときに有効である。

ジョイスティックは自動車や航空機の制御に始まる長い歴史がある。画面上を移動する対象物に追従させるには魅力的な装置である。しかし、不正確で遅いために、カーソルを一定の方向へ移動するのは困難である。

グラフィックス・タブレットは、手を楽な位置に置くことができ、画面から離しておける。印刷物をのせることもできるため、初心者ガイダンスを与えることもできる。キーボードを使わずに長い間この装置を使っていられるような場合には魅力的である。トラックボールよりわずかに速いが正確さが少しばかり劣る。

カーソル・ジャンプ・キーは、ターゲットが少数の場合は有効である。即ち、ポインティング装置はカーソル・キーよりも速い。しかし、それは作業内容に依存する。画面上のターゲットが少なく（2個から10個）、カーソルがあるターゲットからの別のターゲットにジャンプできるような場合などは、カーソル・ジャンプ・キーの方がポインティング装置よりも速くなる。カーソル・キーを用いたときの位置付けの時間は距離とともに急速に増加するが、マウスやトラックボールではその増加はわずかなものにすぎない。

参考文献

- [Abernrthy, 1986] Ergonomically Determined Pointing Device (Mouse) Design, Ergodesign 86 - The Evolution of the Electronic Workspace. 21-24 Oct. 1986 p.311-14
- [Alden, 1972] Keyboard Design and Operation: A Review of the Major Issues, Human Factors, Vol.14, No. 4, pp. 275-293, 1972
- [Burgess, 1988] Helping Users use UNIX, 10th Annual Conference for Computers and Industrial Engineering, 23-25 March 1988 p. 244-8 1988
- [Jackson, 1987] What's behind the Windows, Informatics. (GB) Vol.8, No1 p. 22-4 Jan. 1987
- [Kroemer, 1972] Human engineering the keyboard, Human Factors 14,1, (February 1972), 51-63
- [Martin, 1972] A new keyboard layout Applied Ergonomics 3,1, (1972).
- [Montgomery, 1982] Bringing manual input into the 20th century, IEEE Computer 15,3, (March 1982), 11-18.
- [Person, 1988] Exploratory Evaluation on a Planar Footoperated Cursor-positioning device, ACM CHI '88
- [Shneiderman, 東 (監訳)] ユーザ・インタフェースの設計, 日経マグローヒル社, pp. 192-199, 1987
- [池本, 1988] マウスを用いたユーザ・パフォーマンスの研究, 情報処理学会第36回(昭和63年前期)全国大会, 7 N - 6
- [久野, 1989] 窓はねずみ無しでも操れるか?, 情報処理学会・第30回プログラミングシンポジウム, pp. 93-104, 1月 1989年
- [竹村, 1987] ソフトウェア手法によるポイント効率の改善について, 電子情報通信処理学会論文誌 D Vol.J70 - D No.12 pp. 2402-2409 1987年12月, July-Sept. 1987
- [富樫, 1986] 人はねずみと共存できるか, 情報処理学会・夏のシンポジウム「ヒューマン・フレンドリなシステム」, pp. 121-131, 7月 1986年

- [西中, 1988] 座標入力のためのポイント手法について, 電子情報通信学会論文誌 D Vol. J71-D No. 12 pp. 2604-2612
1988 年 12 月
- [西中, 1989] 画像解像度の違いに対するポイント効率の評価について, ソフトウェア工学 64-20 (1989.2.3)
- [武藤, 1986] マウス感度の切換法の提案と選択時間の解析, 2nd Symposium on Human Interface, Oct. 29-30, 1986
Tokyo, 1212
- [山田, 小笹 (訳), 1981] タイプライタとその入力方法の歴史的考察——日本語タイプライタの開発動向への視点, bit
Vol. 13, No. 7-11, 13, 1981
- [山田, 1983] 専任タイピストによる日本語文入力方式に関する基礎的研究課題, 電子通信学会誌, pp. 939-945, 9/1983
- [山田, 1987] 「日本文入力法研究委員会」からの 8 年の歩み, 情報処理学会文書処理とヒューマンインタフェース研究
会, 12-4, 5 月, 1987
- [吉田, 1986] マウス操作に関する基礎研究, 2nd Symposium on Human Interface, Oct. 29-30, 1986 Tokyo, 1211

3.2 インタフェース媒体

ヒューマンインタフェースの物理的側面を支える様々な技術が重要となってきた。現在の入出力デバイスも10年後にはその解像度やレスポンスタイムなどの改善はあっても理想的なインタフェース実現にはまだまだ制約の多いことが予測される。限られた機能のなかでこれまで先人達が考案し、実用に供してきた手法をふりかえり、考えだされた背景や必然性を自身で理解するとは、我々のめざす次世代ヒューマンインタフェースの研究に役立つであろう。本章では物理的入出力デバイスであるディスプレイをヒューマンインタフェースの窓口として使う場合に考えだされた媒体であるコマンド、メニュー、アイコン、ウィンドウについて報告する。

3.2.1 コマンド言語

アーサー・クラークの“2001年宇宙の旅”でボウマン船長はコンピュータ HAL との会話を音声で行っていた[1]。W・ギブソンの“ニューロマンサー”の主人公コンピュータ・カウボーイは特殊な電極で自分の脳とコンピュータをインタフェースし、ネットワークにジャックイン（没入）する[2]。

我々はコンピュータとの理想的インタフェースをつねに追い求めている。自然言語の理解による方法の一つの手段であるが、機械との対話の言葉としてすべての場合で効率的であるとはいえない。パベッジも早くから機械の振る舞いを記述する記号の必要性を述べていた[3]。FORTRAN, COBOL 等の非会話型から BASIC, APL の会話型までプログラミング言語は問題解決のためのプログラムを作り上げることに大きく貢献したがより直接的に、目的の操作をコンピュータに与え、問題解決までの近道をとるためにはコマンド言語も必要としている。

コマンド言語はメニュー・システムやアイコン・プログラミングと比較してユーザに多くの記憶を強いるという短所はあるが頻繁に使うユーザにとっては利点も多い。特にハードウェアの制限でディスプレイが小さい、レスポンス時間が長い、機能が多く何通りもの組合せが起こる等の条件下で有効である。ユーザはコマンド使用のルールを学ぶ必要があるがシステムの応答を待つのでなく、どちらかと言えばコマンドやコマンドの組合せの入力でタスクを起動させる。ディスプレイのレスポンス時間が短い場合にはコマンドをメニューに組み込んだコマンド・メニューが有効であり、ポップアップメニュー、プルダウンメニュー、バーメニューが成果をあげている[3]。

コマンド言語設計の基本方針について、構成、引数付コマンド、階層構造的コマンドなどを Shneiderman の著書に詳しく述べられているので参考になる[4]。

3.2.2 メニュー

コマンド入力方式では自由度と交換にユーザに記憶を強いる欠点があり、初心者にも受け入れられ易いのメニュー選択システムである。メニューの構造や設計上のポイントなどについては shneiderman の著書[13]に丁寧に述べられている。メニューには構造で分けると、シングルメニュー、リアメニュー、ツリーメニューがある。シングルメニュー中でもバイナリメニュー（yes/noで選択）、多項目メニュー、拡張メニュー（複数ページ）、ポップアップ（プルダウン）メニュー、固定メニュー等がある。メニュー選択方式ではシステムのレスポンスタイムで評価が大きく分かれることがある。メニュー方式は慣れて来るにつれてレスポンスの遅さや繰り返しアクセスに不満が出る。このような上級ユーザのために、タイプアヘッド入力、深いレベルへの直接アクセス、マクロの定義などと述べられている。

具体的な例として Macintosh のプルダウンメニュー、ポップアップメニュー、メニューバーなどが参考になる[14]、[15]。また最近のビデオ・ゲームでもメニュー選択の利用が当たり前のようになっていることは注目に値する。

3.2.3 アイコン

アイコンはプログラミング環境やビジネス・ソフトからゲーム・ソフトまで広い範囲で、その“視覚に訴える親しみ易さ、理解し易さ”という理由から利用されている。人類は歴史的にも古くから絵文字（エジプトとヒエログラフ（象形文字））や、表意文字（中国の漢字）をコミュニケーションの手段として利用してきたので[5]、コンピュータの性能上の制約が解かれた現在から将来にわたり、アイコンによるコンピュータとの会話が益々重要となることが予測できる。

ゼロックスの STAR やアップル社の Lisa, Macintosh ではファイリング・キャビネット、紙屑かご等、オフィス環境で見られるものを絵で表現したアイコンを使い、所謂デスクトップ・メタファを実現し成功をおさめている。しかしこれらの例ではいずれもアイコンの数や使われ方がシンプルなことから初心者にもあまり抵抗なく受け入れられているが、今後機能を増やしたりするうえでアイコンが増えてくると、定義の曖昧さや、同じアイコンが違った意味を持つ可能性などの問題点が指摘されている。

アイコンの記号的正確を最もよく論じているのはパース（Peirce である[6]、[7]）。良く出来たアイコンは誰にでも、国籍や、教育程度に関係なく同じ理解がされるものである。パースは記号を類似記号（アイコン、イコン）、指標記号（インデックス）、象徴記号（シンボル）の3種類に明確に区分して論じている。

(1) イコンとは記号の表現部がその内用部と類似の関係にあるものをいう、

たとえば、“人”という漢字は人間が立っている形をしめす類似記号である。

(2) インデックスとは表現部と内用部が隣接の関係にあるものをいう、

たとえば、煙りと火の関係がそれです。煙の形が火に似ているのではなくて、煙りからそこには火があることを知っているからである。

(3) シンボルとは表現部と内用部が習慣として決められているようなものをいう、

たとえば、咲いている花を“ハナ”と呼ぶのは“ハナ”という表現部が“花”を意味する内容部と類似や隣接の関係をもつのでなく、習慣としてつづかっている。

パースの研究はアイコンの利用に関して奥の深い示唆を我々にあたえてくれる。現実のシステムの中でも、アイコンだけでなく、シンボルやインデックスをも含めることがあるので、絵記号の位置づけを十分検討してゆくひつようがある。

どれだけの数のアイコンがつかえるかという実際的な問題があるが、それは

適当な図形がつかえる可能性、表現の正確さの程度

によるものであり、あまりにもアイコンの数が増えると問題になる。漢字はその典型的な例であり、もはやアイコンではなくシンボルとなった[8]。

しかし漢字の研究はアイコンの数の増大に対処するなにかのヒントをあたえてくれると思われる。一つにはたくさんの数の漢字も一つ一つは体系だった部分構成要素が組合せられているということ。漢字の構造と機能を分析し、分類した“六書”の理論によると、すべての漢字はつぎの6種類に分けられるとされ、細かい解釈がなされている[9]。

1) 象形 2) 指示 3) 会意 4) 形声 5) 転注 6) 仮借

注目すべきは音の部分と意味の部分を合成した字形である形声文字が全体の8割りをこえているようで、表

意文字としての漢字が表音的にも充実した機能をもっているということである。サウンドがアイコンに組み込まれて使われるという状況は容易に考えられることである。

アイコンの表現には曖昧さがつきまとう。たとえばデスクトップ・メタファの例で、紙屑かごの場合に投げ入れた書類は再び取り出せるか否かは大きな問題である。設計者の勝手な判断で定義づけを行って行くことは将来に禍根を残すことになる。現在、国内では日本規格協会情報技術標準化センター (INSTAC) で標準化の作業が進められているのでその動向に注意しておく必要がある [10]。

アイコンの最大の特徴は視覚的であるということである。人間の心理として、目に視えるものは快く感じ、理解し易い。反対に、視えず隠れた存在のものには不快感が付きまとうし、理解もしにくい。この現実からアイコンをプログラミング環境に取り入れ、ビジュアル・プログラミングをめざそうという試みがなされている [11], [12]。

古くは Sutherland の Sketchpad から、Smalltalk でかかれた David Smith の Pygmalion, Gael Curry の PAD, Alna Borning の Thinglab, tanimoto の PICT などが報告されている [13]。

3.2.4 ウィンドウ

コンピュータの性格が進歩してもインタフェースの窓口としてのディスプレイのバンド幅の狭さが問題となっている。我々は本や資料を机の上に広げて一瞥し、必要な部分に目を移してゆくことが出来る。ハードウェアとしてはビットマップ・ディスプレイ上に複数の独立したウィンドウを開き、同時に多くの情報を表示するものである。Smalltalk [16] や、Interlisp-D などのマルチウィンドウ・システムでは複数の重なりあうウィンドウの管理機能があり、いずれもウィンドウ群をスタックとして重なり順に記録し、スタックの最上位のものは全体を表示し、下位のものは隠れた部分を保存し残りを表示するという基本的なアルゴリズムをメニューシステムで起動しながら、つぎつぎとウィンドウを切り替えてゆく。実際の設計に当たっては、処理時間が問題となることも多く、今後さらにハードウェアの改良も期待される分野である。

アプリケーション環境でのマルチウィンドウの利用も様々な分野で進んでいるが、MIT のアーキテクチャ・マシソン・グループのデータランドの実験は何かと参考になる [17]。メディア・ルームのデータランドではそれを表示するワールドビュー・モニタに、顔、地図、テレビ、手紙、本の表紙、カレンダー、電話、電卓などのアイテムが一つのスクリーン上に配置されて映しだされる。利用者は自分の好みに合わせてアイテムの配置を変えられるし、スクリーンのなかで自分の居場所をかえて必要なアイテムにアクセスし、拡大して細部を見ることが出来る。空間的データ管理システム (SDMS) と呼ばれるこの実験システムは“あそこにある、あれが見たい”という普通の人の、検索行動をすなおに実現しようとするトライアルであるが、次世代コンピュータのヒューマン・インタフェースをめざす我々の研究にとっても教えられることが多い。

参考文献

- [1] アーサー・C・クラーク：“2001年宇宙の旅”，早川書房，(昭和52)
- [2] W. ギブスン：“ニューロマンサー”，早川書房，(昭和61)
- [3] Nan C. Shu：“Visual Programming”，Van Nostrand Reinhold，1988
- [4] Apple Computer：“Macintosh Userinterface Guidelines”
- [5] A. C. ムーアハウス：ねずまさし訳：“文字の歴史”岩波書店 (1956-3)
- [6] 米盛裕二：“パースの記号学”，草書房，1983
- [7] 川本茂雄：“ことばとイメージ”，岩波書店，1986

- [8] KAI. A. OLSEN: "Icon in man-machine communications", BEHAVIOUR AND INFORMATION TECHNOLOGY, 1985, Vol. 4, No. 3, pp. 249-254
- [9] P. キュンメル: "西洋からみた漢字文化", ヒューマンサイエンス第5巻, 中山書店, 1984, pp. 215-229
- [10] 山本喜一: "標準化の状況 アイコン", 情報処理, Vol. 28, No. 9, pp. 1153-1157
- [11] Nan C. Shu: "Visual Programming", Van Norstrando Reinhold, 1988, pp. 190-238
- [12] H. Levine & H. Rheingold 共著, 椋田直子訳: "コンピュータ言語進化論", アスキー社, 1988, pp. 313-351
- [13] Ephraim P. Glinert & Steven L. Tanimoto: "PICT: An Interactive Graphical Programming Environment", IEEE COMPUTER, NOV. 1984, pp. 7-24
- [14] ハイパーカード・スタックウェア開発技法 (上) (下), BNN

3.3 出力（メッセージ、音、音声、画像）

3.3.1 メッセージ、音、音声

3.3.1.1 はじめに

本文では、メッセージ出力及び音響信号出力によるヒューマンインターフェースの研究動向について、その調査結果を報告している。以下では、はじめに、現在まで最も一般的な出力形態として用いられてきたメッセージ出力について、特に、メッセージの出力手法として定着しているウィンドウ表示法に関する認知心理学な研究に焦点を当て報告している。各種の具体的な商用あるいはPDS (Public Domain Software) ウィンドウシステムの紹介は参考文献の列挙程度に留めている。次に、これまで多くの研究例がみられないが、近未来の新たなヒューマンインタフェースとして興味深い音響信号出力の応用について、音声ICONに関する研究例を報告している。最後に、U. S空軍関連機関により作成された既存の利用者インタフェースソフトウェア設計ガイドラインの紹介、および西ドイツ GMD (Gesellschaft für Mathematik und Datenverarbeitung) によるヒューマンインタフェースの評価ガイドラインの研究例を報告している。

3.3.1.2 メッセージによるコンピュータインタフェースの研究

(1) 全般的な技術動向

特定の専門分野における専門家に対し、通常慣れ親しんでいる仕事環境や仕事のスタイルを、そのまま或はさらに洗練された形のコンピュータシミュレーション環境として提供する事を目的としたワークテーションの分野では、多くの利用者インタフェース構築技術の研究[1]-[3]が行われており、とりわけ主たるインタフェース方法であるメッセージ出力方式であるウィンドウシステムに関しては多くの実用システム[4]-[5]が提案され、本調査に対しても、多くの示唆を与えてくれるものがある。しかしながら、これらの研究、あるいは各種の実用システムの多くは、高度な知識を持つ専門家が利用する事を前提としており、ごく一般の利用者が特別なコンピュータ等の専門知識を必要とせず利用できるウィンドウシステムの構築には、更に人間が共通して有している認識方式からの研究が必要であると考えられる。本項では、この視点から行われたと思われる。U. Sメリーランド大学に於けるウィンドウ或はマルチスクリーン構築に関する研究を紹介する。

(2) ウィンドウのレイアウトに関する研究例[6]

(図 3.3-1 → 図 3.3-12)

University of Maryland の K. L. Norman, L. J. Weldon そして B. Shneiderman らは、論文“Cognitive Layouts of windows and multiple screens for user interfaces”に於いて、認知心理学的に導き出された“人間が複数の情報を認識理解する7種類のパターン”に基づいた“マルチウィンドウ或はマルチスクリーン上の意味ある表現法 (Cognitive Layout)”を提案している。論文の概要を以下に述べる。

1. Visual Scope

利用者が複数のウィンドウ上に表示された内容を正しく統合的に理解でき、かつ内容を的確に把握できる程度として Visual Scope を定義している。Cognitive Layout の提案は、良好な Visual Scope を実現する事に他ならない。良好な Visual Scope が得られないウィンドウシステム等では、統合的に理解されるべき情報の混乱、意味の不明確性、不適当な情報の位置表示等、数々の不都合が生ずる。

2. Mode of Representation (図 3.3-1)

画面の表示は3種類の Layout、即ち Surface Layout, Machine Layout 及び Cognitive Layout の3種類

の Mode に分類して考える。この内、Visual Scope を実現する上で、最も重要な Mode が Cognitice Layout である。

2-1. Surface Layout

Surface Layout は、概念的には表示物理デバイス上の Layout で、例えば Bit Matrix の分解能、表示可能文字数、ウィンドウのサイズや形態あるいは表現可能な色数等と言った、特別の意味を持たない無機質的な Layout 機能である。Surface Layout の決定に当たっては以下に示す基本的な情報操作が可能な機能を持つ事が必要である。

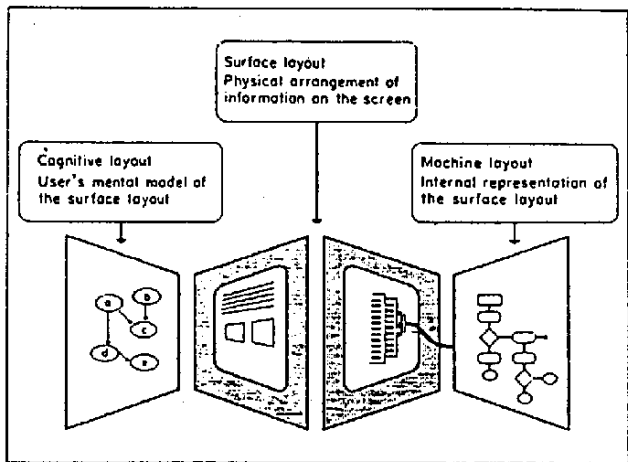


図 3.3-1 Mode of Representation.

1) FUSION (図 3.3-2)

複数のウィンドウ間に渡る長いメッセージ情報の分割形態であり、必要に応じて上下方向の分割、あるいは左右方向の分割が可能な機能である。

2) SYNCHRONIZED AND INDEPENDENT SCROLLING (図 3.3-3)

複数ウィンドウ内のメッセージ情報のスクロールを同期、あるいは非同期で行う機能である。

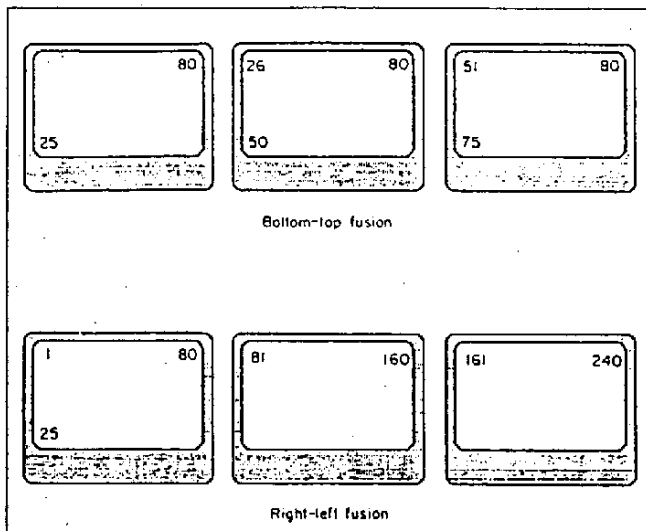


図 3.3-2 Fusion.

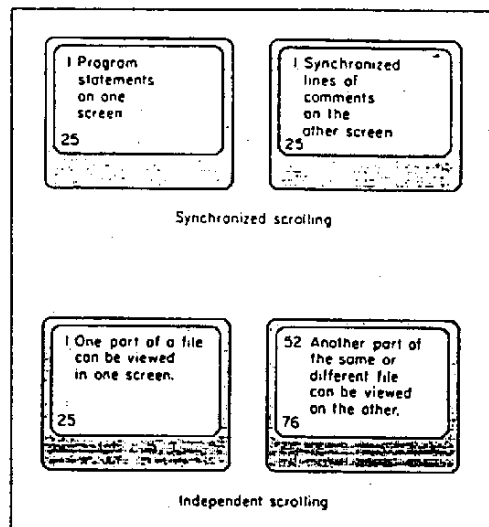


図 3.3-3 Synchronized and Independent Scrolling.

3) TRIGGERED CHANGES (図 3.3-4)

最も基本的な、一つのウィンドウ内の内容を自由に切り替える機能である。

4) SELECTION (図 3.3-5)

TRIGGERED CHANGES の特殊な場合と位置づけており、一つのウィンドウの内容を一定の候補の中から選択する機能である。

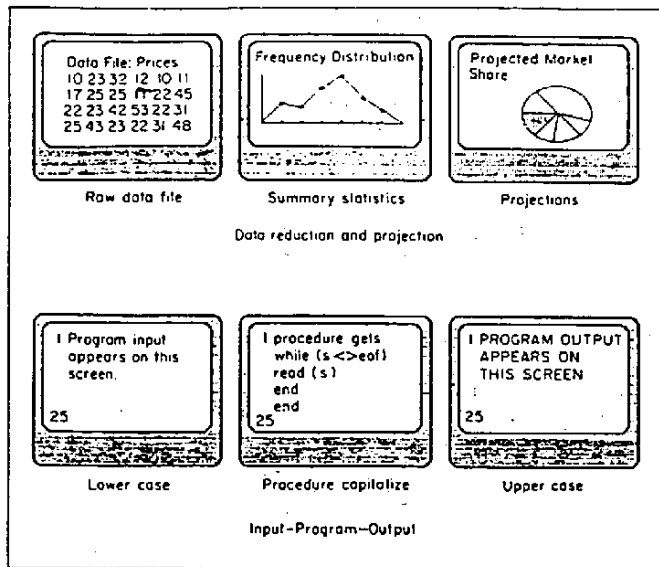


図 3.3-4 Triggered Changes.

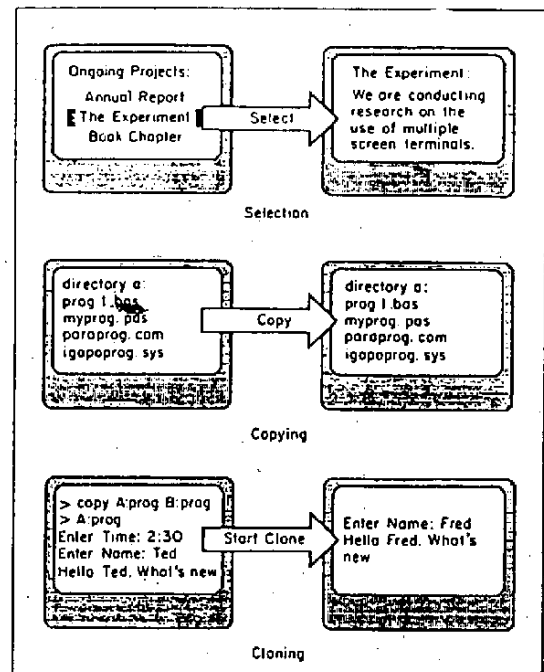


図 3.3-5 Selection, Copy and Clone.

5) COPY

複数のウィンドウ中、現在ワーキングウィンドウとして用いているウィンドウの内容を、参照の為に別のウィンドウにコピーする機能である。

6) CLONE

現在のワーキングウィンドウの内容とまったく同じ内容を、コンカレントジョブとして新たに生成する機能である。

2-2. Machine layout

Machine layout は Surface Layout を実現する為のハードウェア及びO/Sのデバイスドライバ等であり、実際のハードウェアシステムの設計仕様でもある。

2-3. Cognitive layout

Cognitive layout は、利用者の持つ認識モデル、あるいはメンタルモデルに良好に重畳するような画面上の表示形態で、Surface Layout の機能を用いて意味の有る内容表示を行う Layout である。この意味では Cognitive Layout 自身、メンタルモデルの特殊なケースと考えても良い。以下に示されているのは、認知科学的に導き出された意味のある7種類の基本的な表示方法である。

1) LINEAR ARRAY LAYOUT (図 3.3-6)

最も、単純な Cognitive Layout であり、表示情報の複数ウィンドウへの線形空間分割である。必要に応じて、上下方向分割あるいはパノラマ的左右分割等が考えられる。

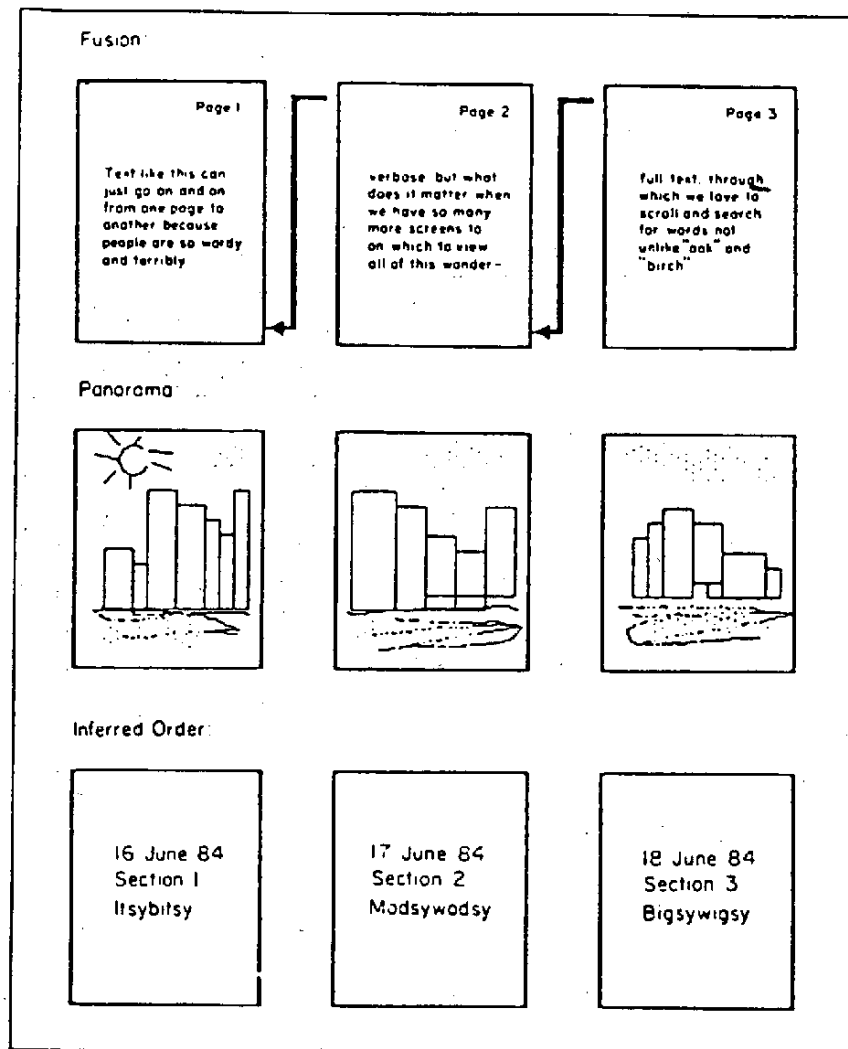


図 3.3-6 Linear Array Layout.

2) INFORMATION INTEGRATION LAYOUT (図 3.3-7)

複数の抜粋された情報の表示により，利用者がやはり表示された特定の問題について問題解決，決断あるいは判断を行うための表示方法である。

3) SELECTIVE ATTENTION LAYOUT (図 3.3-8)

表示されている複数の抜粋された情報の内から，利用者がやはり表示された選択条件を満足する情報を選択する為の表示方法である。

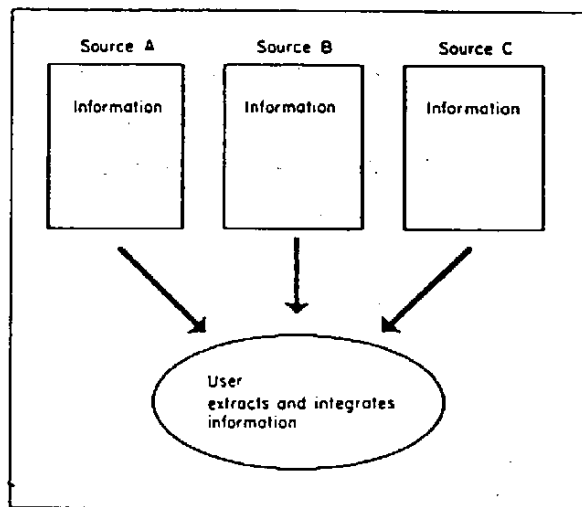


図 3.3-7 Information Integration Layout.

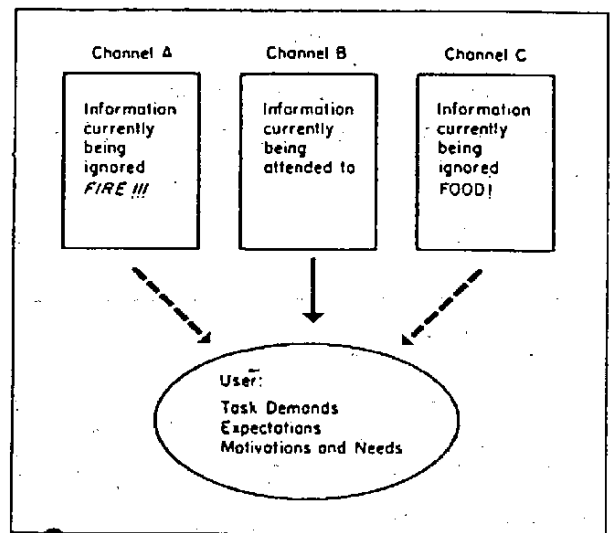


図 3.3-8 Selective attention Layout.

4) LEVELS OF PROCESSING LAYOUT (図 3.3-9)

複数のウィンドウが、連続的な情報の流れを離散的に表示し、利用者にもととの連続的な情報の流れを推論される為の表示方法である。

5) MEMORY STORAGE LAYOUT (図 3.3-10)

Short-term sensory store → Short-term memory → Long-term memory の 3 段階からなる、人間の認知記憶構造モデルに即した応用範囲の広い表示方法である。

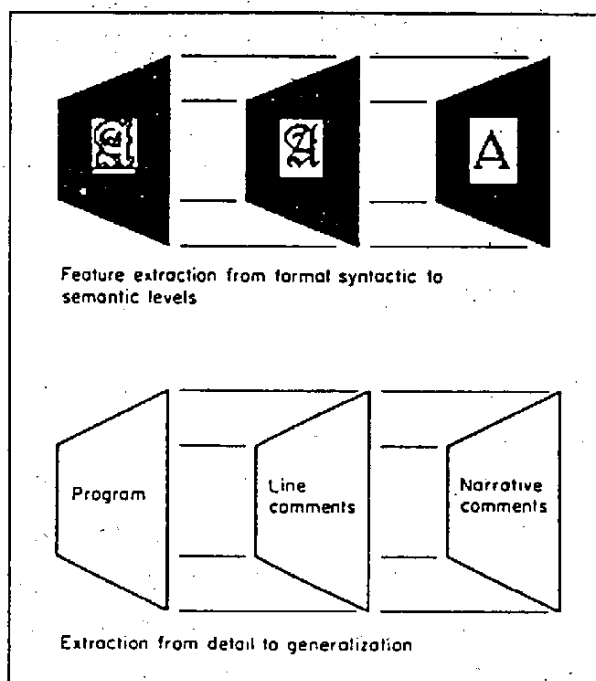


図 3.3-9 Levels of Processing layout.

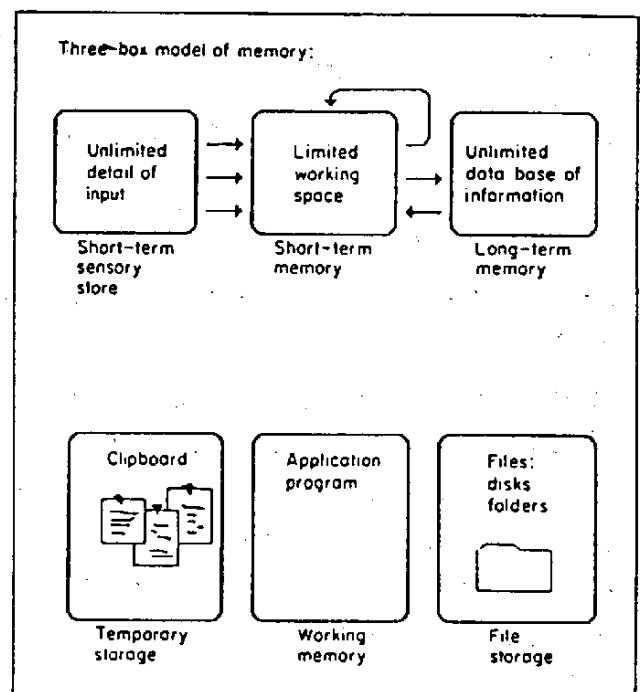


図 3.3-10 Memory Storage Layout.

6) ZOOM IN-ZOOM OUT LAYOUT (図 3.3-11)

視覚的に非常に強力と考えられている、複数ウィンドウによるズームアップを用いた表示方法である。

7) PERSPECTIVE LAYOUT (図 3.3-12)

一つのオブジェクト（主として画像）に対して、複数のウィンドウにより視点を変化させた表示を行い、利用者のオブジェクトの全体像の理解を助ける表示法である。

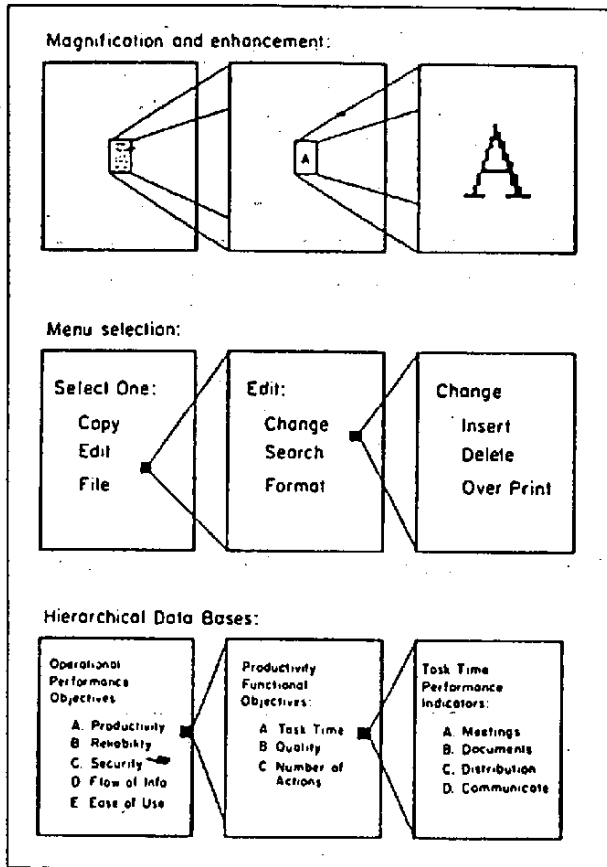


図 3.3-11 Zoom in-Zoom out Layout.

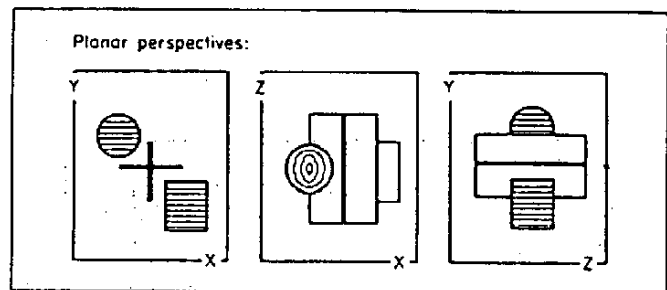


図 3.3-12 Perspective Layout.

3.3.1.3 音あるいは音声によるコンピュータインタフェースの研究

(1) 全般的な研究動向

音響信号とコンピュータの双方に関する研究は、これまで大別して異なった3種類の方向で行われて来たと思える事が出来る。一つは、音自身、音楽或は音声等の特性をコンピュータ技術により解析する試み[7]-[8]であり、次の一つは、目の不自由な人の生活を音とコンピュータを用いて手助けしようとする試み[9]であり、そして最後の一つが、従来から存在しているコンピュータのヒューマンインタフェースに音響信号を導入し、より使いやすいインタフェースを実現しようとする試みである。しかしながら、本調査目的である最後に述べた研究方向は散見できる研究論文も多くなく、他の2種類の研究に比較して十分な研究が行われていないのが現状の様に思われる。本項では、数少ない研究の内から U. C. S. D に於ける音声 ICON に関する研究を取り上げ、以下に概要を報告する。

(2) 音声 ICONS に関する研究例 [10]-[11]

(図 3.3-13→図 3.3-13)

University of California, San Diego 校の William W. Gaver は、論文 "Auditory Icons: Using Sound in Computer Interfaces" により、より利用者により使いやすいインタフェースの実現を目的とした音声 ICON を提

案している。以下に概要を述べる。

1. 伝達情報に対するマッピング (図 3.3-13)

音声 ICON では、伝えるべき情報と出すべき音響音との関連付けを、以下に示す 3 種類のマッピング方式を用いて行う。

Mapping	Visual	Auditory
Symbolic	 for stop	Sirens for approaching ambulance
	 for peace	Applause for approval
Metaphorical	 for organization of relationships	Hiss for snake
	 for horses	Pitch for falling
Nomic	 for scissors	Mailbox sound for arriving mail
	 for tile	Hit wood or metal sound for size of object

図 3.3-13 音声 ICON の意味マッピング。

1) Symbolic Mapping (シンボルのマッピング)

例えば、視覚 ICON が STOP 指示を侵入禁止の標識で表現するように、音声 ICON では握手で賛成すなわち同意を表す様な対応付けである。

2) Metaphorical Mapping (比喩的マッピング)

例えば、視覚 ICON が GO の指示を馬のひずめで表現するように、音声 ICON では馬の鳴き声で同様な指示を行う様な対応付けである。

3) Nomic Mapping (直接的マッピング)

例えば、視覚 ICON では郵便受けの絵でメールの到着を表現する様に、音声 ICON では郵便受けに手紙が入れられた時のボタンと言う音で同様の指示を行う様な対応付けである。

これらの試みは、“人間は日常、周囲からの音を聞くことで多く情報を得ている”と言う事実に基づいている。実際的な音声 ICON は漫画的に自然に発生する音の様なもので、伝えるべき情報と音響音との対応は主として直接的あるいは比喩的に関連付けられている。例えば、コンパイルされたプログラムは金属的な音響音に対応させ、また大容量の状態を重い音響音に対応させる。この場合、コンパイルされた結果のオブジェクトファイルの容量が非常に大容量に成った場合には、大きな金属を叩いたような重い音響音が音声 ICON として発生させられる、と言った具合である。発生させる音響音は、極めて自然な音を出す必要はなく、また最近の疑似音の発生技術の向上により、音声 ICON の実現は極めて容易であると考えられる。

2. 学習の容易性

上記の例で述べたごとく、音声 ICON は従来のピープ音による情報の伝達に比較し、非常に自然であり、

覚えやすい特徴を有している。また、マッピング方式による学習の容易性に関しては、Nomic Mapping → Metaphorical Mapping → Symbolic Mapping の順になると考えられている。

3. 今後の研究課題

今後の研究課題としては以下に述べる3種類の研究内容が考えられる。第1の研究内容は、日常において周囲から伝えられている音の收拾と收拾された音響音の利用価値の有無の検討であり、第2の研究内容は、いかに効果的な疑似音を創り出すかの検討あり、そして第3の研究内容は、直接的な音声ICONと伝えなければならない情報との対応付けが非実用的であった場合に、いかに比喩的な関連付けを創り出すかを検討する事である。

3.3.1.4 ヒューマンインタフェースの評価ガイドライン

(図 3.3-14→図 3.3-15)

ヒューマンインタフェース方式の研究を行うに際し、方式自体の性能を公平かつ正確に評価できる評価基準の存在は重要である。本項では、GMDのEVADIS (Evaluating of Dialogue Systems) と呼ばれるプロジェクトが開発したヒューマンインタフェース特性の評価手続きについて報告する[12]。

```
ITEM No.: 236
CRITERIA: Perspicuity
SYSTEM COMP.: Screen-layout (terminal interface)
SUBSET: 3

ITEM: Are display sectors for user guidance (e.g. help functions)
      separated from those for work data?

COMMENT: Distinctly separated display sectors with different
         meaning avoid user confusion.

ANSWER: - no
        - yes, by split screen
        - yes, by windows
        - yes, by colour
        - yes, by inversion
        - yes, by fonts
        - yes, by others: .....
```

図 3.3-14 評価例 1.

```
ITEM No.: 265
CRITERIA: Transparency
SYSTEM COMP.: status-information (session interface)
SUBSET: 2

ITEM: Does the system indicate to the user whether it expects an
      input or whether it executes processing?

COMMENT: In every situation the user should be enabled to recognise
         whether commands are processed or new input is awaited by
         system.

ANSWER: - Indication about awaited input
        - Indication about current processing
        - No indications
```

図 3.3-15 評価例 2.

(1) GMD-EVADIS プロジェクトの研究例

GMDのEVADISプロジェクトは、論文“A Guideline-oriented procedure for Evaluating interface qualities of Dialogue-Systems (EVADIS)”において、数百種類に及ぶ評価尺度により、対象とするヒューマンインタフェースの特性、性能を総合的に評価し、その結果により、システムデザイナーに対し改良の可能性をフィードバックする事を提案している。数百種類の評価項目は、以下に述べる様に、Structure of

Item-List, Evaluation CriteriaそしてSystem Componentsの3種類の異なった切口の組合せで構成されている。

1. 目的

HCC (Human Computer Communication) における基本的問題を探り、得られた知識により HIC (Human Computer Interface) の設計を行い、また併せて各種の試作機や市販の商用システムの評価も行う事を目的とする。

2. Structure of Item-List

テストアプローチ法による分類であり、以下に述べる5種類のテストアプローチが有る。

1) subest 1 :

マニュアル及びあらかじめ準備されたテストを用いながらの評価。

2) subest 2 :

あらかじめ準備された標準的な対話シーケンスを実行の後の評価。

3) subest 3 :

標準化されたワードプロセッサ等のアプリケーションプログラムの実行後の評価。

4) subest 4 :

標準タスク実行後の評価。

5) subest 5 :

被験者による研究室レベルの評価。

3. Evaluation Criteria

DIN 66.234 規格に基づいた分類であり、以下に述べる9種類の分類がある。

1) Task-Adequancy (タスク妥当性)

O/S等に入用法を強要されず、利用者にとって、自然が感覚で利用できること。

2) Self-Descriptivity (システム自己記述性) :

システム自体が自身の状態を利用者に簡潔に説明出来ること。

3) User-Controllability (利用者制御可能性) :

システム側からの強要が無く、利用者のペースで自在に対話出来ること。

4) Correspondence with user expectations (対利用者期待信頼性) :

利用者が期待した通りのシステム側からのレスポンスが得られること。

5) Fault-Tolerance (フォールトトレランス) :

システム側において、強力なシステム維持機能が得られていること。

6) Transparency (対利用者透明度) :

システムの持つ全ての機能を利用者が容易に知り得、利用出来ること。

7) Learnability (対利用者学習容易性) :

システムの利用法、機能の全容等を利用者が容易に学べること。

8) Perspicuity (対利用者明白性) :

利用者に対する、システム側からの記述がわかりやすいこと。

9) Flexibility (柔軟性) :

出来る限り、個々の利用者の要求に適合していること。

4. System Components

IFIP モデルに基づいた分類であり、以下に述べる 3 種類の評価項目がある。

1) Terminal Interface (入力インタフェース) :

2) Session Interface (対話インタフェース) :

3) Functional Interface (機能的インタフェース) :

利用者からの要求にソフトウェアツールを調和適合させる為のインタフェース。

4) Organizational Interface (組織的インタフェース) :

利用者グループ或は通信グループ内の利用者マネージメントの為のインタフェース。

5. 評価例 (図 14→図 15)

(1) Perspicuity, Terminal Interface に関する項目を Subest 3 で評価する例を図 14 に示す。

(2) Transparency, Session Interface に関する項目を Subest 2 で評価する例を図 15 に示す。

3.3.1.5 ヒューマンインタフェースの設計ガイドライン [13]

U. S 空軍, 計算機資源運用技術計画に基づいて MITRE 社により作成された “利用者インタフェースソフトウェア設計ガイドライン” の内, メッセージ出力に関連する部分の構成の概要を以下に報告する。

(1) 利用者インタフェースソフトウェア設計ガイドラインの紹介

本ガイドラインでは, 利用者インタフェースソフトウェアの設計の為のガイドラインが 6 種類の機能分野に分けて説明されている。6 種類の機能分野は, データ入力, データ表示, 操作手順制御, 利用者案内, データ伝送そしてデータ保護である。本項では, この内, データ表示, 操作手順制御及び利用者案内の構成について以下に述べる。

1) データ表示

(1.0) 一般

(1.1) テキスト

(1.2) データ形式

(1.3) 表

(1.4) 図式

(1.4.1) 目盛

(1.4.2) 散布図

(1.4.3) 曲線と直線のグラフ

(1.4.4) 棒グラフ

(1.4.5) 円グラフ

(1.4.6) 図面と線図

(1.4.7) 流れ図

(1.4.8) 地図と状況表示

(1.5) 書式

(1.6) コード化

(1.7) 表示制御

(1.7.1) 選択

(1.7.2) 枠取り

- (1.7.3) 更新
- (1.7.4) 抑制
- (1.7.5) ウィンドウの重ね合わせ
- (1.8) 設計変更

2) 操作手順制御

- (2.0) 一般
- (2.1) 対話型
 - (2.1.1) 質問応答
 - (2.1.2) 空所補充
 - (2.1.3) メニュー選択
 - (2.1.4) 機能キー
 - (2.1.5) コマンド言語
 - (2.1.6) 問い合わせ言語
 - (2.1.7) 自然言語
 - (2.1.8) 図式対話
- (2.2) 処理選択
- (2.3) 割り込み
- (2.4) 背景の明示
- (2.5) エラー管理
- (2.6) 警告
- (2.7) 設計変更

3) 利用者案内

- (3.0) 一般
- (3.1) 状態情報
- (3.2) 定型的なフィードバック
- (3.3) 誤りのフィードバック
- (3.4) 作業支援
- (3.5) 利用者記録
- (3.6) 設計変更

3.3.1.6 おわりに

以上、メッセージ出力及び音響信号出力にヒューマンインタフェースの研究技術動向について、その調査結果に報告した、メッセージ出力に関する調査では、専門家向けの実用ウィンドウシステムが氾濫している中、人間一般に関する認知心理学的アプローチで地味な研究が続けられている事が判明したのが印象的であった。又、音声を用いた一般的なヒューマンインタフェースの研究例が多くない事実は、少なからず驚きであり、各研究機関の今後の対応に期待したいと考える。

3.3.2 画 像 [14],[15]

(1) はじめに

コンピュータからの出力装置である VDU (Visual Display Unit) は、ユーザへの主要なフィードバックの源となっている。また、VDU の機能、そのなかでも特に表示速度は、ユーザの作業効率、エラー数、そして満足度などヒューマンインタフェースに大きな影響を与えている。したがってディスプレイに対するユーザの期待が非常に高くなってきている。

ディスプレイ制御に関する技術要素 (出力系) を図 3.3-16 に示す。技術要素としては、階層的にハードウェアに近いところから、ディスプレイユニットである表示デバイス、ウィンドウ表現など仮想的表示装置、文字・図形・画像などの出力データの表現の規定、GKS/PHIGS/POSTSCRIPT などディスプレイに表示すべき画面を生成するための AP (アプリケーションプログラム) とのグラフィックス・インタフェースから成る。

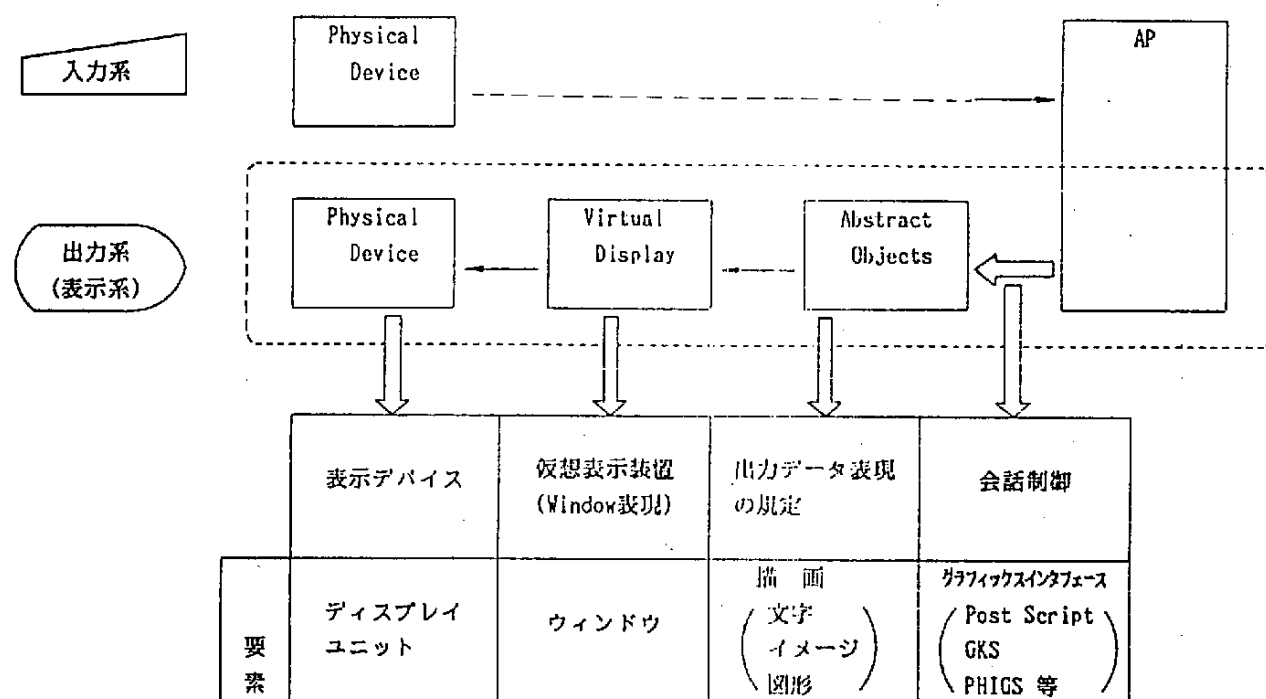


図 3.3-16 ディスプレイ制御の技術要素

ここでは、ヒューマンインタフェースにおける物理的出力ということで、第7章のヒューマンインタフェース関連研究 (画像) と重複しないように、画像の作成、蓄積、検索技術などにはふれず、最もハードウェアよりであるディスプレイ技術とくに高速に画像を生成表示するための画像表示技術を中心に、現状と動向についての調査結果を報告する。

(2) 現 状

1) ディスプレイ制御の現状

図 3.3-17 はディスプレイ制御の現状を示したものである。装置としては、ユーザが要求するアプリケーションに必要な機能を実現するためのディスプレイ制御技術にしたがって、パソコン、OWS (オフィス・ワークステーション)、EWS (エンジニアリング・ワークステーション)、そして GWS (グラフィックス・

ワークステーション)に区分している。ディスプレイ制御技術は前者から後者の装置にいくにしたがって高い機能・性能が要求されている。図中では、これらの装置に対応した機能・アプリケーション、ハードウェアとしてCRTの解像度・表示色・構成要素の現状を示す。

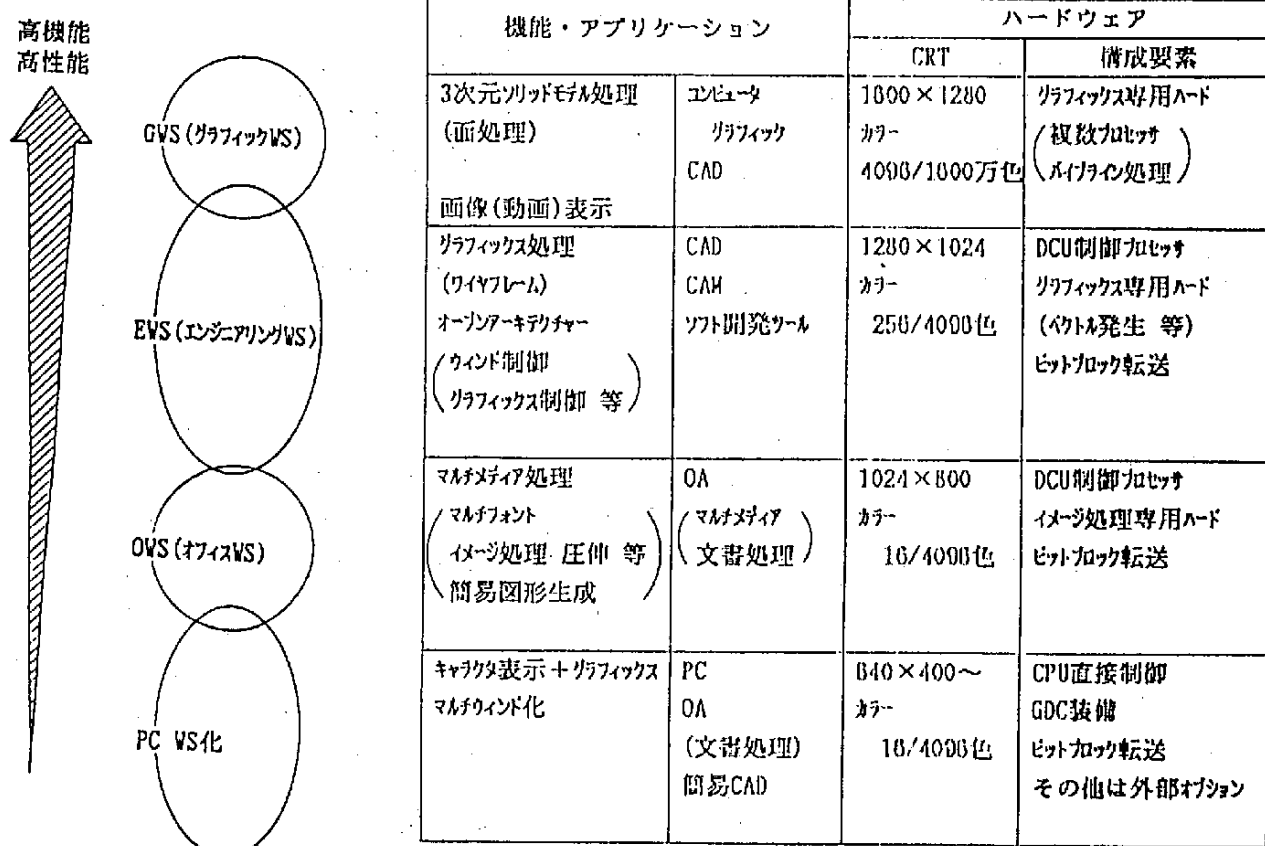


図 3.3-17 ディスプレイ制御の現状

2) グラフィックス・ワークステーションの動向

画像を扱う分野は図 3.3-18 に示すようにビジネス・グラフィックス程度のものから CAD/CAM, ART, そしてビジュアルシミュレーションまで様々である。ユーザが美しい画像を要求する分野においては、装置の高いグラフィックス(画素生成)能力が必要であり、モデルの正確さを要求する分野においては、装置の高い数値計算能力が必要とされる。表示速度がユーザに与える影響が大であることを考慮すると、画像表示装置は利用分野に対応した数値計算能力、画素生成能力を有することが不可欠である。

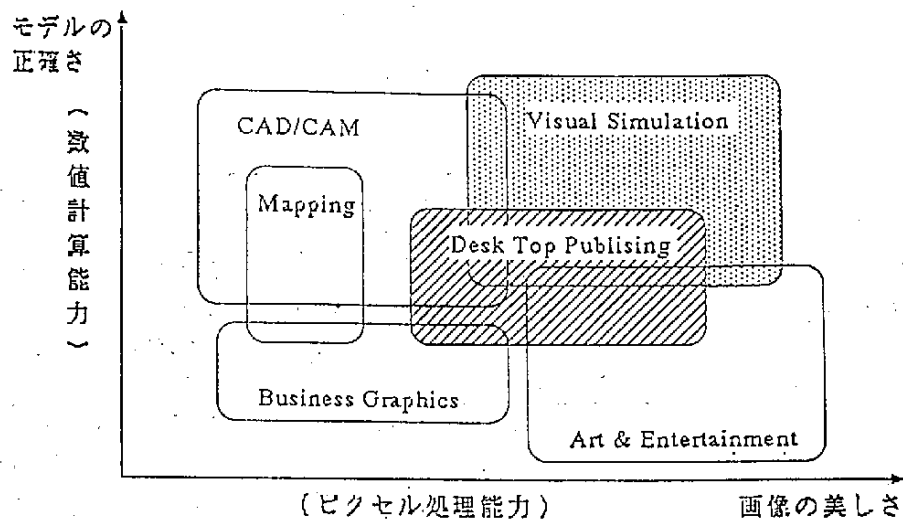
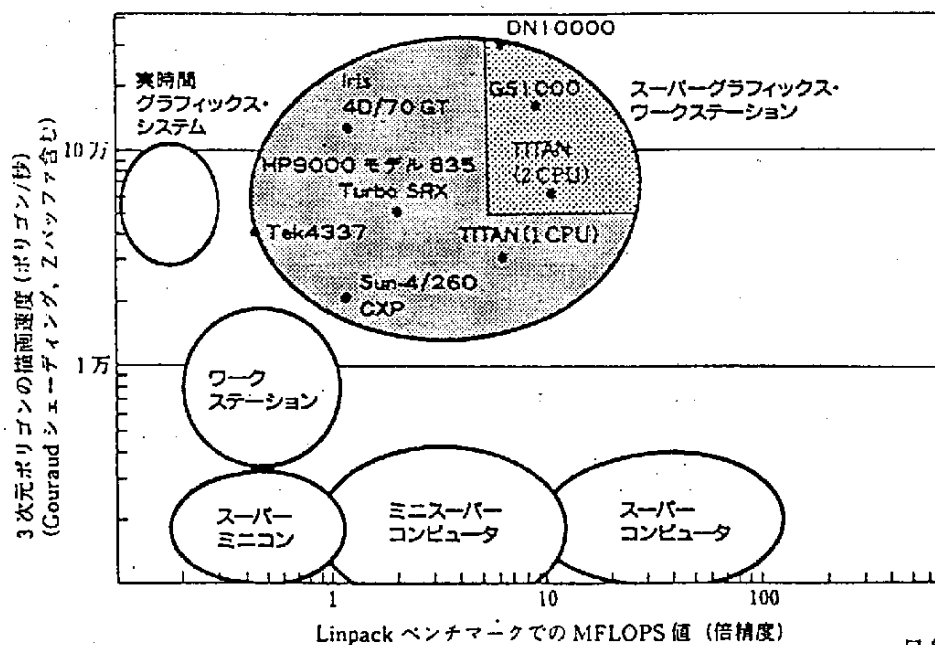


図 3.3-18 利用分野

最近、数値計算能力、グラフィックス能力の両方で従来機種を1桁上回るスーパーグラフィックス・ワークステーションと称されるものが製品化されている。図 3.3-19 に示すように、その性能は10MIPS以上、浮動小数点演算が5 MFLOPS、3次元描画が5万ポリゴン/秒以上のものとなっている。これらの製品は、科学分野における計算結果の視覚的表示（ビジュアライゼーション）を目的としている。代表的なグラフィックス・ワークステーションの性能を図 3.3-20 に示す。



日経エレクトロニクス

1988. 5. 2号より引用

図 3.3-19 スーパーグラフィックス・ワークステーションの位置付け (図中の )

機種名	DN10000 (DSP10000)	TITAN	HP9000 モデル B35 Turbo SRX
開発元	米 Apollo Computer Inc.	米 Ardent Computer Corp.	米 Hewlett Packard Co.
CPU の構成	整数プロセッサ (4万ゲートの RISC プロセッサ), 浮動小数点プロセッサ (ALU, 乗算器などで構成), メモリ管理ユニットなどを1ボードに搭載	整数プロセッサは R2000 (ミッパス社製), 64ビット・ベクトル・プロセッサ などを1ボードに搭載	クロック 30 MHz で動作する RISC CPU 〔整数プロセッサと浮動小数点 プロセッサ〕
整数演算の性能**	40 MIPS(ピーク)	16 MIPS(ピーク)	14 MIPS(ピーク)
浮動小数点演算の性能	5 MFLOPS (Linpack, 倍精度)	6 MFLOPS (Linpack, 倍精度)	2.02 MFLOPS (Linpack, 倍精度)
キャッシュ容量	データ: 64 K バイト 命令: 128 K バイト	データ: 16 K バイト 命令: 16 K バイト	128 K バイト
主記憶容量	8~128 M バイト	8~128 M バイト	8~96 M バイト
マルチプロセッサ構成	最大 4 台の CPU (マルチプロセッサ処理)	最大 4 台の CPU (並列処理機能あり)	不可
内部バス	64 ビット・データ, 転送速度 150 M バイト/秒	64 ビット・バス×2 本, 転送速度 256 M バイト/秒	データ 32 ビット, 転送速度 18.6 M バイト/秒のバス
グラフィックス・プロセッサ	RISC プロセッサ搭載の専用システム (開発中)**	CPU で処理	専用プロセッサを搭載した アクセラレータ**
3次元ベクトル描画速度	40 万ベクトル/秒の見込み	40 万ベクトル/秒 (ベクトルは 12 画素)	24 万ベクトル/秒 (ベクトルは 10 画素)
3次元ポリゴン描画速度**	30 万ポリゴン/秒の見込み	7 万 5000 ポリゴン/秒 (ポリゴンは 40 画素)	5 万ポリゴン/秒 (ポリゴンは 50 画素)
グラフィックス・ライブラリ	PHIGS	Doré	Starbase
フレーム・バッファ構成	80 面 (グラフィックス 24 面, Z バッファ 32 面など)	48 面 (グラフィックス 24 面, Z バッファ 16 面など)	44 面 (グラフィックス 24 面, Z バッファ 16 面など)
OS	Domain/OS	Unix System V リリース 3.0 準拠	HP-UX

日経エレクトロニクス

1988. 5. 2号より引用

図 3.3-20 (1/2) グラフィックス・ワークステーション

機種名	IRIS 4D/70 GT	GS1000	Sun 4/260 CXP	4330 シリーズ
開発元	米 Silicon Graphics Computer Systems Inc.	米 Stellar Computer Inc.	米 Sun Microsystems, Inc.	米 Tektronix, Inc.
CPU の構成	整数プロセッサ R2000, 浮動小数点プロセッサ R2010 (クロックはいずれも 12.5 MHz)	マルチストリーム型のバイプライン・ プロセッサ, 64 ビット・ ベクトル・プロセッサなどで構成	RISC プロセッサ SPARC 浮動小数点コプロセッサ (Weitek 製の 1164/1165)	68020 (20 MHz), 68881 浮動小数点アクセラレータは オプション (98 万円)
整数演算の性能**	10 MIPS(ピーク)	20~25 MIPS(ピーク)	10 MIPS(ピーク)	2.5 MIPS 程度
浮動小数点演算の性能	1.12 MFLOPS (Linpack, 倍精度)	8.6 MFLOPS (Linpack, 倍精度)	1.1 MFLOPS (Linpack, 倍精度)	330 kFLOPS (Linpack, 倍精度)**
キャッシュ容量	データ: 32 K バイト, 命令: 64 K バイト	1 M バイト	128 K バイト	なし
主記憶容量	8~16 M バイト	16~128 M バイト	8~128 M バイト	4~12 M バイト
マルチプロセッサ構成	不可	不可	不可	不可
内部バス	VMEbus	512 ビット幅の スイッチ・ネットワーク	VMEbus	VMEbus
グラフィックス・プロセッサ	約 50 個の専用 LSI など構成	専用のレンダリング・ プロセッサ	アクセラレータ GP2	68020 (16 MHz), 専用 LSI など構成
3次元ベクトル描画速度	40 万ベクトル/秒 (ベクトルは 10 画素)	60 万ベクトル/秒 (ベクトルは 10 画素)	15 万ベクトル/秒 (ベクトルは 10 画素)	34 万ベクトル/秒 (ベクトルは 10 画素)
3次元ポリゴン描画速度**	12 万ポリゴン/秒 (ポリゴンは 100 画素)	15 万ポリゴン/秒 (ポリゴンは 100 画素)	2 万ポリゴン/秒 (ポリゴンは 100 画素)	2 万ポリゴン/秒 (ポリゴンは 200 画素)
グラフィックス・ライブラリ	Graphics Library	Stellar Vision**	SunCore, SunGKS, SunCGI	PLOT10 STI
フレーム・バッファ構成	96 面 (グラフィックス 48 面, Z バッファ 24 面など)	16 面あるいは 32 面	32 面 (グラフィックス 16 面, Z バッファ 16 面)	12 面あるいは 24 面
OS	Unix System V リリース 3.0 準拠	Stellix	SunOS	Utek

日経エレクトロニクス

1988. 5. 2号より引用

図 3.3-20 (2/2) グラフィックス・ワークステーション

これらの製品は主として、浮動小数点の計算能力とグラフィックス描画速度、そのどちらかに力をいれているものが多い。グラフィックス能力に重点をおいている装置に関しては、グラフィックス専用のプロセッサを搭載することにより高速化を図っているものが多い。

(3) 技術動向

1) グラフィックコントローラの動向

図 3.3-21 にディスプレイ装置で使用されている CRTC (CRT Controller) / グラフィックコントローラの動向を示す。初期の単に CRT に対する表示タイミングを生成する機能のものから、ビジネスグラフィックス用などの図形を描画するもの、ファクシミリ用などのイメージの圧縮伸長用のもの、イメージの密度変換 / 回転などを行うもの、そのしてそれらの複合された機能を有するものへと機能 / 性能の向上化が図られている。

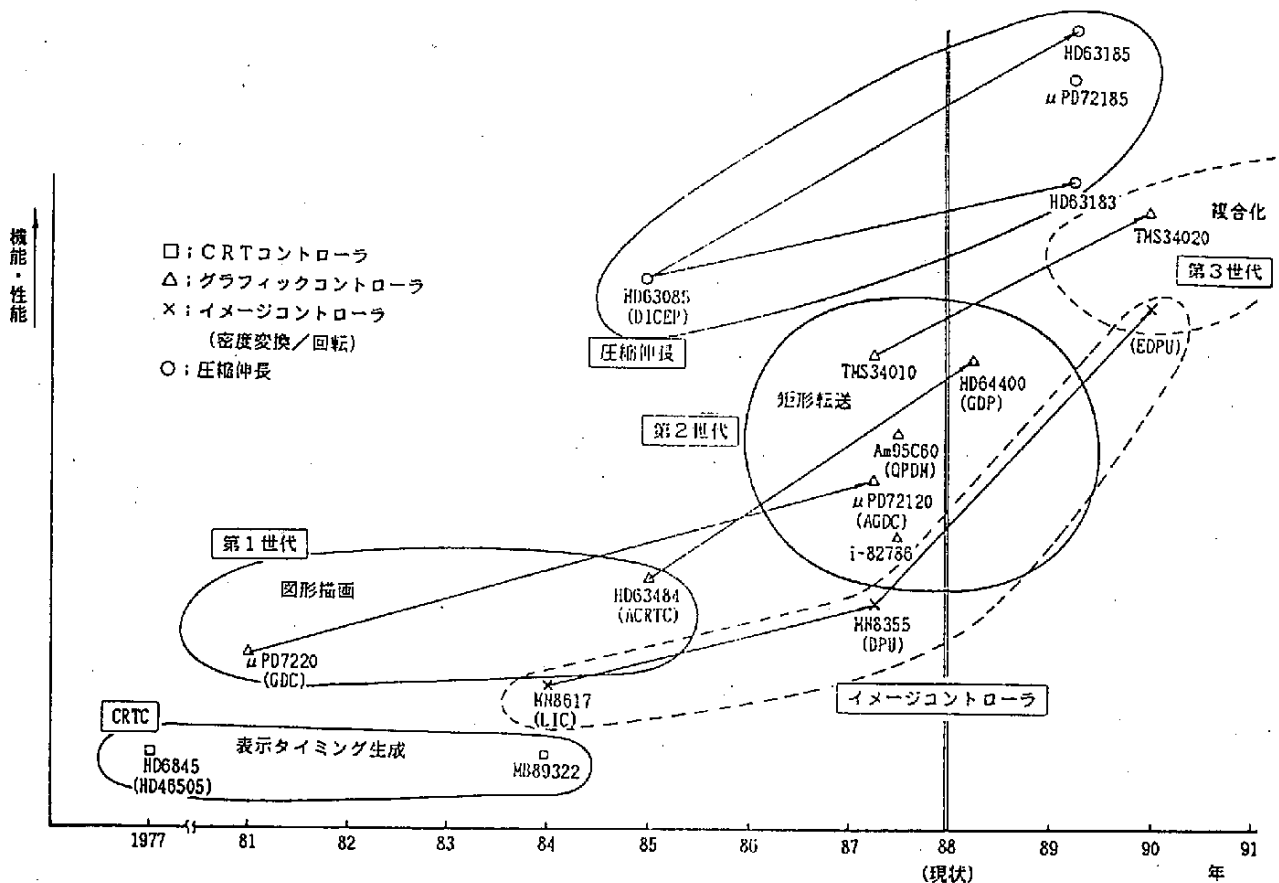


図 3.3-21 CRTC/グラフィックコントローラの動向

2) ディスプレイ技術の動向

図 3.3-22 にディスプレイの区分 (GWS: グラフィックスワークステーション, EWS: エンジニアリングワークステーション, OWS: オフィスワークステーション, PC: パーソナルコンピュータの下位・中上位機) 毎に、前述したディスプレイ技術 (ディスプレイユニット, 出力データの表現) に関して性能・機能の動向を示す。この図は、現在発表されている製品の調査結果に基づく。

G : GWS (グラフィックワークステーション)
 E : EWS (エンジニアリングワークステーション)
 O : OWS (オフィスワークステーション)
 P_H : PC (パーソナルコンピュータ) 中・上位機
 P_L : PC (パーソナルコンピュータ) 下位機

○印は、将来矢印の方向へ移行
 ○印無しは、現状維持

項 目		機 能 ・ 性 能			
		低	高		
ディスプレイ ユニット	表示ドット数	640×400ドットクラス	1024×768ドットクラス	1280×1024ドットクラス	2000×2000ドットクラス以上
		P _L	P _H ○ →	E ○ →	G →
同表表示色	16色 以下	256 ~ 4096色	~ 25万色	~ 1670万色	
		P _L ○ → O ○ →	E →	G →	
文 字	フォント種	1書体・1サイズ	1書体・数サイズ	数書体・数サイズ	
		P _L , P _H ○ →	E, O ○ →		
文字描画速度	4~5秒/画面	1~2秒/画面	0.1~0.3秒/画面		
		P _L	G, P _H ○ → E ○ →		
グラフィックス	図形描画速度 (短ベクトル (10ドット/ベクトル) を描画する速度)	1万未満 短ベクトル/秒	1万~2万 短ベクトル/秒	~10万 短ベクトル/秒	40万~50万 短ベクトル/秒
		P _L ○ → P _H ○ →	O →	E →	G →
図形操作速度 (拡大/縮小 移動、回転)	1未満 画面/秒	約1画面/秒	5~10 画面/秒	10~30 画面/秒	
		P _L ○ → P _H ○ →	O → E ○ →	G ○ →	図形が断続的に動く 図形が動的に動く
塗りつぶし速度 100ピクセル四角形 ポリゴンの塗りつぶし速度	2000 ~ 2500 秒/秒	5000 ~ 10000 秒/秒	~ 100000 秒/秒		
	(約4~5秒/画面 相当)	(約1~2秒/画面 相当)	(約0.1秒/画面 相当)		
塗りつぶし処理の 高機能化 (シェーディング、 ハイライト)	無 し	有 り			
		P _L , P _H , O, E	G ○ →		
座 標 変 換	無 し	2次元	3次元		
		P _L ○ → O ○ →	E ○ →	G →	
イ メ ー ジ	イメージデータの 高速ハンドリング機能 (圧縮/伸長 密度変換、回転)	無 し	有 り (ソフト)	有 り (ハード)	
		P _L G	P _H ○ →	E, O →	
イメージデータ 入力機器接続	無 し	イメージ入力 レベル	階調イメージ入力 レベル		
		P _L	P _H ○ → O ○ → E ○ →		

図 3.3-22 ディスプレイ技術の動向

3) ディスプレイユニットの動向

図3.3.23にディスプレイユニット (CRT, LCD, PDP) の予想される性能・機能の動向を示す。将来さらに高解像度化, カラー化, 大量データの即時表示処理の要求が高くなるゆえ, より高画質, 高速化への対応が必要であろう。

		現 状	1～2年後	3～5年後
CRT	EWS	1280 × 1021	1600 × 1280	2000 × 1600
	OWS	1148 × 754	1148 × 758	1280 × 1024
	PC _B	720 × 512	720 × 512	720 × 512
LCD		640×400ドットクラス モノクロ表示	640×400～800×600 ドットクラス 階調表示, カラー表示	1000×1000ドットクラス フルカラー表示
PDP		640×400ドットクラス 階調表示 (DC型, AC型)	1000×1000ドットクラス 階調表示	640×400ドットクラス カラー化

図 3.3-23 ディスプレイユニットの動向

(4) おわりに

以上, 画像表示装置について, その調査結果を報告した。画像表示装置はハードウェアの技術開発が顕著であり, 将来さらに高解像度化, 高速表示化が進であろう。今回は未調査であったが, 今後ヒューマンインタフェースを考えていくうえで, 現在まだユーザに対して困難を強いている画像生成を容易にするためのユーザインタフェースについても調査が必要であろう。

参考文献

- [1] 上林, ワークステーションインタフェースの構成技術, 情報処理学会誌, Vol. 25, No. 2, pp. 103-110 (Feb. 1984).
- [2] Smith, D. C., Harslem, E. F., Irby, C. H., Kimball, R. B. and Verplank, W. L.: Designing the Star User Interface, Byte, pp. 242-282 (Apr. 1982).
- [3] 淵一博, インタフェースの科学, 共立出版, pp. 111-130 (Apr. 1987).
- [4] Xerox PARC Learning Research Group: Special Issue on Smalltalk, Byte, Vol. 6, pp. 168-194 (1981).
- [5] Sun Microsystems, Inc., Programmer's Reference Manual for SunWindows the SunWindow System, Revision D.
- [6] Kent, L. N., Linda, J. W. and, Ber, S., Cognitive Layout of Windows and Multiple Screen for User Interfaces, Int. J. Man-Machine Studies, No. 25, pp. 229-248 (1986).
- [7] Loy, G., musicians Makes a standard: The MIDI phenomenon, Computer Music journal, 9①, pp. 8-26.
- [8] Dowling, W. J., Melodic Information Processing and Its development, The Psychology of Music, Orlando, Florida: Academic Press. pp. 413-428
- [9] Lunney, D., Morrison, R. C., Cetera, M. M., Hartness, R. V., Mills, R. T., Salt, A. D. and Sowell, I., A microcomputer-based laboratory aid for visually impaired students, IEEE Micro, 3①.
- [10] William, W. G., Auditory Icons: Using Sound in Computer Interfaces, Human-Computer Interaction, Vol. 2,

pp. 167-177 (1986).

- [11] Merra, M. B., Denise, A. S., Robert, M. G., Earcon and Icons : Their Structure and common design principles, Human-Computer Interaction.
- [12] R. Oppermann, B. Murchner, M. Paetau, M. Piepper, H. Simm and I. Stellmacher, Evaluation von Dialogsystemen : der software-ergonomische Leitfaden EVADIS, Walter de Gruyter, Berlin (1988).
- [13] Sidney, L. S., Jane, N. M., Guidelines for Designing User Interface Software, MITRE (Aug. 1986).
- [14] Ben Shneiderman 著, 東 基衛, 井関 治監訳, ユーザ・インタフェースの設計, 日経マグローヒル社 (1987).
- [15] 日経エレクトロニクス, No. 446, 1988.5.2

3.4 マニュアル（印刷物，オンライン）

3.4.1 マニュアルの効用

マニュアルの大切さについては、既に米国 MIL 規定（MIL-HDBK-63038-1，技術マニュアルの書き方等）で技術進歩の激しい対象に、且つ多様なバックグラウンドを持つ人々の為の、マニュアルを統一的に分かりやすくする努力がなされている。マクドナルドを初めとしたマニュアルによる、地域、従業員の違いに左右されない、統一された品質及びサービスが、主として外食産業でその経営戦略上、重要視されてきた。日本でもスーパーマーケット、ファーストフード等を中心とした産業でいち早く、この傾向を取り入れて、成功してきた事が良く知られている。

一般の企業でも、国際化が進むに連れて、多様なユーザの要求に対応できるマニュアルの大切さが、広く認識される様になってきた。この事はプロダクト・ライヤビリティにかかわる狭い意味の製品のマニュアルのみにとどまらず、工場のオペレーションまたはプロセス・マニュアルやさらに、ソフトウェアのマニュアルにいたるひろい範囲にかかわるものである。米国では、はやくから、テクニカルライター、インフォメーション・デベロッパー、テクニカル・コミュニケーション等々、マニュアルの専門家が職業人としての独自の地位を認められ、その技量の研究、向上を図ってきた。日本でも、最近その動きが具体化しはじめ、一例として STC（Society for Technical Communication）東京支部が設立されているし、その他研究会、フォーラムが多数あり、またこれら相互の交流も持たれる様になってきつつある。

マニュアルを辞書で引くと、手から派生したハンド・ブック、中世に教会でつかった祈禱書、武器操作法等の記述が目につく。古代中国による紙及び文字の発明、更に下って版木の発明により書物が出来るに至ったことが知られている。更にこの紙の製造技術が西欧に伝来され、それまでに存在したエジプトのパピルスより高度な紙質をもたらし、西欧文明の基盤、ドキュメントおよびその伝達手段を与える事になる。

このドキュメントあるいは書物のなかでも、要点のみの抜き書き、挿絵等がはいり分かりやすいハンディなカタチになった、ブックレット形式のものが実用的な価値の高いものとして、ひろく使用された事が推定される。

3.4.2 日本におけるマニュアルの歴史

日本にも、仏教伝来とともに、経典、その他の文明が、中国または朝鮮半島を経路としてもたらされた。その後、仏教文明および帰化人のふかい影響のもとに王朝の成立と、それをささえる貴族階級の発展とともに、世界最初の小説といわれる、紫式部の「源氏物語」が創作された。

勿論、仏典の中には、因果経絵巻の様に経文と絵がフュージョンされた分かりやすいものが、早くから存在した。平安時代後期に作られた源氏物語絵巻が、日本で作られた、詞書（ことばがき）と絵が連携した、ハイパーテキスト（後述）と言えよう。一方国家のマニュアルとも言うべき十七条の憲法が、聖徳太子により制定され、その他律令（法律マニュアル）が次々に整備され公布された。より簡単なメモとして、木簡が使用されていた事も知られている。この木簡は貢ぎものの記録、当時の物の値段等が書かれ、その時代の DTP（後述）といえよう。

その後、武士階級の勃興、戦国時代を経て安土・桃山時代に発展し、堺には豪商といわれる階級も出現し、豊臣秀吉の天下統一とあいまって日本固有の文化の基礎が築かれた以上、日本国の成立と発展は、主として王朝貴族または武士階級により、支配された。この時代に連歌、百人一首等、連想的に探索するゲームが完成している。



図 3.4-1 源氏物語 竹河

キャプション：日本におけるハイパーテキストの原型？

原典：平凡社世界名画全集 19 日本絵巻物

p 35 昭和 36 年 4 月 30 日

庶民レベルでのマニュアルは、町人文化の花が開く、江戸時代まで待たなければならない。ただこの時代の寺子屋は、その規模、内容とも高度であった事が再評価されている。藩校または石田梅岩を祖とする心学講舎は、全国的に展開されていた。そこで使われる教科書、マニュアルは、挿絵を含むテキストからなり、版行されたものであった。その他、井原西鶴の「世間胸算用」、松尾芭蕉の「奥の細道」等が版行され、また版木の技術を極致にまで高めた浮世絵の出版も盛んに行われ、後の近代国家化の基盤となった。

明治維新を契機に、日本は一挙に西欧文化を取り入れ、活字印刷による新聞・書籍の出版が広く、なされる事になり、その後二度の世界大戦を経て、通信・交通手段の発達により、国際化時代に突入し、現在に至っている。日本語の世界においても、ワード・プロセッサの発展により個人のレベルでも創作または編集がより身近なものとなり、テキストに加えて簡単なイメージは誰でも容易に入力できるようになった。また一部ではあるが、音声認識やパターン認識が、貨物の配送先の仕分けや郵便番号の識別に実用化されている。

3.4.3 編集の新しい価値

この絶え間無く、進歩し続けている情報社会に於て、情報の創造とは何であるかを、編集の価値を中心に考える[1]。現代社会において進行しつつある一連の情報課程は、通常四つの段階にわけて考えられている。それは、第一は情報の創造あるいは生産、第二は情報の処理、第三は情報の伝達、第四は情報の蓄積である。第二の情報の処理とは、主としてコンピュータによる処理の事と考えられ、技術的にも急速な進歩をとげつつある分野である。第三の情報の伝達については、様々なメディアによる可能性が開発されて、その進歩はとどまる所を知らないようにみえる。第四の情報蓄積については、巨大なデータ・ベースの集積が進みつつある。問題は第一の情報の創造のところにある。この点は天才の神秘的な創造作用のほかは、あとはいっきょに平俗な日常業務のなかにおこまれてしまっている。昔は情報の生産と伝達のあいだには、大きな距離はなく編集者という技術者集団が必要とされることも無かった。

現代では、さまざまなマス・メディアをはじめ、巨大な情報手段の開発が進行しつつある。

発想されたイメージを、情報手段を駆使しつつメディアにおとすときには、技術が必要である。

その作業は、きわめて創造的なものである。ここに一種の創造的技術者集団としての編集者が成立し、創造的作業としての編集という仕事ははじまったのである。

編集という行為は、まさに現代の情報産業社会の夜明けを象徴するものである。情報編集の問題は、内容的には、組織における企画ないしは経営の問題に、原理的に共通するところが多いのである。

3.4.4 デスクトップ・パブリッシング (DTP)

この編集という新しい価値観は、ハードウェアとソフトウェアの進歩にも助けられて、拡大する。米国で新しいブームが起こりつつあり、デスクトップ・パブリッシング、エレクトロニック・パブリッシング、コーポレート・パブリッシングなどの言葉は、すでに日常用語となっている。デスクトップ・パブリッシングという言葉がデビューしたのは、Apple 社の Macintosh 用の Laser Writer というレーザー・プリンター (85 年) の発表とともにである [2]。その源流は、米 Xerox 社のパロ・アルト研究所である。そこでは、高解像度ビットマップ・ディスプレイ、マウス、LAN などの現在のワークステーション、パーソナル・コンピュータの基礎技術が開発された。続いて開発された、米 Xerox 社の Star は、

- ・アイコンなどコマンドでは無く図形による指示
- ・必要最小限の情報の提示 (複雑な機能は徐々に)
- ・マニュアルの読み直しを避けた統一されたコマンド
- ・WYSIWYG (What You See Is What You Get)

などの四つの重要な概念を製品のなかに取り入れていた。この考え方は、Apple の Macintosh へと受け継がれ、さらに高性能のワークステーションにも引き継がれる。さらに、パブリッシングの出力は、印刷物にかぎらずに、電子的に入力、編集、制作したテキストやグラフィックスを、電子的に配付することも、示唆している。ニューメディア、電子メディアの概念も含まれている。



図 3.4-2 DTP の出力の一例

キャプション：LightSpeed Design Workstation のレイアウト

原典：日経 EP 6/22/87 p 172

日本では、フォント及び日本語の違いから、やや遅れてスタートしたが、日本語ワープロ、日本語コンピュータの進歩によりこのエリアも発展してきている。最近の新製品の特徴は、

- ・低価格の専用ワープロでも画像を扱えるようになった
 - ・米国のデスクトップ・パブリッシング・ソフトウェアの日本語化が始まった
 - ・パソコンの高機能化としての各種デスクトップ・パブリッシング・ソフトウェアの開発
 - ・漢字フォントや日本語を扱えるページ記述言語の標準化の動き
- などである。

製造業、特に航空宇宙、自動車、軍事、エレクトロニクス、およびコンピュータ業界の各企業、エンジニアリング企業、そして科学研究機関などは、電子出版システムをオペレーター・マニュアルや、修理・保守マニュアル、パーツカタログ、そして価格表の作成などに利用する利点を、商業出版産業以外で認識した最初のユーザ・グループであった。こうした企業にはすでに、集中化した技術文書作成部門、編集業務、グラフィックス部門、あるいは出版グループが存在していることが多いので、電子出版システムの導入は容易である。

技術文書は、一般的な企業内出版とは以下の点で異なっているのが特徴である。

- ・数百ページの長さになりうる
- ・データベース全体で、何十あるいは何百もの文書が存在しうる
- ・頻繁に改訂が行われ、長期の絶え間ない保守が必要
- ・通常、データベースの40%程度までが改訂中の状態にある
- ・CADシステムによるグラフィックスとテキストを組み合わせる必要がある
- ・しばしば複合ページ番号や複雑な索引付けが必要

このような種類の文書作成にあたって極めて役立つ電子出版の機能としては、ページ、段落、図、表、脚注、および引用文献の番号を自動的に変更する機能、イラスト、表および脚注をテキストの特定の場所に固定する機能（これにより、固定されたものは、テキストとともに移動し、番号変更も自動的に行われる）、変更箇所表示（テキスト外側の縦の線）機能、最後まで修正を加えられる機能などがある。この種の環境下では、大型から中型の専用マルチユーザシステムが普通であり、高性能レーザープリンターや印刷機用のタイプセッターが出力として付加されている。

このように電子化された技術文書作成システムを有効利用して、製本されたマニュアルの在庫を極小に抑えた、オン・デマンド・プリンティングの例が報告されている[3]。多種多様な製品マニュアルを、海外用多言語にわたり在庫させることは、改訂がたとえ無いとしても、企業負担がおもく、また実際の個々のオーダーは小量であることが多い。この例の電子出版システムは、機械翻訳システム、ワード・プロセッサ、イメージ・プロセッサおよびこれらの最終段としての出版システムからなりたって、電子的出力はマグネティック・テープなどに記録される。オーダーを受けた時点でレーザー・プリンターなどの出力それに続いた自動製本プロセスを経て、顧客先に発送される。

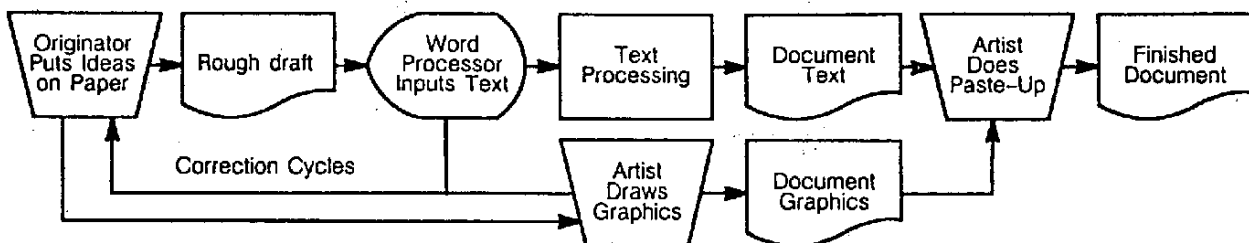
日本でのデスクトップ・パブリッシングを取り巻く環境は、この1-2年で大きく変貌した[4]。米国でデスクトップ・パブリッシングが普及し始めた86年ごろ日本語DTPの普及にはフォントの取り扱いや、処理速度など技術面の問題がネックになるとみられていた。しかし、ベクトル・フォントの採用、ハードの性能向上、レーザー・プリンターの普及などにより、急速に解決されつつある。

このため今では、DTP普及のキープポイントとして素人向けの使いやすさが重要視されるようになってきた。素人が簡単にレイアウト機能を使えるようにした標準的なレイアウト・フォームを用意し、操作性を改善している。ただし、DTP専用機にくらべてレイアウト機能、フォントの種類等でまだ開きがあるが、ワープロの機能拡大のながれにのった、動きがみられる。

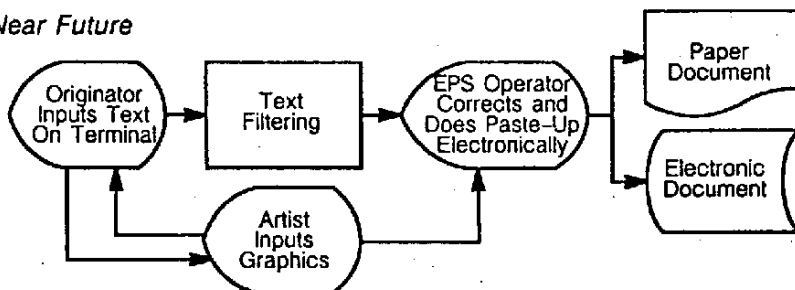
コーポレート・パブリッシングに関しては、ハードウェアの進歩とともに、ユーザが専門家ではなく、素人になる傾向にある[5]。この結果、出版の数は増えるが、デザインとコミュニケーションの質が落ちることもある。また現在、近い将来とも、テキストはその初めの創作者の部門により直接パブリッシュされ、セントライズされたサービス部門には、あまり仕事を依頼しない傾向にある。近い将来には、グラフィックスのような洗練された仕事まで、インテリジェントな機械の進歩により部門内で行われてしまう。

Documentation Workflow

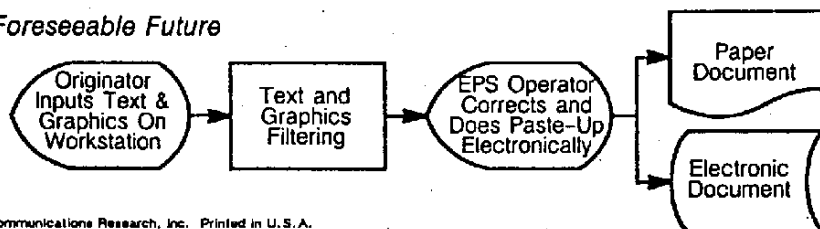
In the Recent Past



In the Near Future



In the Foreseeable Future



Symbol Key:

Manual Process

On-line Process

Computer Process (Batch)

Paper

Electronic Storage

EPS: Electronic Publishing System

© 1987 by Bell Communications Research, Inc. Printed in U.S.A.

図 3.4-3 ドキュメント作成法の動向

キャプション：過去・近未来・将来

原典：35th ITCC ATA-143

3.4.5 ハイパーメディアの発展

次世代または、将来のパーソナル・コンピュータにとっても、製品を支えその性能を引き出すマニュアルそれ自体の、使い勝手の良さが強く要求される。コンピュータの専門家、まれに使用する素人に近い利用者、または子供に至る多様なユーザ層にとって、マニュアルの使い勝手の良さとは、ユーザのレベルに合わせて「疑問点のみを、素早く、分かり易く教えてくれる事」であろう。ユーザにとって、自分が判らない点には、相手が人間であれコンピュータであれ、教え方が不親切にかんじられ、また既によく知っている事を解説されると、長々しくうるさく感じるものである。

いわゆる人工頭脳、AI、ニューロチップ及び音声認識の進歩や、光ディスクファイル、電子カメラ等ハードウェアの進歩にも依存するが、現在のヘルプ機能を大きく超えたオンライン・マニュアル、On-line

Information が、手軽に利用出来る事が期待される。その時点では、ユーザのレベルと疑問を察知し、適切な答え、または方向づけを、テキスト/音声(音楽)/グラフ/ビデオの手段で、迅速にユーザに伝え且つユーザの理解をフォローする機能が要求されるであろう。

この場合の迅速さは、印刷されたマニュアルに例えれば、単に頁を順番にめくる早さでは無く必要な頁に直行する事であり、またユーザの再利用等を図りマニュアルはユーザによる再編集が可能な事も必要となる。ここで再編集とは、ユーザによる追記や、その場に至る道筋が登録できる事であり、ユーザは二度とおなじ思考過程を繰り返すことがない事である。

以上述べてきた、ハードおよびソフトの進歩を、ファイルの相互的、構造的側面から考えてみる。現時点でも、一般雑誌等に Hypermedia, Hypertext, Hyper Card 関連の記事が掲載される様になり始めた。例えば[6],「ハイパーメディアはすぐそこに」とのタイトルで、コンピュータとビデオと写真とサウンドを強力に混ぜ合わせたハイパーメディアは、大学で使われ、家庭にも 89 年末に入り始める。

ジョージ・ルーカスの映像やコーネル大学の医学解剖のスクリーン上での実習等エンターテインメント及び教育がテレビ以上の影響をうける。ハイパーメディアは光ディスクを使いサウンドのみならず、文章、静止画、ビデオ・スクリーンを切り貼りできる。光ディスクは多数の情報を小さな図書館なみに記憶でき、ハイパーテキストというソフトウェアがこのデータを管理し、ユーザは索引や目次を参照しないで次々に必要な情報にアクセスできる。

他のハイパーテキストの記事例では[7], ハイパーテキストの基本は一種のデータベース・マネージメントシステムであり、ユーザと情報スクリーンを連想リンクで結ぶもので、進んだレベルではソフトウェア環境としてユーザの共同作業、コミュニケーション、知識獲得を助けるものである。人間の様に参照リンクを通じて素早く直感的に必要な情報を出し入れする。

ハイパーテキストの歴史は、ルーズベルト大統領の科学アドバイザーであった、V. Bush の考えに始まるマンハッタン計画の技術の蓄積とその利用をマイクロフィルムを使用したシステムで構成した。後に '60 台前半 SRI の D. Engelbart がコンピュータを使ったシステム、On-Line System (又は NLS) を開発した。彼は又マウスを開発、及びフィルターの概念を提唱、この結果早くデータベース間の移動が可能になった。一方 T. Nelson はオンラインネットワークで Bush の考えを実現した。Xanadu というシステムと共にハイパーテキストという言葉も始めて提唱した。サンマイクロのワークステーションで動く(Unix)Xanadu が 1989 中に出荷される予定である。

ハイパーテキストは[8], 伝統的な情報記録の替わりになるもので、書籍の様な印刷物をユーザの指定す

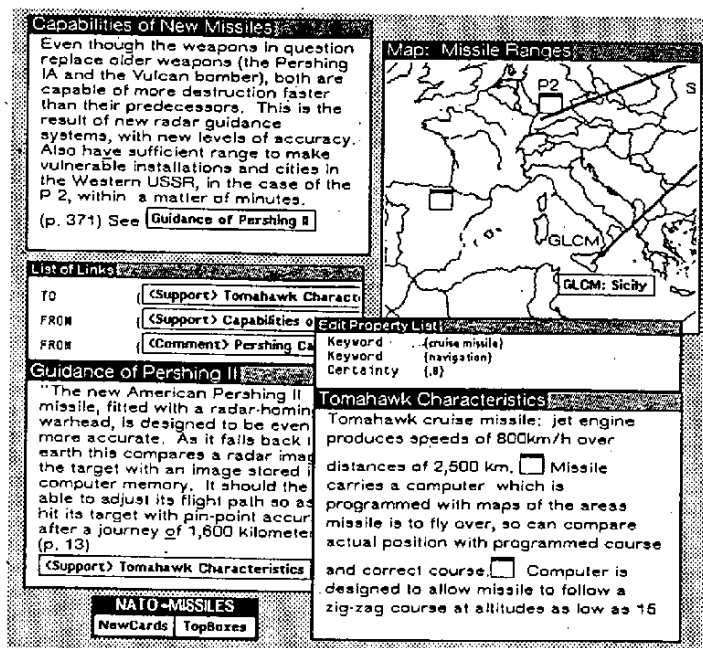


図 3.4-4 Note Card の例

キャプション: パーシング・ミサイルの関連情報

原典: Byte 10/88 p 239

る複雑な形に適応し易いハイパーテキスト階層に変換する。医学用テキストの原稿をXeroxのNote Cardsを用いてハンドブックにする例が述べられている。

ハイパーテキストは[9]製品/テクノロジーというよりコンセプト/テクニックであり、情報処理の見方/考え方のありかたである。我々のシステムの中には種々のアプリケーショングラフィックス/データベース/文章/スプレッドシートが混在している。この中には相互関連のあるもの/またはほぼ同じものがある。この関係のあるものを結ぶのがハイパーテキストの目的である。更にすすんで、あるfull textがあればその全てに関連するものにリレーショナルにリンクする様になる事が期待される。

これらハイパーテキストの主なアプリケーションとつぎの3分野になろう[10]。



図 3.4-5 CAI に応用した例

キャプション：BTRON での適用例，CAI は良い応用分野

原典：日経エレクトロニクス 2/20/89 p 170

- ・教育（インストラクション）
- ・電子出版とオンライン・データベース
- ・思考/著作ツール

このハイパーテキストのコンセプトを最初に商品化したのは Apple 社の Hyper Card ソフトウェアで [11]，今後の Macintosh の一部となり教育アプリケーションに適用されよう。この具体例として [12] USC で “Electronic Notebook” というプロジェクト名で新入生用の作文プログラムを拡張して，心理学部のデータベースに利用，人類学部が灌がい方法の評価をバリでフィールドテスト中である。また米国憲法をハイパーカード・インタフェースを通じてアクセスできる様拡張中である。

ハイパーメディアの機器そのもののサービスマニュアルをハイパーメディアを利用して作成した例も報告されている [13]。障害のある部品を画面上に表示された，プリントボード上でマウス等で指定すると，その部品の部品番号，診断方法，仕様等が音声またはテキストでユーザに知らされる。技術情報のコミュニケーションに使われた場合，

- ・より洗練されたオンライン・ドキュメント
- ・多層化されたトレーニング材料

- ・自分のペースでの教育ソフトウェア
- ・光ディスクまたはマウストレイジによる多様化されたユーザ・インタフェース
- ・パーツ・カタログ、サービス・インストラクション、アート・グラフィックス、地図、字体等を含むビジュアルなデータベース
- ・各種の音響効果、演説、音楽、動物の鳴声、外国語等のサウンド・データベース
- ・“スタックウェア”またはサードパーティのハイパーメディア・パッケージ

等を創造する事ができる。

準 AI といえるハイパーテキストと、エキスパートシステムを PC 上で動くレベルで比較した報告もある [14]。テクニカル・ライター、秘書、プログラマーが、この比較に参加したが、大多数の人々がハイパーテキストに使いやすさの点で好意的であった。エキスパートシステムは余りにも、タイプ量が多く、メニュー方式が、より好まれた。一方ハイパーテキストの欠点として左右の矢印のみに限定された操作、ウィンドの重なりが面倒に感じられた。エキスパートシステムはそのルールについての、予備知識が必要であるが、ハイパーテキストはワードプロセッサを操作できる人は、テキストをリンクする為のマークを付けるだけなので、より簡単である。

3.4.6 日本におけるハイパーメディアの状況

日本でも既に販売または、開発中のハイパーメディア・システムが報告されている [15]。BTRON (実身/仮身) と現在プロトタイプとして静止画、動画を松下 M500 上で稼働中である。三菱の PSI-II 上の静止画をあつかう推論機能を加えた、知識メディアステーション。NEC の静止画、動画の CAI ガイダンス (PC 9800 + ビデオ・プロセッサ)。日立静止画をあつかう、視覚的なユーザ・インタフェースを主な用途にする Concept Browser 等が主なハイパーメディア・システムとして開発されている。実用例として、東大病院では Hyper Card を使った診断支援システムを開発した。比較的症例の少ない急性肝炎に関し、過去の症例や治療課程をカードにまとめ、検索できるようにした診断そのものは、医師に任せるという考えである。

今まで述べたように、優れた機能と将来への可能性を持ったハイパーメディアも、現在まだ多く課題が残されている [16]。例えばマルチメディア・アプリケーションを開発するための「オーサリング・システム」。マルチメディアまたはデータベースを生かしたアプリケーションを効率よく開発できなくては意味がすくない。ところがそのために不可欠のオーサリング・システムに関する研究は、まだ手つかずの状態に近い。また既存の情報システムとマルチメディア・システムをいかに融合させるか、AI 技術を取り込んで人間にとってどこまで自然な操作環境にするか、など技術面での課題は多い。またマルチメディア・アプリケーションの普及には、知的財産権の問題がクローズ・アップされてくる。オリジナル映像を改変したり、またはそのままでも、本や雑誌などのコピーをイメージとして蓄積、関係者に簡単に配付できるような環境が自然に出来上がってしまう。標準化も重要なテーマである。CD-ROM にしても CD-ROM XA との関係、光磁気ディスク同士の互換性が問題である。データ圧縮方式の標準化も 90 年代に向けて、大きな問題として浮上しよう。

3.4.7 米国におけるハイパーメディアの開発状況

米国におけるハイパー関係の主な研究・開発活動の状況を示すと、以下の通りとなる [17]。

- ・ベル研 (Bellcore) Superbook (テキスト・ブラウザー)
- Telesophy (オンライン文書システム)

.....Thoh II (意味認識付ハイパーテキスト)

- ・ブラウン大学.....Intermedia (対話型学習システム)
- ・カーネギー・メロン大学.....Notes (ハイパーテキスト・ライターのためのツール)
.....ZOG (マルチユーザのためのハイパーテキスト・システム)
- ・MAD インテリジェント (コモン Lisp のハイパーテキスト, ニューヨーク株式市場で使用中)
- ・MCC gIBIS (問題解析ツール, サン・WS 上で稼働)
.....Plane Text (ユニックス版, 一般用システム)
- ・メリーランド大学.....Hyperties (教育・対話型エンサイクロペディア)
- ・ノース・カロライナ大学.....WE (対話型編集システム)
- ・サウス・カルフォルニア大学...DIF (ソフトウェア開発用ツール)
- ・ザナドゥ CoXanadu (ワールドワイドなハイパーテキスト・ライブラリー)
- ・ゼロックス (PARC)Note Cards (情報解析サポート・ツール)

3.4.8 ユーザ・インタフェース

ここでハイパーメディアおよびハイパーテキストを離れ、コンピュータのユーザ・インタフェースを考える。大型コンピュータのアプリケーションのユーザ・インタフェースも当然使いやすさ、ユーザビリティが、要求される[18]。

- ・ PF キー, コマンド, パネルに対する標準化
- ・ アプリケーション自体の中にドキュメントがより取り込まれている事
- ・ ユーザが仕事で何時でもオンライン・ヘルプが利用可能な事
- ・ ドキュメントの専門家がユーザが見えるもの全てに係わる事

が必要である。この内ユーザが見るものとは、

- ・ プログラム・パネルのレイアウトとそのフォロー
- ・ プログラム中のコマンド
- ・ PF キーの定義
- ・ ユーザへのコンピュータ・メッセージのテキスト
- ・ 印刷されるもの全て

である。

オンライン・マニュアルやオンライン・ヘルプにしても、人間とコンピュータの対話即ちインタラクティブ・インタフェースの上に成り立っている。このインタフェースのあり方を、分類しそれぞれのインタフェースについて、共通の評価基準で比較検討した研究がある。[19] この評価に関係する要素には、

- ・ 初めての人にとっての使いやすさ
- ・ たまに使う人にとっての使いやすさ
- ・ よく使う人にとっての使いやすさ
- ・ 抜け道 (fast path) 有ること
- ・ ユーザによる調整・修正が可能なこと
- ・ 選択の余地または程度
- ・ オブジェクト以外の余計なユーザの仕事負担
- ・ 仕事の早さ

などがある。一方、インタフェースの分類については、

- ・ コマンド型
- ・ シミュレーション型
- ・ スプレッドシート型
- ・ ワークエリア（ウインドー）型
- ・ 階層型
- ・ 複合型

がある。

コマンドインタフェースは、はやくからあるインタフェースで、ユーザがコンピュータに望むアクションをコマンドあるいはインストラクションの形でコンピュータに与える。ユーザは仕事のステップを指定でき、コンピュータは与えられた指示に従うか、できないので別の指示を要求するか、何れにしても人とコンピュータとの対話から成り立っている。この型のインタフェースは、最初は馴染みにくいだが、習熟したユーザにはパワフルで、新しいコマンドの加えやすさなど良い点があるが、たまに使う人や初心者には難しい。

シミュレーションインタフェースは、全てがユーザの制御下にある様にみえる。例えばフライト・シミュレーションの様に、ユーザは飛行機の操縦を全て制御し、その結果に責任を持つようシミュレートされる。この型は、煩雑なコマンドやメニューがなく、飛行機の操縦桿の様に、直裁で明快である。一方このためのプログラムが複雑で、ユーザがシミュレーションする対象に詳しいことを前提にする必要がある。

スプレッドシートインタフェースは、一番ポピュラーなものでユーザは、スプレッドシート上の枠内にデータを入れるか、枠と枠の関係式を指定するだけでよい。従って非常に入りやすく、習熟するにつれて、複雑な仕事も容易にできる様になる。他方抜け道があまり無く、プログラムの修正も与えられた機能の組み合わせしか出来ない。

ワークエリア（ウインドウ）インタフェースは、いくつかの関連した（あるいは無関連の）情報を同時にユーザに提供し、ユーザは必要に応じて新しい関係をつくっていく。自由度が高く、経験を積む程使い易くなり、仕事のグラフ化など、異なった側面から、同時的・多角的に利用することができる。但し、全く初めてのユーザにはとっつき難く、余りに多くの情報が同時に与えられるため、ユーザがオーバーロードになり易い。

階層型インタフェースは、例えば連続するメニュー方式のように、ユーザフレンドリーを目的とした階層型のパネルを示し、つぎつぎにユーザに選択肢から選ばせていく。このため容易に使うことができ、過ちをおかすことも無い。また他のインタフェースと共存して利用することができる。他方、高度なユーザにはかえって不便で、また一時に一つの仕事しかできない。

複合型インタフェースはアプリケーションが複雑になった場合、いくつかの基本インタフェースを組み合わせ構成する。各々のインタフェースの長所・特徴をアプリケーションに合わせて採用する必要がある。

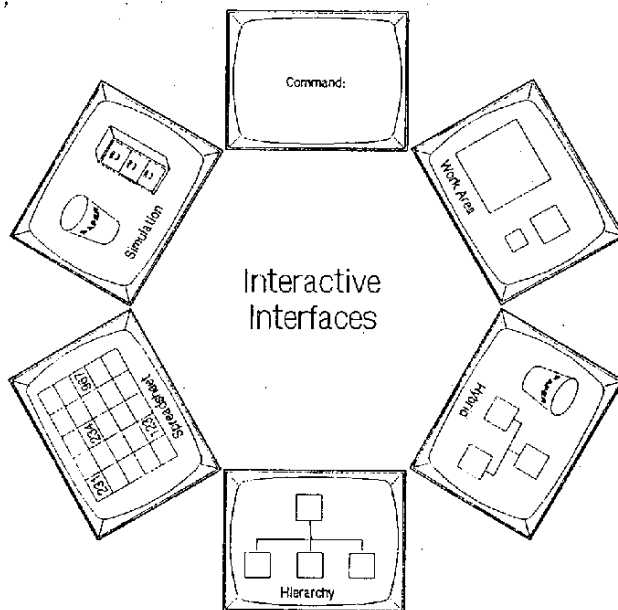


図 3.4-6 インタフェースの分類

キャプション：ユーザ・インタフェースの様々

原典：35th ITCC ATA-94

3.4.9 ユーザ・インタフェース統一の例

一方、ワークステーションもしくは端末にむかっている人とのユーザ・インタフェースをアプリケーションによらず共通/統一化して人に要求処置を分かりやすくしている共通ユーザ・アクセス (CUA) が IBM により提唱されている [20]。適切に設計されたエンド・ユーザ・インタフェースをアプリケーション・プログラムに利用すると、アプリケーションは学習・利用しやすくなる。エンド・ユーザ・インタフェースにアプリケーション間およびシステム間で一貫性を持たせると、一つのアプリケーションまたはシステムで得られたテクニックは通常他のシステムやアプリケーションにも応用できる。CUA は、ユーザがアプリケーションを見たりアクセスするためのウインドウを提唱し、使い易さの向上と一貫性により、ユーザが容易に利用できるアプリケーションの数は増加する。

CUA は人とコンピュータ間の対話用に定義されている。機械が人とのコミュニケーションをどうとるか、つまり、スクリーン・レイアウト、カラーや強調表示の利用、メッセージ、およびヘルプ情報を含め、コンピュータがオペレーターに提示するものを CUA は扱う。また、CUA にとって重要なことは、ユーザと機械とのコミュニケーション方法で、キーボードのレイアウトや使用方法、マウスの使用、スクロール、選択方法などが関連する。使い易さを向上するガイドラインとルールが開発され、これらの規則を広範囲に適用する連続性が学習を容易にする。CUA はインテリジェント・ターミナルの機能を活用すべく設計されている。

具体的には [21]、Main Menu の Title は一行目に中央にそろえる。この Panel の Instruction は左詰に配置する。白丸をラジオ・ボタンといい、項目のなから一つだけえらべる。ユーザが項目を選択する黒い点がボタンの中に表示されること。一画面に納まらないリストはスクロール・バーを表示し、ユーザに表示域外に情報が存在することを知らせる。パネルの最上位にあるバーをアクション・バーといいユーザの要求できるアクションを単一語句または複数語句の形で表示する。アプリケーション間で共通の機能は最下位の機能キーに割り当てる。(例えば、F 1 = ヘルプ、F 3 = 終了、F 10 = アクション、ESC = 取り消し)

その他、ユーザの応答不用の通知メッセージ、バックグラウンドが黄色の警告メッセージ、赤のバックグラウンドの重大メッセージはユーザの処置が必要なこと等が統一されている。但し、これら以外の色は、ユーザの好みで自由に選ぶ事が出来る。

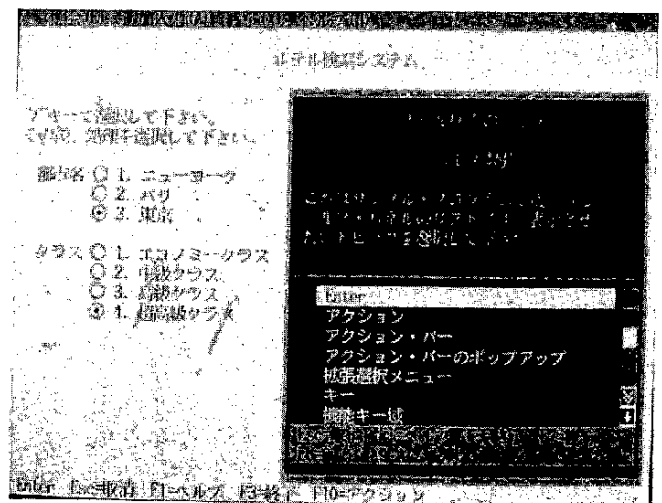


図 3.4-7 共通ユーザ・インタフェースの例
キャプション：ホテル検索システムのヘルプ画面と
スクロール・バー

原典：SAA Common User Access, IBM SC
26-4351

3.4.10 オンライン・インフォメーションの保守・情報検索への利用例

オンライン・インフォメーションをコンピュータの情報検索や保守に応用されている。IBM の AS/400 にその一例をみると、客先の端末から直接メニュー型のスクリーンで、センターにある必要な情報に到達することができる。具体的には、

- ・製品発表ニュース，サプライ商品ソフトウェア・カタログなどからなる製品情報
- ・汎用性の高い質問・回答集，処理中の質問などの技術情報
- ・マニュアルや，講習会・催物の案内

などにアクセスできる。保守サポート機能については、「コピー・スクリーン機能」により客先の端末の画面そのものを保守要員の端末にコピーして問題解決を早めたり、「表示装置パス・スルー機能」で保守要員が遠くから客先の機械を操作して解決をはかることができる。

ソフトウェアに障害原因がある場合には，その障害がすでにセンターに登録され，既知の修正策があれば，オンラインで客先に送られ即座に解決される。

3.4.11 将来のオンライン・インフォメーションの要件

以上述べてきたことから，今後のオンライン・マニュアルはどうなるか，を予想して見る事にする。将来のパーソナル・コンピュータは，それぞれの時点でのハードウェアまたはソフトウェアの進歩を取り込んで，たえまなく高機能化と使いやすさを，安いコストで，おおきな矛盾なしに達成していくであろう。現在でも，音声や画像はごく普通のこととしてパーソナル・コンピュータに取り込まれている。人間にたとえれば，良き教師のように，話して分からせ，黒板にテキストや絵を書き生徒（ユーザ）に，必要なことを理解させている。さらに生徒の進捗度・理解度を把握している。その時点での利用できるメディアを適宜使うにしても，良き教師のすべきことは最低限カバーする必要がある。将来のオンライン・マニュアルに必要な事は，

- ・マウスやキーボードによる入力に加え，音声入力・指示の受け入れ
- ・音声によるユーザにたいする警告，お願いなど
- ・創作・編集された画像・音楽の記録
- ・画質を落とさない拡大または縮小
- ・ビット・イメージのベクトル化または，テキストをイメージから取り込み
- ・電子カメラとの連動或は，TV等の動画の表示と取り込み
- ・香り・味覚など五感の感知と表出
- ・ホーム・オートメーション，家庭電気製品とのリンク
- ・手を取る機能（例えば，楽器の演奏，習字の筆を一緒に運ぶなど）
- ・ニューロ・チップ等によるコンピュータ・ユーザ相互の学習
- ・あいまい情報の処理
- ・ユーザのパフォーマンス・レベルの把握とそのレベルでの対応
- ・大型 HDTV の利用またはウィンドウ・サイズの縮小・拡大の自在化
- ・コンピュータ・ネット，ISDN，または電話とのリンク

など，あらゆる上下，左右の情報源を共同し，目的にたいして最適なメディアを駆使する事であろう。

3.4.12 本当に必要なもの

何にしても，将来のパーソナル・コンピュータは人間にとって役にたつ道具，助脳道具，の役割をはたすことが期待される。一般に道具は文化が進むほど，細分化されるのが普通である。このパーソナル・コンピュータという助脳道具は，ユーザの好みに応じて，汎用からプロ用の細分化された仕事まで応じられ，さらにパラレルに仕事を進める能力迄を本来的に持っている道具なのである。日本で民芸品の美を，最初に評

価した柳宗悦によれば[22]、道具の持つ美しさは、おもに次の点からくると指摘した。

- ・ 日常的な意味での実用性
- ・ 簡明で健康的
- ・ 素材の性質を（ときには欠点を含め）あるがままに利用

これら日本の各地で民芸品を作った工人達は、芸術家とは違って意識的に美しいものを作ろうとした訳ではない。結果として美と実用性がマッチして、それを使う人々の心を捕らえて、愛着をもってその品物を日々使われたものである。

助脳道具の歴史はまだ始まったばかりであり、また本来的/潜在的能力が従来の道具より画期的なため、本質を見失いがちになるが、パーソナル・コンピュータも、そのオンライン・インフォメーションも実用性を恒に重視する必要がある。この実用性についても、助脳具というからには、民芸品と同様かそれ以上に、それを使う人々の感性に訴えるものが欲しい。この感性に訴えるものと、実用性との間には、いくつかの段階があろうが、三蔵法師が天竺にお経を悟空達と求めた苦難の旅の末、一度は白紙のお経を最高のものとして戴いたが、次善のものとして、再度字の書かれたお経を貰い直した寓意を、思いおこす必要があろう。

参考文献

- [1] 梅棹忠夫 中央公論 9/1988 p 142-53
- [2] 日経 EP 6/22/1987 日経マグローウヒル社
- [3] Y. Nakajima Demand Printing 35th ITCC ATA-105
- [4] 加古川 日経コンピュータ 2/13/1989 p 87-98
- [5] J. Gillette Trends and Responsive Strategies 35th ITCC ATA-143
- [6] Newsweek (Switzerland) 10 Oct. 1988
- [7] J. Fiderio BYTE Vol.13, No.10 Oct. 1988 p 237-44
- [8] M. Frisse BYTE Vol.13, No.10 Oct. 1988 p 247-53
- [9] P. Gillman Comput. Libr. (UK) Vol.2, No.2 Sept. 1988 p 12-13
- [10] Chia jer Tsai J. Educ. Technol. Syst. Vol.17, No.1 1988-1989 p 3-14
- [11] R. Parker 8th Annual TCEA State Conference, Proceedings 1988 p200
- [12] S. K. Kinnell Online '88 Conference Proceedings 1988 p89
- [13] A. Rockley 35th ITCC ATA-10 Hypermedia A Web of Thought
- [14] K. Hansen 35th ITCC ATA-28 Comparison An Expert and a Hypertext System
- [15] 稲葉則夫 日経エレクトロニクス 2.20.1989 (No.467) p. 167-74
- [16] 田口潤 日経コンピュータ 1.2.1989. p 32-50
- [17] BYTE Vol.13, No.10 Oct. 1989. p.268
- [18] R. Lenehan 35th ITCC ATA-24 User Interface for Mainframe Application
- [19] R. Grice 35th ITCC Interactive Interfaces ATA-59
- [20] E. Wheeler IBM Systems Journal 1988 Vol.27 No.3
- [21] SAA Common User Access Panel Design and User Interaction SC 26-4351, IBM
- [22] 柳宗悦 手仕事の日本 岩波文庫 p 230

4. ヒューマンインタフェースにおける ユーザモデルと認知心理学的アプローチ

4.1 はじめに

従来、コンピュータの役割は、シミュレーションなどの数値計算に典型的に見られるように、人間のある機能を「代行」させることに重点が置かれてきた。これに対しこれからは設計などの高度に知的な作業をコンピュータと協調して行うといった、コンピュータによる人間の「支援」という側面が大きなウェイトを占めて来よう。当然の事ながら、このような状況下では人間とコンピュータとの対話の柔軟さ、言い換えればユーザフレンドリなインタフェースがこれまでとは比較にならないほど重要になる[3]。

ところが、現在のコンピュータインタフェースが知的で柔軟とは言い難く、これを改善しようとする様々な試みがなされてきた[1]、[2]。ここで「知的」とは、インタフェースが類似の不適切な応答を繰り返さないとか、ユーザにコンピュータのより便利な使用法を示唆できるなどを意味し、「柔軟」とは、インタフェースがユーザの誤りのある程度許容できるなどを意味するものとする[4]。

このような知的で柔軟なインタフェース構築のためには、ユーザの行動特性をより良く知る必要がある。そのためには、ユーザとコンピュータとの実際の対話状況を観察することがまず必要である。実験室内での完全にコントロールされた状況下では、伝統的な実験心理学のアプローチが適用可能であるし、実験結果の蓄積も多い。しかし後述の物理インタフェースなどを除けば、インタフェースとユーザとの対話を、このような狭い枠で捉えることは不可能である。言い換えれば、単なる刺激反応という入出力関係だけでは不十分であり、それらの間の内部構造をも同時に扱う認知心理学的アプローチが必要となってくる。このような認知心理学的分析によりユーザの行動特性を抽出し、これに基づいてユーザ行動の背後にある意図を推測できるモデルを構築することが重要である[3]。

インタフェースにおけるユーザモデルとしては、さまざまなレベルのものがある。カードらはインタフェースを、ユーザとの物理的な対話と捉える物理インタフェース、ユーザを情報処理器と捉える認知インタフェース、コンピュータのメンタルモデルとして捉える概念インタフェース、問題解決を扱うタスクインタフェースの4つのレベルに分けている[5]。研究の興味の中心はより抽象レベルの高い分野に移行しつつあると言えよう、シュナイダーマン[2]も、インタフェース設計に際して、問題解決なども扱えるように伝統的心理学的理論を拡張すべきであると主張している。

別の視点として、ノーマンは人間とインタフェースとの関わりを対話の時間的順序に沿って捉える、ユーザ・アクティビティの7段階モデルを提唱している[1]。これは、図4-1に示すように対話過程を①目標・②意図・③行動決定・④実行・⑤認知・⑥解釈・⑦評価の7段階に分割するものである。ここで④実行と⑤認知の間に物理的なコンピュータが存在する。

インタフェースにおけるユーザモデルは、これらさまざまなレベル・段階ユーザを支援できるものでなければならない。このために、ユーザの意図を推測できることが重要であることは、既に述べた通りである。

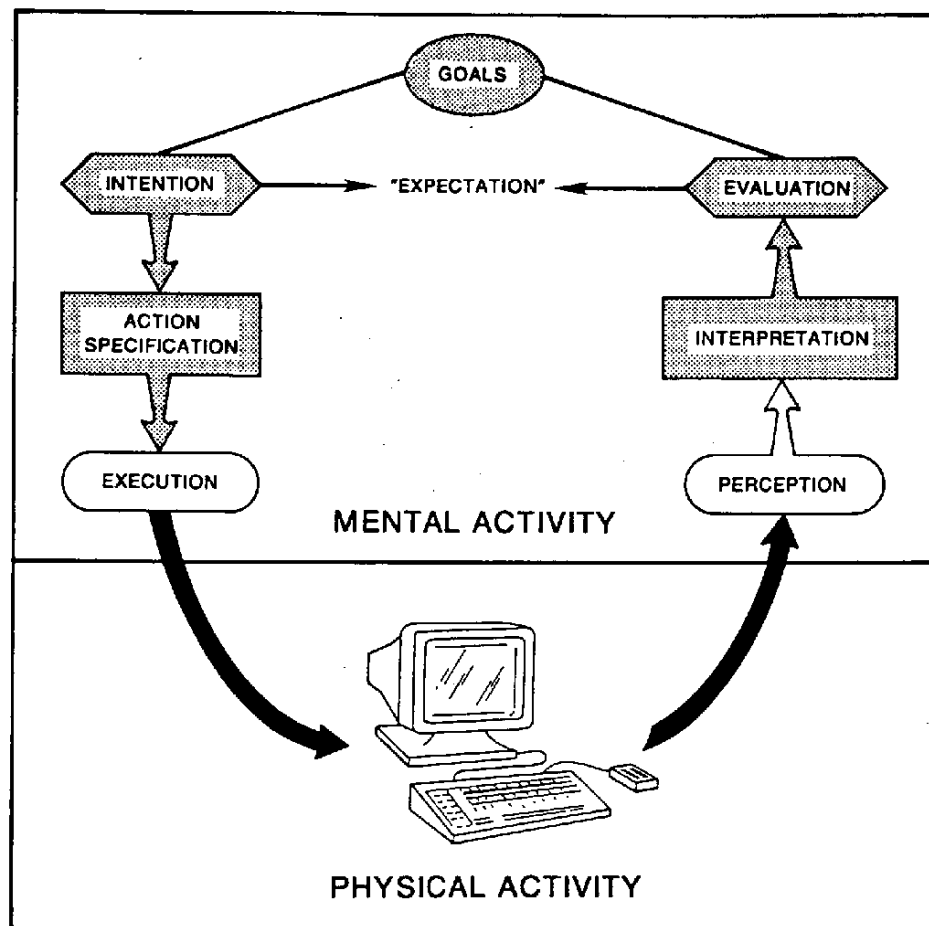


図 4-1 ユーザ・アクティビティの7段階モデル [1]

4.2 ユーザモデルの分類・特徴

インタフェースのためのユーザモデルに要求される特徴を挙げてみよう。第一に、ユーザ行動の学習が行えるモデルである必要がある。第二に、個人差が大きいのでユーザ一般ではなく特定個人のモデルである必要がある。第三に、人間とコンピュータとの対話の場で、ユーザモデルを実時間で構築する必要がある。これらのため通常のモデル化方法とは異なった方法が必要になると考えられる。

ユーザモデルのタイプとしては、大別するとルールを用いるモデル[12]と、コネクショニズムに基づくモデル[6], [7]とが考えられる。前者のルールベースモデルでは、コンピュータとユーザとの対話に基づくユーザの行動とルールとの対応づけ、新しいルールの獲得およびルールの再構成といったことがモデル化の重要なポイントとなる。ユーザの行動とルールとの対応づけに関しては、ユーザのある行動に対応するルールが一般に複数個存在することが問題を複雑にしている。

後者のコネクショニストモデルは、以下で述べるようにインタフェースに関連するいくつかの特長を潜在的に有していることから、有望なアプローチであると考えられる。インタフェースへの応用という観点からコネクショニストモデルの特徴を整理してみよう[4]。

第一に、複数の制約を満足する解が自然に浮かびあがってくるというコネクショニストモデルの特徴が、一つ一つでは決定ではない多くの断片的情報に基づいてユーザの意図を理解するためのモデルとして適していると考えられる。もちろん与えられた情報からのみではその意図がユニークに定まらないことも多いが、解が制約を満足している程度がユニット活性度の大きさで表わされることから、このような場合こそコネクショニストモデルが適していると考えられる。後述のタイプミス訂正モデルでもこの特徴が用いられている。

第二には、分散表現を用いるメリットとして、一般化がより柔軟に実現できることが期待される。すなわち上記のルールの獲得・再構成がリンクの重みの変化という連続的な形で陰に表現されるもので、ユーザのふるまいのより柔軟な学習が期待できる。ただしこの利点は、学習後のモデルの理解が困難であるという欠点と裏腹の関係にある。

第三の同じく分散表現を用いているため、過大な学習をさせた場合のパフォーマンスの低下がゆるやかである。

第四に、平均的ユーザのモデルを出発点として特定ユーザの行動の学習が進む結果、当該ユーザに適應すべくモデルが逐次修正され、カスタムインタフェースの実現が期待できる。言い換えれば、長期間使い込むほど該当ユーザの特性を一層取り入れたユーザモデルが形成される。

第五に、実行時には活性度の伝搬を計算するだけでよいので計算時間は短くてすみ、さらにモデルがハードウェア化された場合には並列計算が可能となり、実時間性が要求されインタフェースという場に適している。もちろん学習には長時間を要するが、これはバックグラウンドジョブとして実行すればよい。

第六に、不完全なパターンが与えられても完全なパターンが想起されるというコネクショニストモデルの特徴から、ミスタイプされたコマンドを入力として与えると、正しいコマンドが出力に連想されてくるといったことが期待できる。

第七に、インタフェースでは同時進行的なジョブが多く見られるが①、このような並列的なジョブの記述にコネクショニストモデルが適している。レベルは多少違うが、たとえば、モータコントロールの並列性をモデル化することにより、タイプミスの発生をうまく説明している例がある[13]。

4.3 情報処理モデル [14]

カードらは、人間を情報処理器として捉える立場に立ち、テキスト編集などの対話的タスクを例にとり、エキスパートユーザの対話時の認知スキルの研究、とりわけ対話の時間的側面に焦点を当てた研究を行った。

人間を、長期記憶、作業記憶、プロセッサから構成されている情報処理器と捉えたモデルとして、図 4-2 のようなものを提案している。人間とコンピュータとの対話の典型例としてテキスト編集が選ばれる。さまざまなテキスト編集システムを比較したベンチマークテスト、およびユーザの個人差の研究がなされている。テキスト編集システムでのユーザの情報処理モデルとして、GOMS モデル（ゴール、オペレータ、方法、方法選択規則）を提案している。また GOMS モデルを単純化したものとして、タスク逐次時間を予測する工学的モデルも提案している。さらにテキスト編集を認知スキルの一例と捉え、ここでのアプローチを、VLSI のレイアウト問題などのより一般的な問題解決へと拡張する試みもある。

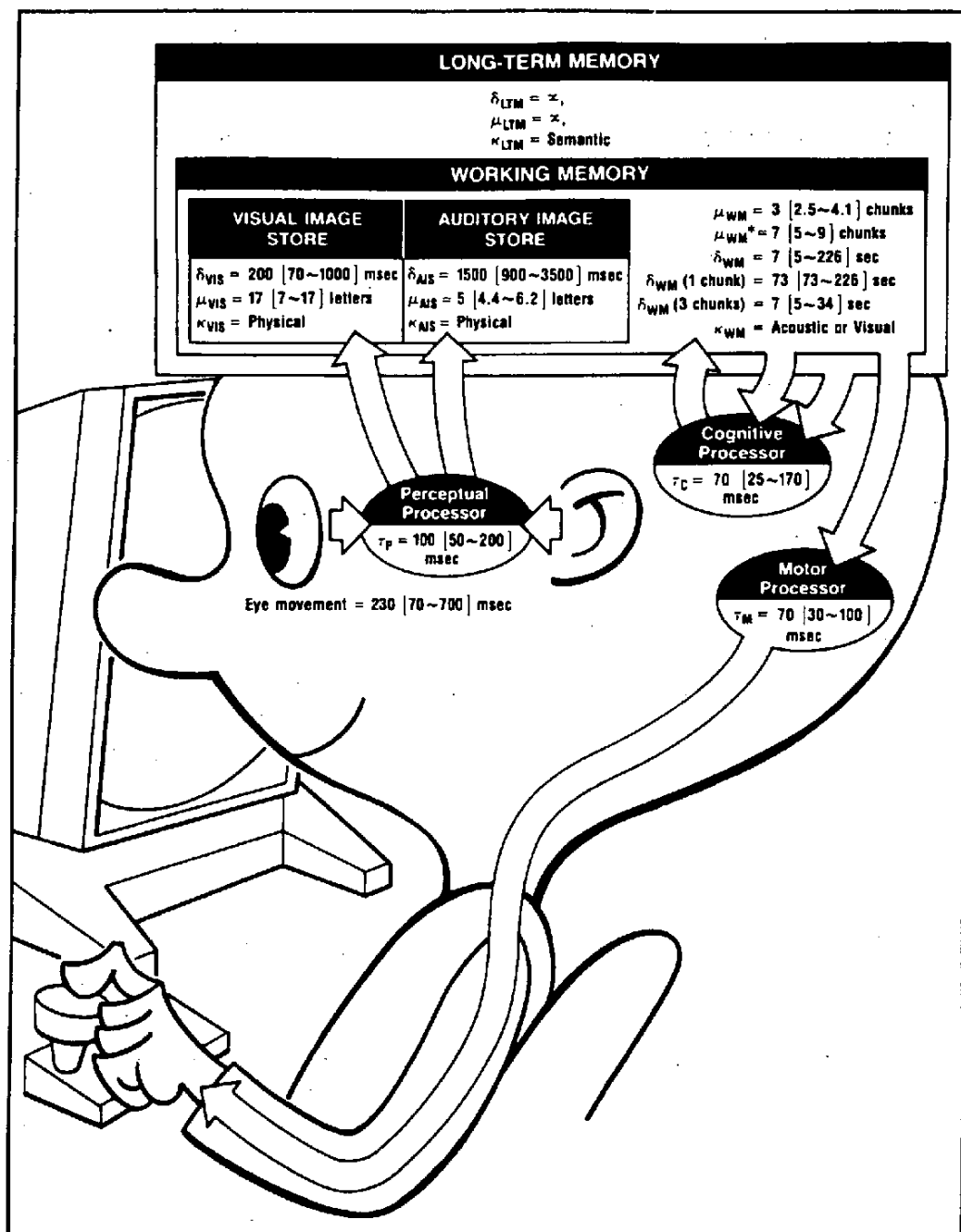


図 4-2 ユーザの情報処理モデル [14]

4.4 ルールベース形モデル

4.4.1 ACT *モデル [11]

人間の高度の認知機能のモデル化の基本的考え方として、機能別に個々の能力があるのではなく、単一の心的システムで各機能の実現が可能であるという仮説をアンダーソンは採っている。プロダクションシステムがこの計算アーキテクチャとして適切であるとし、ACT *モデルを提案した。

ACT プロダクションシステムモデルの一般的枠組みは次のようなものである。これは図 4-3 に示すように、作業記憶、宣言的記憶、プロダクション記憶の三種類の記憶から成っている。符号化過程は、外界からの情報を作業記憶に取り込む。行動過程は、作業記憶中のコマンドを行動に変換する。これらの二つの過程は外界とのつなぎの部分に相当し、ACT 理論の中核ではない。蓄積過程は、作業記憶の内容を宣言的記憶に恒久的に記録したり、既に宣言的記憶に記録されているものを強化したりする。検索過程は、宣言的記憶から情報を取り出す。照合過程では、作業記憶中のデータがプロダクションの条件部と比較される。実行過程では、照合されたプロダクションの実行部を作業記憶に入れる。プロダクションの照合と実行とを併せてプロダクション適用と呼ぶ。

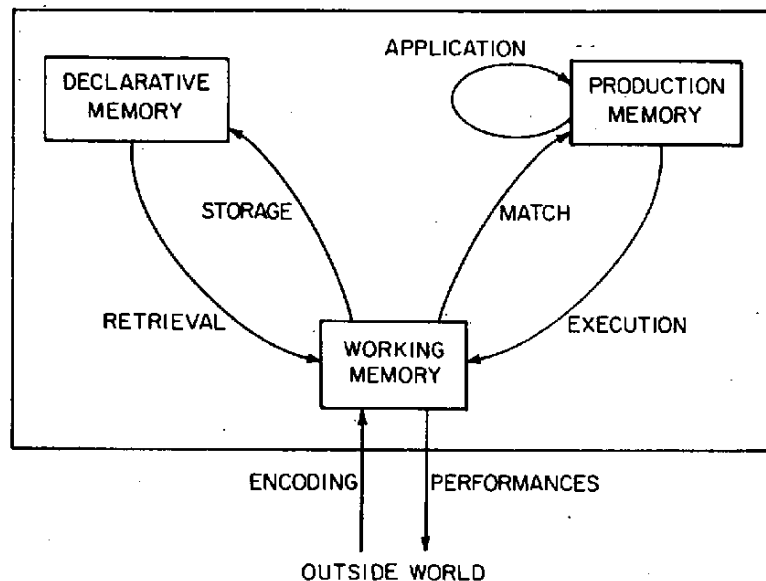


図 4-3 ACT プロダクションモデルの一般的枠組み [11]

ACT *モデルは 14 個の基本的仮定を置いている。宣言記憶中のユニットは活性度を有しており、ユニット間のリンクには強度が定義されている。ユニットの活性度はリンクを通じて伝播する。プロダクションにも強度が付与されており、その適用が成功すると強化される。活性度は、プロダクションの選択に影響を与える。プロダクションのコンパレーションやチューニングが行われるなどが特徴的である。

この ACT *モデルを、空間イメージの表現と操作、想起、定理証明、文字認識、言語獲得などの問題に適用している。

4.4.2 三段論法における思考のモデル [3]

実時間での個人別のモデル化に適した方法として、石川は計算機との対話を通じて「人間の振るまい」とあらかじめ用意した「特性ルール群」とのマッチングに基づいて人間の特性を実時間的にモデル化する方法

を提案した。このモデルは、いかなる状況の下で人間がどのような振る舞いをするかをルールで記述するものである。その際、計算機はシミュレーションとか自動的データ収集といったことだけでなく、途中プロセスの推定、過去の履歴に基づいた的確な質問の呈示、ルールで説明できない部分の抽出などを行っており、より積極的に計算機の推論能力を活用しているのが特徴である。

このモデル化に際しての困難な点は、人間の振るまいにマッチするルールが一般には複数個存在すること、複合誤りが存在すること、および人間を対象とするために勘違いやキーの押し間違いなどの混入が避け難く純粋なデータが得にくいことである。ここでは、このような状況の下でも真の誤り原因を究明できるためのアルゴリズムを提案している。更に、この様な方法論およびアルゴリズムの有効性を実証するために、定言的三段論法を対象として、誤答に対して誤りの原因を想定してこれを検証するのに有効な問題を呈示するという手順で誤りの原因を究明するための実験システムを作成した。このようにユーザの応答にしたがってシステムの応答を変えられるのはモデルトレーニングの特長であり、ユーザインタフェースとして持つべき特性の1つである。

多くの場合ある回答を説明できる誤り原因が複数個存在するもので、これらの中からどれが真の誤り原因であるかを決定しなければならない。ここでは被験者実験により得られた一連の回答を可能な限りで説明できるような誤り原因を決定するアルゴリズムを提案した。このアルゴリズム自体は定言三段論法を対象とするものであるが、ユーザの意図を理解するためには1つだけの応答だけでは不十分であり一連の応答を調べる必要があること、および勘違いなどのためにどうしても理解しきれない応答が残りがちであるということでの考え方は、モデルトレーニングがうまく機能する上で不可欠である。

この実験システムを用いた被験者実験を行い、この結果を、どのように考えて結論に到達したかという被験者からのプロトコール情報と対比させた結果、ここでの提案の有効性が実証されら。また実際の人間の特性の個人差が大きいことか確認され、個人別モデルの必要性が実証された。

4.5 コネクションリスト形モデル

4.5.1 タイピングのモデル [13]

カードの分類では物理レベルに相当するものとして、ラメルハートによるタイピングのモデルがある。ここでの目的は人間の熟達した動作を研究することである。タイピングの基本現象として、キーストロークの時間に関して次のようなことが分かっている。第一に、人間は非常に速くタイプできる。速いタイピストのストローク間の平均時間は 60 msec であり、神経の伝達速度に近い。これではストローク間のフィールドバックはあまりありえない、第二に、異手間のストローク間隔のほうが同手間のストローク間隔よりも短い。第三に、同手間のストローク間隔は次のキーまでの距離の関数である。第四に、ストローク間隔はコンテキストに依存する。第五に、隣接したストロークの間に時間間隔には負の相関がある（とりわけ異手の場合）。

また、誤りパターンとして、次に三つを考慮している。第一に、隣あった文字の互換（transposition errors）であり、たとえば *bacause* を *bacuase* とする間違いである。特に異手間の誤りが多い。第二に、二重打ち（doubling errors）があり、たとえば *school* を *scholl* とする間違いである。第三に、alternation reversal errors と呼ばれる上記の組合せのような間違いがあり、たとえば *these* を *thses* とする。

このような特徴を考えると、指の移動は並列的でなければならない、また通常のスピードでタイプするためには、少なくとも 8 文字先を見ていなければならない。そこで、基本的フレームワークとして、Activation Triggered Schema system (ATS) と呼ばれるものを用いた。このモデルはスキーマの集まりであり、それぞれのスキーマは活性値を持っている。モデルの入力は文字列であり、最大の活性値を有するスキーマが実際の指と手の動きに相当するとしている。

4.5.2 タイプミス訂正モデル [4]

石川は、コマンド言語システム形インタフェース [2] におけるユーザ支援を目指して、コマンド入力支援およびコマンド系列特徴抽出のためのモデル化を試みた。

前者はコマンド入力中にタイプミスがある場合、入力文字情報などに基づいてユーザの意図を推測してその訂正を行なうためのモデルであり、ユーザの誤りがある程度許容できる柔軟なインタフェースの持つべき機能の一つである。これは、カードの分類では物理インタフェースに近く、7 段階モデルでは④実行を支援するものと考えられる。

タイプミスに関しては、①隣りあった文字の互換、②1つの余分な文字の挿入、③1文字の脱落、④列・行・左右置換誤りの4種類を考慮の対象をした、①②③を「時間的誤り」、④を「空間的誤り」と呼ぼう。

タイプミス訂正モデルは、タイプミスのあるコマンドが入力として与えられた場合、正しいコマンドが出力されることを目的とするものである。モデル化のポイントは、問題の構造をうまく表現するようにネットワークの構造、入出力表現をうまく選定することにある。この問題の場合には、入力層と隠れ層の間に粗符号化層を設け、入力層との間に入力の時間的あいまいさを表現するための特殊なリンクを張ってある。また、空間的誤りを表現するためにタイプ入力の符号化に工夫を凝らした。このように問題の構造を事前にネットワークに埋め込んでおき、これに多くの入出力パターンを訓練例として与える。訓練例としては、ミスタイプされたコマンド（入力）と訂正されたコマンド（出力）のペアを与えるとともに、正しくタイプされたコマンド（入力）と同じコマンド（出力）のペアも与えた。これらの誤りパターンを学習しこれから一般化を行い、未経験のタイプミスに対しても適切な訂正を行えることが期待される。タイプミス訂正モデルの基本的構成を図 4-4 に示す。

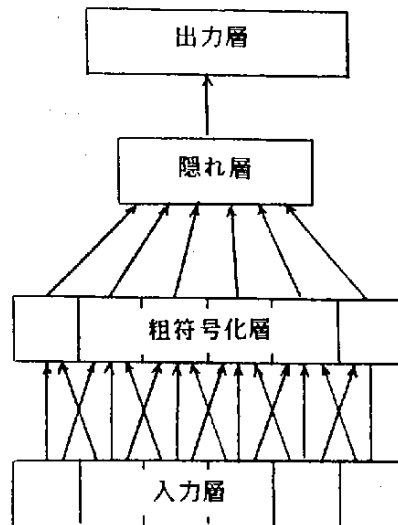


図 4-4 タイプミス訂正モデルの基本的構成 [4]

実験の結果、長いコマンドの訂正能力は十分高いのに対し、短いコマンドでの訂正能力は誤り率で約 20 % であり、予想された通り必ずしも満足すべきものとは言えない。

このため入力文字情報以外の情報をも併せて考慮することで性能の改善を図ろうと試みた。第一に、入力文字情報に加えて過去のコマンド系列というコンテキスト情報を併せ利用する。しかし現実には、コマンドの生起確率、コマンド間の遷移確率などの統計的性質が系列によって大きく変動するため、コンテキスト情報の利用はあまり有効ではなかった。

第二に、該当ユーザのタイプミス特性情報を用いた。タイプミスにはユーザ固有の特性があり、ジョブ内容に左右されるコマンド系列よりは、タイプミス特性のほうが安定的と考えられるからである。実系列を用いた評価を行った結果、比較的良好なタイプミス訂正能力が得られた。

4.5.3 コマンド系列中の特徴抽出モデル [4]

系列特徴抽出モデルは、コマンド系列の特徴に基づいて便利なマクロコマンドの示唆をユーザに与えらるためのモデルであり、知的なインタフェースの持つべき機能の一つである。これは、カードの分類ではタスクインタフェースに近く、7 段階モデルでは③行動決定を支援するものと考えられる。

コマンド系列が与えられた場合、どのコマンドが頻繁に出現するか（一次生起確率）、あるコマンドの後にはどのコマンドが良く現われるか（一次遷移確率）、二つの引き続いたコマンドの後にはどのコマンドが良く現われるか（二次遷移確率）、長さ 2 の部分系列がどの程度頻繁に出現するか（二次生起確率）などの特徴を抽出するモデルの構築を考える。

特徴のうち遷移確率を求めるモデルとしてバッファモデルを用いることが出来る。一般に、遷移確率の場合の数は高次になると組合せ爆発的に増加するが、バッファモデルを用いることにより、この困難を回避することが可能である。図 4-5 に二次の遷移確率を求めるためのモデル構築を示す。

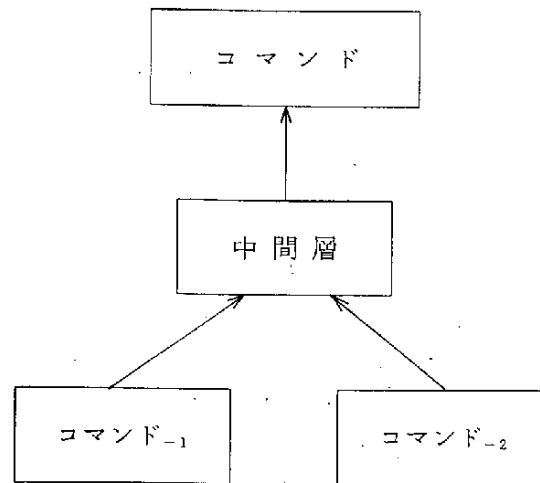


図 4-5 二次の遷移確率を求めるためのバッファモデル [4]

長さ 103 の実際の UNIX のコマンド系列で、一次生起確率および一次・二次遷移確率を上記の方法で求め、理論値と比較を行い十分な精度で求まることを確かめた。またこのコマンド系列中に頻繁に出現する部分系列を求めた。この結果、長さ 2 の部分系列として例えば “lf → cat” , “cat → rm” , “cat → lf” , “rm → lf” が、長さ 3 の部分系列として “cat → rm → lf” , “rm → lf → cat” が顕著な部分系列として見つかった。同じ長さの別の系列では、長さ 2 の部分系列として “cd → lf” が、長さ 3 の系列として “cd → lf → vi” が見つかった。これらはいずれも良く出て来そうな部分系列であり、“rm → lf” , “cd → lf” , “cat → rm → lf” などはマクロコマンドとして登録すると便利であろうと考えられる。

4.6 認知心理学的アプローチ [14]

人間-コンピュータシステムにおいて、ユーザ、コンピュータ、タスクの三つがシステムの構造要素である。心理学は、事後的な評価だけでなく、設計の際に用いられるべきである。この枠組は次の三つに分類される。

① 人間-コンピュータシステムの構造と性能

人間-コンピュータシステムは、その構造変数によってさまざまに変化する。構造変数は、タスク、ユーザ、コンピュータの三種類に大別される。たとえば、タスク領域は当然大きな影響を与える。ユーザも、その知的能力、認知スタイル、経験、使用頻度、方法・概念等の知識、タイプ能力などが大きく変動するので、これらを構造変数と考える。コンピュータに関しても、その対話スタイル、コマンド体系、ディスプレイ内容、入力装置、反応時間など大きな影響を与えるので、これら構造変数とする。

最終的関心は、構造そのものではなく、その性能である。性能にも多くのものがあるので、これらを性能変数と呼ぶ。システムに関する基本的性能変数としては、システムがどのタスクを行えるか、これが可能になるまでの学習時間、タスクを実行するのに要する時間、誤りがどの程度発生するか、また誤りがどの程度必然的か、タスクがどの程度うまく出来るか、予期しない状況下でどの程度ロバストであるかなどがある。この他にも、疲労やストレス下でも性能、ユーザの作業記憶・長期記憶の必要量、ユーザの使いやすさに関する主観評価などがある。

構造変数と性能変数との関連を求めるものは、心理学の役割である。

② 人間-コンピュータシステムの性能を予測する性能モデル

人間-コンピュータシステムの設計に際しては、構造的制約と性能目標を与えられる。構造的制約に関しては直ちにチェックが可能であるが、性能に関してはそうは行かない。そこで、構造変数から性能変数を予測する性能モデルが必要となる。この性能モデルにもさまざまなものがある。実験的モデルには実際のシステムあるいはその物理的模型をユーザが使用してその性能を測定するものである。記号モデルは、計算的モデル、シミュレーションモデル、代数的モデルなどを用いて実際のユーザを用いることなく性能を予測するものである。データベースモデルは、あらかじめ測定あるいは計算してあるデータを取り出して利用するものである。チェックリストもこの一種である。これらのモデル設計のさまざまな過程で用いられる。

③ 設計過程で性能モデルを用いる設計関数

上記性能モデルは、構造変数を与えられた場合に、性能を予測するものである。しかし、設計問題は、性能目標を与えられた場合に、構造を生成することである。性能から構造を写像するような関数は一般には得られない。従って複雑な設計過程を、それぞれ設計部分問題を解く設計関数の集合であると考ええる。

設計関数は、評価、パラメトリック設計、構造的設計の三つに分類出来る。評価とは、構造が指定されている場合、性能を求めることに対応する。パラメトリック設計は、構造がほぼ指定されている場合、定量的パラメータを決定することである。構造的設計は、性能目標を達成するために構造の一部を変更することである。

心理学を用いると枠組みの中で、性能モデルが最も重要な位置を占める。言い換えれば、心理学の役割は設計者に性能モデルを提供することである。

4.7 おわりに

4.6で述べたのは、認知心理学と言うよりはむしろ心理学的アプローチに近い。しかし、入出力関係だけを扱う心理学的アプローチはむしろ不十分であり、入出力間の内部構造をも扱う認知心理学的アプローチがぜひ必要となる。内部構造を採るための情報源として、プロトコール情報が有力である。

アンダーソン[9]も心理的研究方法論に対して問題提起を投げかけている。反応時間などを扱うインプリメンテーションレベルと、問題解決の場でどのようなアルゴリズムが用いられているかを扱うアルゴリズムレベルとに分けるマー[10]の分類を紹介した上で、これまでの研究がほとんど前者のレベルであったのに対し、今後は後者にもっとウェイトを置くべきである主張している。具体的な情報源としては、プロトコール情報を重視している。このような考えに基づきアンダーソンはACT*モデル[11]をLISP学習[12]に適用している。

しかし、このプロトコール情報に関しては、その信頼性をめぐってさまざまな論争が行われた[15]。一つの新しい方法として、コンピュータの推論能力をもっと積極的に活用し、いわば「コンピュータ支援プロトコール」のような方向を摸索することが考えられる。

入出力間の内部構造を扱う方法論の一つとして、ピリシンは複雑等値性の概念を提案している[16]。これは入出力だけが合っていれば良いとする弱い等値性よりも、ずっと強い制約をモデルに与えたものである。

このように、認知心理学的アプローチの方法論に関しては、これからの課題が多く残されている。

ユーザフレンドリなヒューマンインタフェース実現のためにユーザモデルが用いられると言うのは、魅力的ではあるが、これまでがモデルを導入する試みがいくらか見られるだけで、まだ実用とははば遠いのが実状である。実用に至るまでにはまだ多くの基礎的研究が必要であり、真に実用的なユーザモデルの実現に向けての研究の進展を期待するものである。

参考文献

- [1] Norman, D.A. & Draper, S. W. : User Centered System Design, Lawrence Erlbaum Associates (1986).
- [2] Shneiderman, B. : Designing the User Interface, Addison Wesley (1987).
- [3] 石川, 他: 三段論法における人間の思考過程のモデル化, 人工知能学会誌, Vol. 3, No. 4, 486/493 (1988).
- [4] 石川, コネクションストアプローチによるインタフェースのユーザモデル, 人工知能学会誌掲載予定
- [5] Card, S. K. & Moran, T. P. : User technology: From pointing to pondering, Proceedings of the ACM Conference on the History of Personal Workstations, 183/198. Palo Alto (1968).
- [6] 石川真澄: コネクションズムの展望 (1) 概論, 情報処理, 29-7, 666/672 (1988).
- [7] 石川真澄: コネクションズムとシミュレーション, シミュレーション, 7-2, 83/90 (1988).
- [8] Rumelhart, D. E. et al: Parallel Distributed Processing, Vol.1 & Vol.2, The MIT Press (1986).
- [9] Anderson J. R. : Methodologies for studying human knowledge, Behavioral and Brain Sciences, 10, 467-505 (1987).
- [10] Marr, Vision, Freeman (1982).
- [11] Anderson, J. R. : The Architecture of Cognition, Harvard Univ. Press (1983).
- [12] Anderson, J. R. : Production systems, learning, and tutoring, ch9 in Klahr, D. et al eds: Production System Models of Learning and Development (1987).
- [13] Rumelhart, D. E. & Norman, D. A. : Simulating a skilled typist: A study of skilled cognitive-motor performance, Cognitive Science, 6, 1/36 (1982).
- [14] Card, S. K., Moran, T. P., & Newell, A. : The Psychology of Human-Computer Interaction (1983).

- [15] Ericsson, K. A. & Simon, H. A. : Protocol Analysis ... Verbal reports as data, The MIT Press (1984).
- [16] 佐伯監訳：認知科学の計算理論，産業図書（昭和 63 年）

5. ヒューマン インタフェースにおけるメンタルモデル

5.1 はじめに ー要約に替えてー

メンタルモデルそのものは心理学、実験心理等で古い歴史を持つ。

Kenneth Craik (1943) は既に次の様に書いている。

「生体が外界の実在について“縮尺されたモデル”を頭の中に持っていて、それを様々に動かしてみることができるなら、生体にとっていろいろなことが可能になる。さまざまな可能性を試してみ、その中の最良のものを選ぶこと……………(途中省略)……………危機に際して、それにすべての面から、より有能、安全、かつ十分に対応することなどは可能になる。」

一方コンピュータ・システムあるいは情報処理システムのヒューマン・インタフェースにおけるメンタルモデルが話題となり重要な研究対象となり始めたのは、1970年代後半のからである。これは Xerox の PARC における Workstation 概念の形成、および製品の開発と軌を一にして始まった。

本章ではメンタルモデルに関する6つの論文をトピック的に採り上げ、その梗概を紹介する。

- | | |
|------------------------------|---|
| (1) Johnson-Laird | 「認知科学におけるメンタルモデル」 |
| (2) Norman | “Some Observations on Mental Models” |
| (3) Rasmussen | “Mental Models: Aggregation, Abstraction, and Analogy” |
| (4) Boulay, O' Shea and Monk | “The black box inside the glass box: Presenting computing concepts to novices” |
| (5) Bayman and Mayer | “Instructional manipulation of users' mental models for electronic calculators” |
| (6) Yong | “The machine inside the machine: users' models of pocket calculators” |

(1)の Johnson-Laird の論文はメンタルモデルについての基礎的概念を考察したもので、認知科学形成に大きな役割を果たした論文である。人間の高級な知的活動である三段論文による推論、あるいは自然言語からの意味の解釈、イメージと心像の関係を説明する理論においてメンタルモデルが決定的に重要な役割を果たすことを論証している。

(2)の Norman の論文もまたメンタルモデルの基本的概念を考察したものであるが、Johnson-Laird に比べて多少プラグマティックである。メンタルモデルのヒューマン・インタフェース上での扱いを論述しており、ユーザのメンタルモデルは特性として、信念 (Belief) システム、外界との対応関係 (Observability)、システ

ムの動作の予測提供 (Predictive Power) の3つを持つべきことを主張している。

(3)の Rasmussen の論文は、比較的最近 ('86年) の論文である。人間は複雑な物理環境とのインタラクションにおいて、経験済みの結果を転用して (即ちモデル変換して)、新規情報の必要性を最小にするようメンタルモデルを修正しながら、認知処理を行っていることを述べている。このモデル変換として、Aggregation, Abstraction および Analogy の3つを人間は採用していると主張している。

(4)の Boulay 等の論文は、プログラミング言語を初心者に教える際のポイントとして Simplicity と Visibility の2つが重要と主張している。Simplicity はアナロジーにより実現され、Visibility は説明にテキスト、図のような判り易い表現導入により実現されると指摘している。

(5), (6)の論文は、計算器の操作を正しく行う上でシステム構造、動作イメージを与えておくことの重要性を実験的に論述したものである。実際の世界でのメンタルモデルの働きの有用性を示唆したものである。

5.2 メンタルモデルに関する6つの論文

5.2.1 Mental Model in Cognitive Science [1]

- (1) Johnson-Laird はメンタルモデル (Mental Model) の提唱者であり、この方面の第一人者である。(英 Sussex 大教授)

メンタルモデルの概念は心理的実験の結果およびコンピュータ・プログラムの開発から得たアイデアの両方を基礎に導出したもの。

- (2) メンタルモデルは、『生体（特に人間）が外界の実存について頭の中に持っている“縮尺されたモデル”』である[7]。
- (3) 本論文中で Johnson-Laird は3つの問題について論じている。
- 1) 心理的表現とはどういう形式によるのがよいか。また、イメージと命題の集合とは同じか、違うか。
 - 2) 日常的な推論の心理的過程はどのようなものか。その中にどのような推論規則が含まれているか。
 - 3) 言葉の意味をどう表現するか。意味表現は、意義素に基づく辞書による形をとるのがよいか、それとも命題の形をとる意味公準の集合に依拠するのがよいか。
- (4) 構成は「推論とメンタルモデル」「意味とメンタルモデル」「イメージ、命題そしてメンタルモデル」[8]-[10]「記述のレベル」「メンタルモデルと命題表現についての実験」および「結論」となっている。

Johnson-Laird は、推論、意味を正しく、深く捉えるにはメンタルモデルをベースに考究すべきであり、メンタルモデルは有用であり、本質的な概念であることをそれまでの様々な学説・考察と比較・検証しながら主張している。

説得力のある論説である。

- (5) 推論については、与えられた前提から結論を出すには論理以外の世界の原則が必要であり、これはメンタルモデルで操作することにより得られると言う。

三段論法推論の理論として具備条件としての

- 1) 習慣的反応の偏りを説明できること
- 2) 限量詞つき言明の扱いが可能なこと、および拡大利用できること
- 3) 子供がいかにして演繹的推論能力を獲得するかを説明できること
- 4) 形式論理学の発展と矛盾しないこと

の4条件もメンタルモデルが満すと主張。

- (6) メンタルモデルと命題表現の関係として

- 1) 命題表現は、事態を記述するもので、つまり心的言語による記号例である。
- 2) メンタルモデルは事態を表現するもので、その構築において2つの道を持ち、一つは上記の命題表現 即ち言語情報から構築され得るし、もう一つは知覚情報から直接構築できる。
- 3) メンタルモデルは情報を類推的に表現できる構造を持つ。即ちメンタルモデルが表現された次元空間の次元上の変数を制御することにより、直接構成したり、操作したり、走査することが可能である。

命題表現にこのような柔軟性はなく、表現の要素間に予め敷かれた一定の方向に走査できるのみである。

- (7) 意味については、例えば自然言語の文の意味を解釈する場合、命題表現では対応する心的言語で命題に翻訳し、心的語いの意味公準を用いて推論するが、自然言語のもつあいまいさから推移律が壊れる状況を克服できない。メンタルモデルでは、語をより原素的な単位へ分解して意味を把握する手続き的意

味論で明示するので、上記のような問題はない。

(8) 分解型意味論と意味公準との関係として

- 1) 「右」とか「左」とかの語の外延について、メンタルモデルでは、その構築および評価の一般的手続きとして分解型の手続きにより明示可能である。

意味公準では上記の語の外延を行える道具を持たない。

- 2) 推移律の気まぐれさのような現象もメンタルモデルではその構築から自然に出る性質であるが、意味公準ではこのような現象を処理できない。

意味公準は語の論理的性質や、語間の論理的関係を群述する規則に過ぎない。

- (9) イメージとメンタルモデルの関係：メンタルモデルは言語情報をもとにしても知覚情報をもとにしても、どちらからも構築されるが、イメージはメンタルモデルの知覚可能な部分の構築要素に対応する。

- (10) 以上簡約した如く、メンタルモデルは演繹的推論の基礎となる心理的な原則を説明することが可能なモデルである。

また、言語とイメージ、それらとメンタルモデルの関係は上記の如きであり、コンピュータのディスプレイ表示における文字列、アイコン等の表象の人間側での捉え方を考える上で基礎を提供する概念がメンタルモデルである。

5.2.2 Some Observations on Mental Models [2]

人は環境や技術的人工物とインタラクトする時、自分自身や対象についてのモデルを形成する。これを内的メンタル・モデルといい、インタラクション理解のために予想や説明の力を与える。メンタル・モデルにおいて4つのものを考慮しなければならない：すなわちターゲット・システム (t)、ターゲット・システムの概念モデル (conceptual model) ($C(t)$)、ターゲット・システムについてのユーザのメンタル・モデル ($M(t)$)、メンタル・モデルについての科学者の概念化 ($c(M(t))$) である。概念モデルはターゲット・システムについての正確で、一貫した、完全な表象を提供し、教師やデザイナー、科学者や工学者によって形成される。しかし、メンタル・モデルはインタラクションを通して形成されるものであり、機能的ではあるが正確である必要はなく、結果を得るまで修正を続ける。ユーザの技術的背景、類似システムにおける経験、人間の情報処理システム構造などに制約されるモデルである。

(1) Some Observations on Mental Models.

Observations of Calculator Usage

メンタル・モデルについて一般的特徴

1. 不完全
2. モデルを働かせる能力は制限されている
3. 不安定 (人はシステムの詳細を忘れやすい)
4. 確定的な境界がない (類似のデバイスやオペレーションが混同する)
5. 非科学的 (身体的エフォートが少なく心的エフォートが省けるなら、迷信的行動パターンを持続する)
6. 極度に儉約的 (心的複雑さより身体的アクションを好む)

(2) On modeling a Mental Model

On the Relationship between Conceptual and Mental Models

メンタル・モデルをモデル化する研究において必要とされる3特性。

1. Belief system :

メンタル・モデルは物理的システムについての belief を反映しているので、 $C(M(t))$ は人の belief システムのモデルを含むべきである。

2. Observability :

メンタル・モデルの変数・状態と、観察し得る物理的システムの外観・状態との間に対応関係がある。 $C(M(t))$ においても同様の対応関係が存在すべきである。

3. Predictive Power :

物理的システムの振舞いの理期と予測を人に提供することが、メンタル・モデルの目的である。 $C(M(t))$ は、そのようなメンタル・モデルを人が利用できるような、適切な人の情報処理と知識構造モデルを含むべきである。

(3) The System Image

理想的にはシステムは、システムと人のインタフェース全体を支配する概念モデルを基にしてデザインされる。その結果、システム・イメージはユーザにとって一貫した凝集力のあるわかりやすいものになる。このようなシステム・イメージは、マニュアル、教育、操作を通して一貫しているべきである。システム・イメージが概念モデルと一致しているなら、教示者が概念モデルを与えるとメンタル・モデルも一致した物となる。その場合の概念モデルは、Learnability, Functionality, Usabilityの基準を満たしているべきである。

5.2.3 The black box inside the glass box : presenting computing concepts to novices [4]

プログラミング言語を初心者に教える際のポイントとして、①simplicityと②visibilityが重要であると主張している。

著者たちは、プログラミング言語の構成要素から概念的に構築されるコンピュータのモデルという意味で、システムイメージに対応する概念を、notional machineと呼んでいる。彼らによると、simplicityを表現するためには、例えばアナロジーを導入することにより、notional machineに関して新たに学習すべき要素数を少なくすることが必要である。また、visibilityという概念は、必ずしも視覚化することを意味しているのではなく、notional machineの要素や機能を、テキストや絵によって判り易く表現すること (commentaryとも呼ぶ) を指し、その重要性を指摘している。

なお、標題のblack boxという用語は、ここでも一般的な意味と同じに使われているが、ここでは多少積極的に、システムの実際の内部動作を知る必要は無いという見方を代表しているようである。またglass boxという用語は、そうした暗箱について、入出力動作特性だけから、そのイメージを構成しようとする (また、それで充分ではないかという) 見方を代表しており、利用者に対しどのようなシステムイメージを構成させるかという点において注目すべき考え方といえよう。

5.2.4 Mental Models : Aggregation, Abstracting, and Analogy [3]

knowledge-based ドメインでの環境とのインタラクション・プランニングは、構造的知識と心的表象 (エレメントの構造的形態とそれらの機能的関係について) に依存している。これらのメンタル・モデルは、認知的タスク分析との関連から抽象階層における様々なレベルで存在する。

複雑な物理的環境とのインタラクションにおいて、先行結果の転用を最適に、新規情報の必要性を最少にするようにメンタル・モデルを特定のタスクに適合するように修正することが、情報処理や短期記憶容量の

制限から逃れる効率的な方法である。人間の認知処理はこの様なモデルの up-dating を伴うモデルの変換、内的動的な世界モデルによって定義される文脈に依存している。

(1) モデル変換のための戦略：

1. Aggregation：

表象におけるエレメントは文脈との familiarity が増すように、同モデル・カテゴリー内でより大きなユニット（チャンク）に集大成される。（part-whole 次元でのモデル変換）

2. Abstraction：

システムや環境特性の表象は、抽象のより高レベルのモデル・カテゴリーに変換される。重要な変換は means-end 次元に沿った機能的階層における変換である。

3. Analogies-use of ready-made solutions：

解決が既知であるかルールが解決に利用可能なモデル・カテゴリーに、表象は変換される。

(2) Aggregation

現実生活環境の複雑さを処理する方法は、物理的項目をより高次元の対象に集大成するか、高レベルの意図によって制御されるルーチンにアクション・パターンを統合するかである。aggregation によって事態を構造化して、問題を変換する。（図 5.2-1：whole-part と抽象－具象次元）

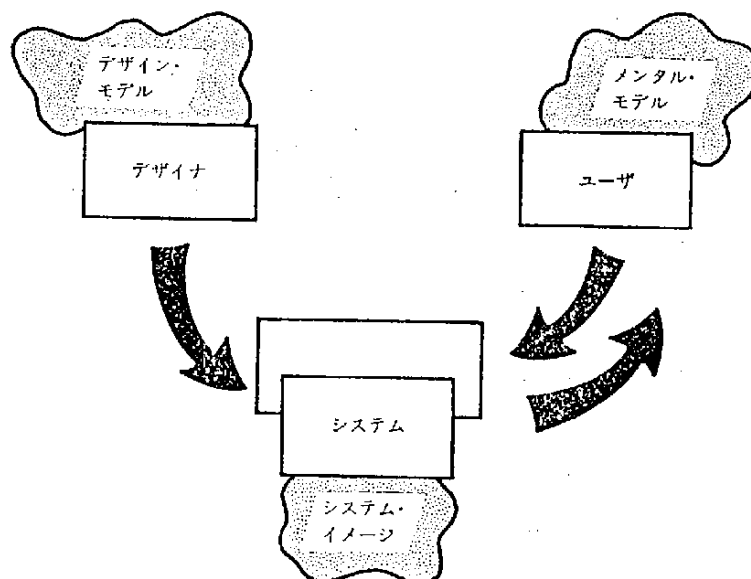


図 5.2-1 デザイナは、デザイン・モデルをもとにシステム・イメージをデザインする。ユーザはシステム・イメージからメンタルを形成する。

メンタル・モデル…ユーザが頭の中に描くシステムのモデル

デザイン・モデル…システムの設計者が頭の中に描くシステムのモデル

システム・イメージ…ユーザがシステムとのインタラクションより得るシステムの動作、入力特性やマニュアルの内容等より描くシステムのイメージ

出典：「インタフェースの科学」 知識情報シリーズ 5

淵 一博監修，共立出版

(3) Abstraction

機能的 means-end 階層において、システムの機能的特性は抽象のいくつかのレベルに沿って概念で示される (Fig. 2)。システムの物理的形式や配置を示す最低レベルから、物理的過程や機能の表象、機能的特性をより一般的概念で示すレベルなどである。

人間-機械システムにおける人間の重要な機能は、妨害や誤りの修正にある。この様な診断的検索において、不適切な機能の原因は抽象レベルの“bottom-up”で説明され、適切な機能の理由は機能的目的からの“top-down”で引き出される。また、技術的システム・デザイン・タスクには数個の抽象レベルでの表象が用いられる。システム・デザインは様々なレベルの熟考間のインタラクションの過程で、レベル間に many-to-many のマッピングが存在する。すなわち、ある目的は多くの物理的形態によって提供されるし、ある物理的システムは多くの目的を供給し様々な効果を有する。また、他の人との未知のインタラクションでは top-down 予期・常識的表象が役立つ。

(4) Analogy ; Transfer of Results and Rules

抽象の各レベルでは reasoning は、特定な型のモデルと情報処理ルールに依存する。これらを獲得するための戦略として、ある物理的形態に対する高レベルのモデルを、別の形態に関する問題解決に再解釈して用いる analogy が有効である。

(5) Problem Solving in a Means-end Abstraction Hierarchy

問題解決を考慮する場合、次の2種類の対象を区別する。物理的システムは causal システムであり、law of nature に従い抽象階層の bottom-up reasoning (functional analysis) で説明・予期される。一方、高度に自動的機能や自己体系化し適応的特徴を持つものは、意図的システムであり機能的分析では説明できない。このシステムの例としては人間や社会システムがさらに複雑な制御システムや情報処理システムも含まれる。制御の意思決定は top-down reasoning が基礎となる。

(6) Problem-Solving Strategies

Rumelhart & Norman (1981) は、4 カード問題と続いて同じ概念を実世界に近い状態 (bill-signing) で仮説検証タスクを与えた結果、後者での正当率が高かった。初めの問題は慣れていない事態で general reasoning 過程に頼るが、後者は事態の理解から特定の概念的制約を用いる。別の説明によると、前者は単一の抽象レベルでの formal, logical arguments すなわち syllogistic logic を基礎としている。一方、後者では文脈が意図的システムを定義してより高レベルでの推論が可能になり、top-down のモデル修正が解決に導くように問題を転移し得た。

この場合は analogical reasoning を用いているが、抽象階層のレベル変換は functional reasoning でも強力なツールである。この様な practical, functional reasoning は、単一レベルでの formal logical reasoning (syllogistic reasoning) との比較で哲学や心理学の分野で多く論議されてきた。

物理的デバイスの機能理解においては、目的や理由を考慮して全体の中での機能を理解する抽象階層における top-down 思考が、ある解決が正しいか否かという判断が要求されるとき有効である。

(7) Implications for the Design of Decision Support Systems

1. Structuring Representation in Database :

従来はボトムアップに過程の状態を提供する物理的状态のデータが提示されてきた。しかし、機能的情報や意図を示す情報などのようにトップダウン思考に必要な情報を、一貫した means-end 階層で示し、系統だった purpose / function / equipment 関係を位置づけるよう構造化すべきである。

2. Structuring Information Presentation :

タスクの諸相で考慮のレベルは means-end 階層において変動する。それゆえシステムからのデータは、レベルの機能構造のメンタル・モデルに合致した形式で提示され利用可能であるべきである。

3. Predicting Decision-Error Models :

decision エラーは確率的事象でなく問題空間の構造に依存している。decision エラーのモデル化と予想のためには、制御対象の means-end 関係を系統的に表示することが必要条件となる。

4. Planning of Experiments :

抽象階層のレベル変換に基づいた問題解決事態で real-life タスクや、文脈を付加した問題を実験に導入する。

(8) Skills, Rules, and Knowledge in Problem Solving

問題解決は knowledge-based 変形と、慣れているモデルの枠組みの rule-または skill-based 操作に依存している。knowledge-based レベルから rule-または skill-based レベルへの問題解決の転移は、操作のためのルールが利用可能な抽象レベルでのシンボルリック表象が再解決され得る時に生じる (図 5.2-3)

(9) The form of Mental Models

大別すると常識推論に用いられるモデルと、変数と関係に関してより formal モデルがある。前者では対象と状態と事象を基礎とした因果推論がなされ、診断問題事態のような伝播分析や、システム・デザインのようないくつかの非公式で粗略な reasoning に用いられる。他方、formal, deterministic reasoning は定義された物理変数の定量的表象の関係ネットワークのモデルと基礎としている。システムの正常で意図的連続的機能記述に適していて、数学的モデル化がこの枠組みでなされる。

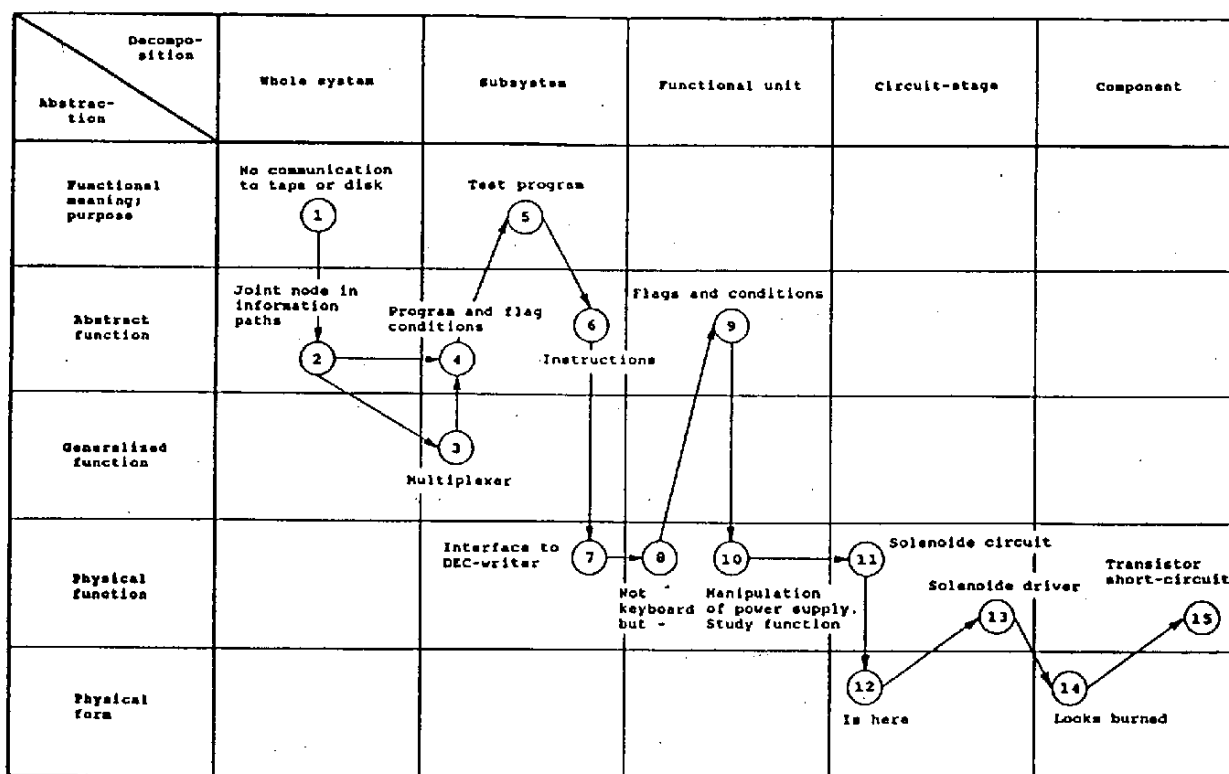


图 5.2-2 计算机故障排除者的注意力焦点可以随部分-整体维度和抽象程度的不同而变化。该图说明了一个实际诊断任务的跟踪过程。

出典：参考文献[3]

Functional purpose
Production flow models, system objectives, constraints, etc.

Abstract function
Causal structure: mass, energy, and information flow topology, etc.

Generalized functions
"Standard" functions and processes: feedback loops, heat transfer, etc.

Physical functions
Electrical, mechanical, chemical processes of components and equipment

Physical form
Physical appearance and anatomy; material and form; locations, etc.

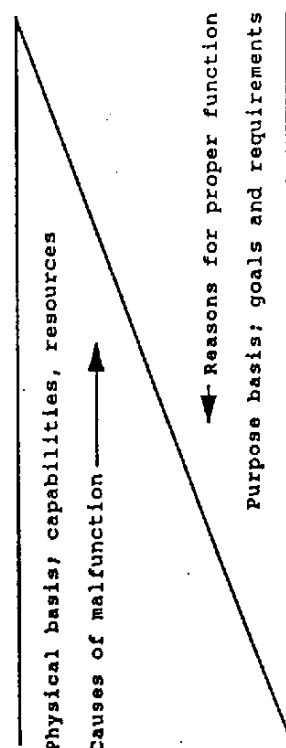


图 5.2-3 系统的功能特性可以在抽象层次中用几个层次表示。

出典：参考文献[3]

Knowledge-based performance Depends on semantic content * Symbolic model, strategy, and data transformation used to achieve stated goal * Search for transformation rules to find path to solution while considering goal, direction to goal, discrepancy between goal and actual state of affairs * Used for problem solving and for monitoring and correction of rule- and skill-based activities			
* Stereotype operation on the physical world * Recognition of states and recall of stored rules * Goals are implicit in task context	Oriented towards appearance of physical world	Rule-based performance Depends on syntactic form	* Stereotype operation on symbol and model states * Model states act as signs associating familiar rules * Goals are implicit in task context
* Automated sensorimotor manipulation of physical objects		Skill-based performance Depends on physical form	* Automated sensorimotor manipulation of symbols appearing as artificial objects

図 5.2-4 Skills, rules, and knowledge in a manual task and in problem solving.

出典：参考文献[3]

5.2.5 Instructional manipulation of users' mental models for electronic calculators [5]

メンタル・モデルは、マシンの要素と操作ルールから成るメタファであり、新しいコンピュータや計算機言語の使用学習においてユーザが形成する。入力と出力の間で見えない (invisible) 情報処理ステートや変換をユーザに想像させる。ユーザがマシンの現在の状態を推測し、与えられたコマンド系列がどの様にある出力を生じるかを推論しなければならない時に重要になる。

メンタル・モデルの 3 次元

- (1) Usefulness：ユーザによる洗練された逐行を支援する程度
- (2) Completeness：要素と操作ルールに関する詳細を全て含むとモデルは完全といえる
- (3) Veridicality：要素がマシンの実際の物理的デザインに、操作ルールがマシン言語に対応付けられる程度

第一の次元はユーザに重要であるが、他の 2 次元は電子工学の熟練者にとって重要である。

Usefulness の次元で異なるモデル

- (1) Incorrect performance：簡単な式では矛盾しないが、それを超えて推定すると誤った逐行を導くモデルを有している (ナンバー・キーを押すとすぐ式の数値が求められる)
- (2) Impoverished performance：簡単な式を超えた推定が常に正しいとは限らないようなモデル (等号キーを押すと答えが得られる)
- (3) Sophisticated performance：簡単な式で矛盾が無く、複雑な式も有用に推定できるモデル (演算キーが数値を求めるのに効いている)

(1) 実験

電卓の有用なモデルをユーザに明白に教えることができるかを決定することが目的である。内的レジスタ

を示す図式の紹介を与える 2 種類の実験条件と、内的レジスタの記述が無い条件を設定した。図式紹介がユーザにより有用なメンタル・モデルを構築させるなら、図式が提示されたユーザは sophisticated 逐行が多くて他は少なくなるであろう。

(2) 方法：

被験者は 72 名のカジュアル・ユーザで、図式が提示される条件の line 群と stack 群、提示されない条件の no model 群の 3 群に 24 名ずつ振り分けられた。実験前に年齢や電卓についての経験などの質問を行っている。

まず被験者に各群で内容の異なる教示冊子を与えた。1 頁目は共通で 4 機能を持つ電卓が示されている。2 頁目に、図 5.2-1 に示す内的レジスタの図式をそれぞれ提示した。line 群は数値のためのラインと演算子のための箱が交互に並ぶレジスタと答えを表示するディスプレイレジスタが、stack 群には数値のための x^- , y^- レジスタとディスプレイレジスタ、そして演算子レジスタが提示されている。課題は最後のキー・プレス後に、各レジスタとディスプレイに保持されている数値と記号を示すことである。3 頁目には例題が示されている。no model 群には他の群の 2 頁目に相当するものはなく、最後のキー・プレス後にディスプレイされるものを書き込む課題を与えた。全群同じ問題を 25 個、冊子で提示した。問題は 3 種類に分類され (図 5.2-2)、分析された。

(3) 結果：

図 5.2-2 に示される 3 種類の問題に対して、それぞれ incorrect, impoverished, Sophisticated である戦略を用いた人数を分析した。カイ自乗検定によって問題 1-10 と 11-23 において、内的レジスタ図式を提示した群が、提示しない群より有意に多く sophisticated 戦略を採用していることが示された。

(4) 結論：

1. ユーザは同様の経験をしてもそのメンタル・モデルは異なっている (no model 群の結果)。
2. ユーザはしばしば impoverished または incorrect モデルを発展させる。
3. 教示を介在させてより有用なメンタル・モデルの発展を促進することができる。

Mayer & Bayman (1981) は、プログラミング・エキスパートがノビスより sophisticated 戦略を好むことを示している。本実験でカジュアル・ユーザが簡単なモデル提示だけで、エキスパートのような傾向を示すことが認められた。さらにメンタル・モデルをユーザに明白に教示することが可能で、そのモデルは sophisticated 逐行のレベルを上昇し得た。さらに大多数の個人の助けになる良いメンタル・モデルを研究しなければならない。

(5) 電子計算デバイスに対する教示の与え方と評価の方法に対する推奨：

1. ノビス・ユーザに内的な要素と操作ルールについての単純で有用なモデルを与える。
2. モデルを hands-on 経験やマニュアル教示と関連づけるようにユーザを導く。
3. モデルの有用性は評価されるべきである。

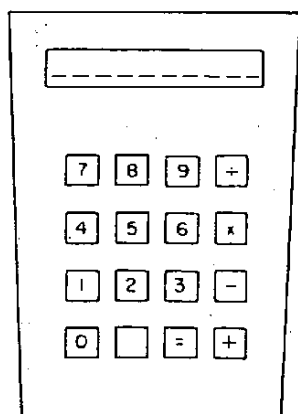


図 5.2-5 The four-function calculator.

出典：参考文献 [5]

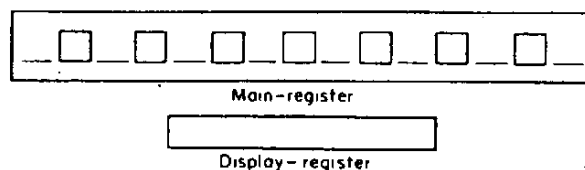


図 5.2-6 The line model.

出典：参考文献 [5]

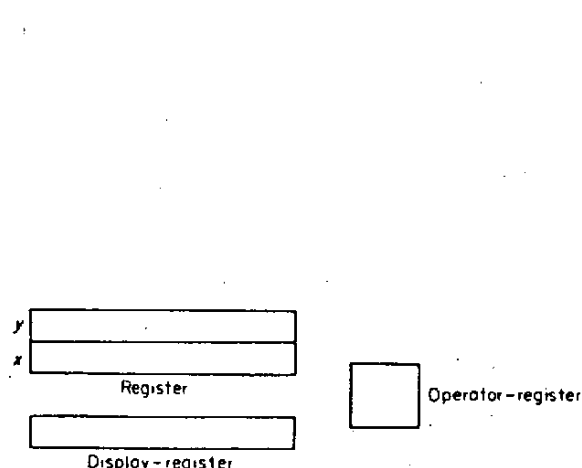


図 5.2-7 The stack model.

出典：参考文献 [5]

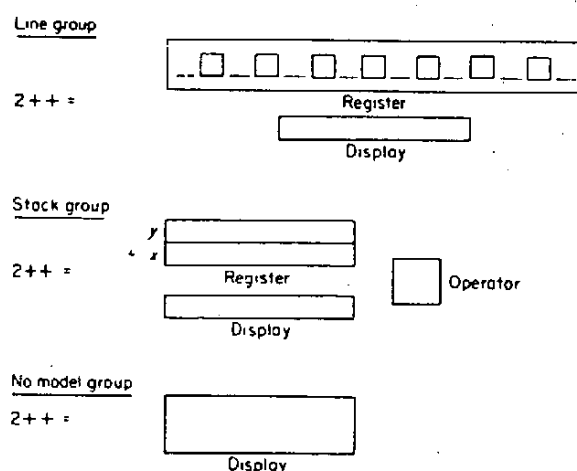


図 5.2-8 Examples of problems of line, stack and no model groups.

出典：参考文献 [5]

5.2.6 The machine inside the machine : users' models of pocket calculators [6]

対話型デバイスのユーザは、そのデバイスに対する“概念モデル” (conceptual model) を形成する。このモデルの意味は曖昧なものであるが、いずれにしろユーザはデバイスの動作を理解しやすくしたり、操作の助けとなるような何等かの内的表象ないしメタファを抱く。この論文はポケット計算機概念モデルについて検討したものである。

逆ポーランド記法 (Reverse Polish Notation : RPN), 四則演算 (Four Function : FF), 代数演算 (Algebraic : ALG) という 3 種類のポケット計算機の内部動作は、演算子や演算数の入力に対する反応を予測する暗黙的レジスタ・モデル (implied register models) と呼ぶ概念モデルによって表現され得る。これは RPN 計算機の場合、4 個の演算数レジスタからなるスタック・モデルという形をとれば、計算機の動作は完全に予測できる。

一方、FF 計算機については、演算数と演算子のレジスタそれぞれ 1 個からなるモデルによって単純加算、累算、連鎖計算 ($2 + 3 = + 4 =$) は予測できても、定数項を含む計算 ($2 + 3 = 4 =$ の結果は 9 でなく 7

となり、3が定数として扱われる)は予測できず、移動ポインタを加えたモデルが必要となる。この定数項計算は逆数を求めるには便利とされているが、逆数計算では新たに上下2段レジスタ・モデルでないと予測できない。しかも、演算数と演算子の並びが複雑でユーザには到底理解できない。第3番目の計算機ALGの算術演算子(括弧を含む)には優先順位があり、計算させるための言語的“ステートメント”を記述すればいいだけで、計算機の内部動作に関するモデルを形成する必要はない。ステートメントの記述自体がモデルと見なせる。

こうした暗黙的レジスタ・モデルは、FF計算機の応答を記述することはできるが、このモデルをユーザが形成するのは困難であり、ユーザにとっての計算機の使いにくさを表現することはできない。また、ALG計算機の動作を説明することもできない。ユーザのアクションと計算機の動作(タスク)との対応関係を記述するためには、Moran (1981)の“Command Language Grammar”(CLG)のような解析法、インタフェースの“level”といった概念が必要とされる。次のタスク/アクション・マッピング(Task/Action Mapping)はこのCLGに類似したものであり、①計算の実行についてはタスク・アレナ(Task Arena)、②ユーザの行うアクションについてアクション・アレナ(Action Arena)、③計算機の構造記述については抽象的マシン・アレナ(Abstract Machine Arena)という3つのレベルから検討する。アレナは心理学的な構成概念であるが、これには実体(entities)と手続き(procedure)に関する記述が含まれている。マッピングとは、これらのアレナの対応関係を記述するものであり、タスク・アレナには、前述の累算、定数項計算、逆数計算などがあり、計算順序を変えるなどの操作は手続きに含まれる。アクション・アレナには計算機への入力や計算機からの出力などがある。抽象的マシン・アレナは計算機内部の計算過程を示すものである。

FF計算機のタスク・アレナとアクション・アレナのマッピングを考えると、両アレナに1対1対応がある場合とない場合が生ずるが、RPN計算機では直接的に対応している。計算機のデザインの良否およびユーザの使いやすさは、こうした両アレナの対応性が影響している。ただし、計算課題には、written formula taskとlocally structured basic calculations taskと2種類のあり、前者の場合にはFF計算機が、後者のばあいにはRPN計算機がマッピング分析から適切であることがわかった。すなわち、後者の心的なレベルでの計算タスクでは、RPNはユーザの計算機に対する操作と実行されるタスクが直接的にマッチする可能性を与えるのである。ただし、定数項のある計算では直接的マッチは無理である。マッチしていない場合には、心的な変換操作を経てタスクとアクションの対応をとる必要があり、それが計算機の使いにくさの原因となる。

参考文献

- [1] Johnson-Laird, P. N. : Mental Model in Cognitive Science,
Norman D. A (Ed), Perspectives on Cognitive Science, Cambridge U. P
(邦訳:「認知科学におけるメンタルモデル」, 認知科学の展望, 第6章, PP. 185-250).
- [2] Norman, D. A. : Some Observations on Mental Models,
In Gentnev, D. & Stevens, A. L (Ed), Mental Models, LEA, PP. 7-14, (1983).
- [3] Rasmussen, J. : Mental Models: Aggregation, Abstraction and Analogy,
In Information Processing and human-machine Interaction, PP. 117-148, (1986).
- [4] Boulay, B. D. et al : The black box inside the glass box : Presenting Computing
Concepts to novices, Int. J. Man-Machine Studies, Vol. 14, PP. 237-249, (1981).
- [5] Bayman, P. et al : Instructional manipulation of users, mental models for electronic calculators, Int. J. Man-Machine Studies, Vol. 20, PP. 189-199, (1984).
- [6] Young, R. M. : The machine inside the machine : users, models of pocket calculators, Int. J. Man-Machine

Studies, Vol. 15 PP. 51-85, (1981).

- [7] 高野陽太郎：心像の概念的考察，心理学評論，Vol. 24, PP. 66-84, (1981).
- [8] 宮崎清孝：メンタル・イメージは絵か命題かー認知心理学でのメンタル・イメージ論争について，教育心理学年報，Vol. 19, PP. 112-124, (1980).
- [9] Richardson, A. : Mental Imagery, Routledge & Kegan Paul, (1969).
- [10] Anderson, J. R. : Arguments concerning representations for mental imagery, Psychological Review, Vol. 80, PP. 249-277, (1978).

6. ヒューマンインタフェースにおける問題解決支援

6.1 個人レベルでの問題解決支援

コンピュータは、人間の知的活動を助ける道具であり、人間が問題を解決するのを支援してくれる道具である。

人間は、問題解決に向けて、部分問題に分け、その部分問題を解決するためにコンピュータシステムを活用し、コンピュータと対話しながら、問題解決を図っていく。

このような時、次の3レベルのヒューマンインタフェースがある。

まず、第1は、ユーザがコンピュータシステムの使い方、機能について理解する「システムを理解するためのヒューマンインタフェース」である。

次が、問題解決のために、ユーザとコンピュータとが対話する時、ユーザにとって使い易くするよう「気くばりのヒューマンインタフェース」である。

最後が、コンピュータが知識と問題解決手法を持ち、ユーザの状況を把握し、問題解決に導く「知的ヒューマンインタフェース」である。

以下、この3レベルのヒューマンインタフェースについての技術動向を、事例主体に述べる。

6.1.1 システムを理解するためのヒューマンインタフェース

コンピュータシステムを問題解決に活用する時、まず、ユーザは、コンピュータシステムの使い方、機能について理解する必要がある。これには、ユーザから見て、ツールがどのように見えるか、ユーザがツールに対してどのようなメンタルモデルを形成して理解するかが重要なポイントとなる。

従来、コンピュータとユーザとのインタフェースは、コマンド操作によるものが主体であった。また、操作の結果も簡単なメッセージを出す程度のものが多かった。

このようなインタフェースでは、ユーザは、ツールの各コマンドをいちいち覚えなないことにはそのツールが使えない、操作結果がわからないから、問題解決に向けて次に何をすべきかその場で考えることができない、といった欠点があった。

このような欠点を克服するものとして、「ツールの使い勝手を人間の習慣に合わせた」ツールが出現してきた[1]。人間が（コンピュータなしで）行っている作業、例えば、紙とエンピツで行っている作業を、コンピュータ上で、手で行うのと同じような感覚で行えるようにするのである。コンピュータ上に対象としているものを出来るだけ人間の世界にあるものと同じイメージで表示し、その上で人間の手が行うのと出来るだけ同じ操作で直接操作し、その結果を表示している対象物上に直接反映する。このようなインタフェースは、「ダイレクトマニピュレーション」[2]、あるいは「WYSIWYG」(What You See Is What You Get)と呼ばれる。

このようなインタフェースによって、ユーザは、コンピュータ上に表示されている物に対応する人間世界にある実体を思い浮かべ、その実体の持つ機能、性質から、コンピュータ上の表示物（ツール）の機能、性質、操作方法を、人間の習慣から連想することができる。ユーザがコンピュータの表示物に対するこのよう

なメンタルモデルを持つことによって、「何も教えてもらわなくても、今までの習慣に従って、コンピュータを操作できる」ようになる。

以下にこのような具体例を示す。

(1) ディスプレイエディタ

テキストエディタは、かつてはラインエディタであった。今やディスプレイエディタであり、EMACS [3][4], vi, SPF [5]等が使われている。もはやラインエディタに戻ることは出来ないだろうし、ワープロが普及したのもディスプレイエディタのおかげである。

B. Shneiderman [2]は、ラインエディタと比較したディスプレイエディタの特長を以下のように述べている。

- ・テキストがフルスクリーンに表示され、内容が見易く、探し易い
- ・最終結果が表示され、間違いが分かり易い
- ・カーソルが見え、操作のフォーカスポイントが明らかである
- ・矢印キー等によるカーソル移動操作は自然である
- ・各操作に対応したラベル付ボタンは複雑なコマンドを覚えるのを不要にする
- ・操作の結果がすぐ反映され、操作結果の確認作業が不要である。

```

EDIT --- SPFDEMO.MYLIB.PLI(TESTDIR) - 01.03 ----- COLUMNS 001 072
COMMAND ==>                                     SCROLL ==> HALF
000700      IF ERROR-FLAG THEN
CC          DO;
000900          PUT FILE(SYSPRINT)
001000          EDIT('TEST NOTES ERROR RETURN FROM D-I-R')
001100              (COLUMN(21),A(34));
001200          PUT FILE(SYSPRINT) SKIP(2)
001300          DATA(ERROR-FLAG,CARD-IMAGE,ANSWER);
001400          ERROR-FLAG = '0'B;
CC          END;
001600      ELSE
001700      ;
001800      END;
A_          ELSE
002000      ;
002100      END;
002200  END TESTDIR;
***** ***** BOTTOM OF DATA *****

EDIT --- SPFDEMO.MYLIB.PLI(TESTDIR) - 01.03 ----- COLUMNS 001 072
COMMAND ==>                                     SCROLL ==> HALF
000700      IF ERROR-FLAG THEN
000800      DO;
000900          PUT FILE(SYSPRINT)
001000          EDIT('TEST NOTES ERROR RETURN FROM D-I-R')
001100              (COLUMN(21),A(34));
001200          PUT FILE(SYSPRINT) SKIP(2)
001300          DATA(ERROR-FLAG,CARD-IMAGE,ANSWER);
001400          ERROR-FLAG = '0'B;
001500      END;
001600      ELSE
001700      ;
001800      END;
001900      ELSE
001910      DO;
001920          PUT FILE(SYSPRINT)
001930          EDIT('TEST NOTES ERROR RETURN FROM D-I-R')
001940              (COLUMN(21),A(34));
001950          PUT FILE(SYSPRINT) SKIP(2)
001960          DATA(ERROR-FLAG,CARD-IMAGE,ANSWER);
001970          ERROR-FLAG = '0'B;
001980      END;
***** ***** BOTTOM OF DATA *****

```

図 6.1-1 SPF のコピー操作例

(2) スプレッドシート

スプレッドシートは、集計表をコンピュータ上に実現したものである。Visicalc, Lotus 1-2-3[6] がこれの代表的なものである。

スプレッドシートでは、集計表がディスプレイ上に表示され、その表上で、例えば、カラム 4 はカラム 1 からカラム 3 までの合計と指示しておくと、カラム 1, 2, 3 の値が変わると自動的に合計を再計算し、カラム 4 に表示する。

D7: =SUM(D3..D6) 入力

	A	B	C	D	E	F
1		更新				
2	映画名	1週	2週	3週	合計	
3	京愁	31,818	21,206	14,480	67,504	
4	第三の男	20,064	19,930	17,392	57,386	
5	北北西に進路を取れ	0	20,713	13,755	34,468	
6	1984年	16,401	24,719	16,962	58,082	
7	週間合計	68,283	86,568	62,589	217,440	
8						
9						
10						
11						
12						
13						

図6.1-2 Lotus 1-2-3 の画面例

(3) 情報検索

情報検索を一覧表の形でコンピュータ上に実現したのが QBE である。

一覧表がディスプレイ上に表示され、そのカラムに検索したい条件を指定すると、その条件にあったレコードが一覧表の形で表示される。

(4) ハイパーメディア

ハイパーメディアは、本、文献等にある注釈や引用を用いて、関連ある他の情報を探す操作をコンピュータ上に実現したものである。このようなツールには、NoteCards [7], HyperCard [8] がある。

ハイパーメディアの構成とメカニズムは次のようになっている。情報を小さなまとまりのある単位に分ける。このひとまとまりの情報をノードと呼ぶ。このノードと関連のあるノードをリンクで結ぶ。こうするとネットワーク状に関連付けられたデータベースが出来上がる。このハイパーメディアデータベースに対して、情報を入力するためのエディタや、情報の関連付けを表示するブラウザなどのデータベース管理機能を付ける。この時、扱う情報がテキストなのがハイパーテキストであり、テキスト、グラフィック

File Edit Go Tools Objects

Corporate Staff Directory See Map

Name: Timothy Newton ← →

Position: Vice President, Corporate Planning [Icon]

Department: Executive Administration Sort

Extension: 5999 [Icon] Print

Location: North Building, 312

find ↑

図6.1-3 Hyper Card の画面例

ス、イメージ、音などのマルチメディアまで拡張したのがハイパーメディアである。

このようなハイパーメディアシステムでは、ユーザは、情報を表示し、そこにある関連情報を注釈を見るように、次から次へと調べるといった自然な操作ができる。

(5) ウィンドウシステム

ウィンドウシステムは、プログラムからよく使われる汎用的なヒューマンインタフェースの部品群を提供している。その部品群の中に、人間の世界にある実体を表現して、その機能、操作を連想させる部品が幾つ

も見られる。

例えば、OPENLOOK [9]には、プッシュピン、エレベータがある。

プッシュピンは、ポップアップメニューを一回だけでなく、何回も使う場合、その間中、ポップアップメニューを表示し、すぐ使える状態にしておきたい場合に使う。ポップアップメニューについているピンを指すことによって、そのメニューが表示され続けられ、ピンを抜くことによって、そのメニューは選択すると消えることを意味する。

エレベータは、ウィンドウに表示されている部分がどこであるかを示し、そのエレベータを動かすことによって、ウィンドウ内に表示される部分を変える部品である。ウィンドウ内の表示部分のコントロールをエレベータに見立てて行っている。

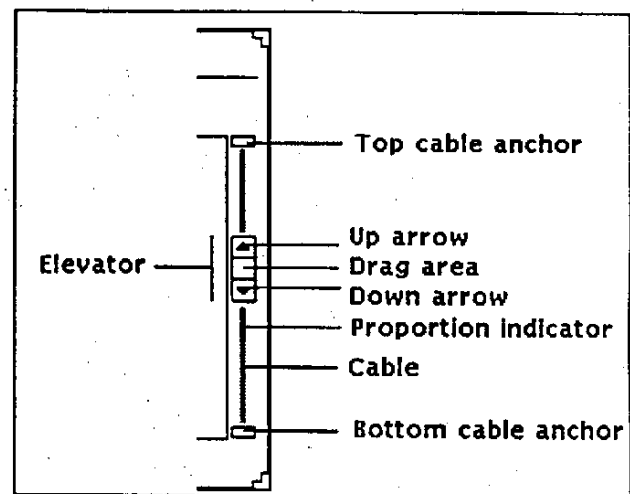


図 6.1-4 OPENLOOK のエレベータ

6.1.2 気くばりのヒューマンインタフェース

前節では、問題解決を支援するツールの機能をユーザが理解するためのヒューマンインタフェースについて述べた。ユーザがツールの機能を理解した後、ユーザはコンピュータシステムと対話しながら問題解決を図って行く。この時、コンピュータシステムの使い易さ、操作性が重要である。

使い易くするためには、当然、前節で述べたことも重要である。

また、キータッチを最小にするといった人間工学（エルゴノミックス）的アプローチがある。これについては、本節では深く言及しない。

本節では、使い易くするためのアプローチとして「気くばりのヒューマンインタフェース」[10]について述べる。

コンピュータシステムにとって、ユーザは人間であり、その人間に対して気くばりをして上げるのは当然であり、エラーに対する気くばり、次に予想される操作に対する気くばり、ユーザ好みの操作性に対する気くばり等を行う必要がある。以下のこのような気くばりについてのべる。

(1) エラーに対する気くばり

ユーザがツールを使う場合、必ずしも正しい使い方をするとは限らない。間違ってエラー等を生じたりする。このような状況に対して、それを親切に教える、間違いを少なくするといったエラーに対する気くばりが必要である。

C. LEWIS と D. A. NORMAN は、「Designing for Error」[11]で次のように述べている。

エラーには間違いによるものと、過失によるものがある。

過失によるものには、モードエラー（Mode Error）、記述エラー（Description Error）、捕獲エラー（Capture Error）があり、これらのエラーを起こさないようにするための方策を述べている。例えば、モードエラーは、システムがどのようなモードにあるかを誤って判断することによって起きるエラーであり、これには、システムのモードをユーザに明確にフィードバックすればよい。また、操作に似ているために、意図し

た行為とは違った操作をしてしまうエラーが記述エラーであり、それには、システムが操作体系を互いに区別し易い体系にすればよい。

また、エラーの示し方について、停止 (Gag)、警告 (Warning)、何もしない (Do noting)、自己修正 (Self correct)、話し合い (Lets' talk about it)、教えて (Teach me) の6種類があり、それぞれの良い例を示している。例えば、何もしない例として、Pinball Construction Setを上げ、そのシステムでは、色を変えることが出来ないものに対して色を変えようとしても、何も起こらないが、これでも、ユーザは、操作しても何も起こらなければ、エラーであると気付く。また、自己修正の例として、INTERLISP DのDWINを上げ、このシステムでは、スペルミスは、自動的に修正してしまう。それでよければ、そのまま続け、それで悪い場合には、UNDO 操作により、戻せば良い。

(2) 構造エディタ

エラーが起こるのを少なくし、入力情報を分かり易くするアプローチに構造エディタがある。

構文エディタ [12] は、編集対象をテキスト (単なる文字の列) としてではなく、編集対象が持つ意味的な情報を構造的に表し、ユーザからの編集操作をこれらの構造に対する操作として実現しているエディタである。

構造エディタには種々のものがあるが、入力方式からみると、テキストエディタと同様にテキストを直接入力するものと、テンプレートと呼ぶ一種の型紙に従って入力するものがある。

前者の例には、MENTOR、EMACS等があり、後者の例には、C. P. S. 等がある。

MENTORは、PASCALのソースプログラムのテキストを文字列として読み込んで、内部ではそれを構文解析し構文木を作成する。この過程で構文上の誤りが検出されれば、ただちにメッセージが表示される。

CPSは、PL/IのサブセットであるPL/CSプログラムを対象としている。ユーザはエディタに用意されたテンプレートを選択し、必要な箇所に情報を埋め込んでいく。例えば、IF文のテンプレートを選択すると、IF (condition) THEN statement ELSE statementが表示され、condition、statementの部分を埋めていく。

このようなテンプレート型の構造エディタは、確かに間違いが少なく、分かり易い利点をもつ。しかし、ユーザは、常に構造を意識して編集しているわけではなく、構造を無視してカーソルを移動するなどテキストベースで操作していることがある。したがって「テキスト+構造」型のエディタが妥当であるとの意見がある。

(3) 次に予想される操作に対する気くばり

問題解決する時、それを部分問題に分け、部分問題を段階的に解いて行って、最終的に目的とする結果を得る。このようなやり方をする場合には、システムは、出力結果に対して、ユーザは次に何を行おうかを先読みし、それに対する操作が行い易いように準備してあると使い易い。例えば、会議室の予約システムにおいて、指定日時、場所の会議室が既に予約済の場合、次に行われる操作として、他の日時の空き状況を見る、他の会議室の空き状況を見る、既に予約済の担当者を知る (既に予約済のが変更できないか交渉する) といった操作が予想される。このような場合には、このような操作がすぐに行いやすいように、例えばメニューを用意しておく。

ハイパーメディアは、情報の中に他の情報とのリンクをはる手段によって、このような気くばりの操作を実現するしぐみを提供しているとも言える。

(4) ユーザの好みに合わせた気くばり

操作性をユーザの好みにあった方法が選べるようにするといった気くばりも必要であろう。

EMACSは、このような気くばりを実現する仕組みを備えている。EMACSはシステムの内部にキーマップというテーブルを持ち、キーとコマンドとの対応を管理している。ユーザはこのキーマップを自由に変更することができる。これによって、キー操作をviの操作に変えることが出来たり、LISPとC言語向きに適した操作性に変更することができる。

6.1.3 知的ヒューマンインタフェース

問題解決を図る過程で、コンピュータシステムを比較的単純な情報検索、演算、編集を行う道具として使うだけでなく、コンピュータ自身が知識、問題解決手法を持ち、問題解決まで行う高度な知的システムがある。このようなシステムでは、ユーザに対して、コンピュータシステムが行う問題解決の過程をユーザに示すことがポイントとなり、このような観点から「知的ヒューマンインタフェース」について述べる。

(1) エキスパートシステム

エキスパートシステムは、コンピュータ自身が知識を持ち、それを用いて推論処理を行い、専門家と同じように知的な仕事をするシステムであり、既に数多くの実用的なエキスパートシステムが構築されている。

このようなエキスパートシステムによって、ユーザは高度な意味を持った結果を得、専門家と同じように問題解決をすることができる。そのエキスパートシステムのヒューマンインタフェースは、以下のような特徴がある。

単純な作業をツールでは、ユーザは、得られた結果に対してどのようにしてその結果が得られたのかを理解しているが、エキスパートシステムでは、得られた結果に対して、どのようにしてその結果がえられたのかがつかめない。従って、ユーザの求めに応じて、何故その結果が導かれたかを示す説明機構インタフェース、あるいは問題解決の過程をユーザに示しながら、結果を導く問題解決過程表示インタフェースが必要となる。

第2の特徴は、単純な作業をするツールでは、何を入力すればよいかわかるユーザには分かっており、ユーザ主導型のダイレクトマニピュレーションの方が使い易いのに対して、エキスパートシステムでは、ユーザは問題解決に対してどのような情報をインプットする必要があるかが分からない。従って、システム側から、ユーザ入力を促すシステム主導型のインタフェースが必要である。例えば、自動車の故障診断システムでは、システム側から、故障の主症状を聞き、それから原因を推定し、検証するために細部の状況を聞くことが行われる。

第3の特徴は、システムの柔軟性である。エキスパートシステムでは、専門家の問題解決に関するノウハウ知識が入っている。このノウハウは、まわりの情勢が変化するとつれ変更されるし、システムが成長するにつれて変更していかなければならない。このため、知識（特にノウハウ知識）に対して、頻繁な変更に耐えられるようにインタフェースを用意する必要がある。

(2) 自然言語による自動プログラミング

問題解決に向けて、既存の問題解決ツールでは、ユーザにとって満足のいく結果が得られないケースが出てくる。そのような場合、要求仕様を入力し、問題を解決してくれるプログラムを自動的に作成することによって解決を図ろうとするアプローチが考えられる。

要求仕様をどのように書くかが問題であるが、要求仕様の段階では曖昧な部分が多く、形式的に記述するのが困難である。従って、自然言語による要求仕様記述からプログラムを作成するインタフェースとなる。このようなインタフェースにおいては、要求仕様自体の論理的な曖昧さ、自然言語自体が持つ曖昧さを解決しなければならない。

このようなアプローチのシステムとしてKIPS [13], [14]がある。KIPSでは、曖昧な箇所を検出した時点で、ユーザに対してメッセージとして幾つかの可能性のある候補を示し、ユーザにその中から選択してもらうようになっている。また、KIPSが理解した要求仕様をプログラムチャートによる図形表現で出し、ユーザが意図していることとの確認を得る仕組みとなっている。

(3) ビジュアルプログラミング

既存のツールでは満足がいかない場合、ビジュアルなプログラム部品を合成することによって、問題解決ができるシステムを作ろうとするビジュアルプログラミングのアプローチもある。

このようなシステムの例として、Intelligent Pad [15]がある。このシステムでは、予め、ボタン、スイッチ、アナログメータ、テーブル、バーチャートといったビジュアルなプログラム部品が用意されており、これをIntelligent Padと呼ぶ。台紙に紙を張りつけていく操作と同じように、Intelligent Padを張り合わせることによって、ビジュアルなプログラムシステムを作り上げていく。このように、Intelligent Padでは、プログラム部品であるPadによって、問題解決を図る過程がビジュアル的に示されている。

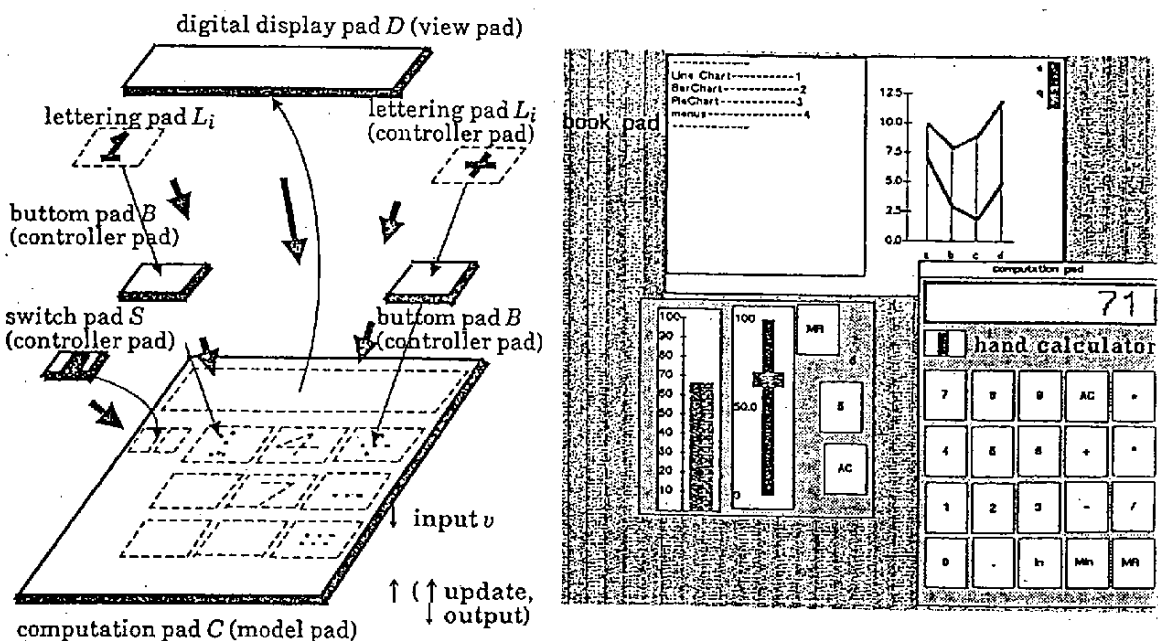


図 6.1-5 Intelligent Padにおける電卓の作成とその画面例

(4) 知的CAI

問題解決システムにおいて、真に知的なヒューマンインタフェースとは、単に問題解決を行うシステムではなく、人間の知的活動を強化、拡大するのを目的とするものでなくてはならない。すなわち、知識を生み出す過程のインタフェースが必要であり、知的CAI的なアプローチが必要である。

これを目指すものとしてGUIDON 2 [16]がある。これは、「発見的分類」を行うという問題解決の能力を

教育する汎用なgenericなシステムである。対象としている分野は、感染症の診断と処方提案する（このエキスパートシステムとして MYCIN がある。）のを教育することである。これは、HERACLES という「発見的分類」を行う分野独立なシステム、学習者の解答を解釈して、彼が用いたであろうプランニングを再構成する IMAGE、エキスパートの入力に対し、自分の欠けている知識、違っている関係を検出する ODYSSEUS、アニメーションとウィンドウを利用して、HERACLES の推論戦略を学習者に見られるようにする GUIDON-WATCH、学習者自身が問題解決オペレータを操作し、結果をグラフィックな表示によってみながら、問題解決の筋道を成長させてゆくことができる支援環境である GUIDON-MANAGE、学習者がシステムによる問題解決過程を評価することによって、システムの知識ベースを改定することができる GUIDON-DEBUG から構成されている。

参考文献

- [1] ハイパーメディア・システム, NIKKEI ELECTRONICS 1989. 2. 20, pp 167-174
- [2] B. Shneiderman: Direct Manipulation: A Step Beyond Programming Languages, IEEE Computer, August 1983, New York: IEEE Computer Society.
- [3] R. Stallman: GNU Emacs マニュアル, bit 1988. 2 別冊
- [4] 斎藤康己: 拡張可能な画面エディタ EMACS, 情報処理, Vol.25, No.8 (1984)
- [5] 田畑哲彦, 杉山紀子: 汎用キャラクタディスプレイのための画面エディタ, 情報処理, Vol.25, No.8 (1984)
- [6] Lotus Development Corp.: 1-2-3 入門, 12.1987.
- [7] Xerox: NoteCards 1.1 Reference Manual, Mar. 1985.
- [8] D. Goodman: THE COMPLETE HYPERCARD HANDBOOK Vol.1, BNN, 1988. 3
- [9] Sun Microsystems, Inc.: OPENLOOK Graphical User Interface Functional Specification Prerelease Version, July 15, 1988.
- [10] 日本機械工業連合会, 新世代コンピュータ技術開発機構: 昭和59年度先端的コンピュータに関する調査研究報告書〔技術動向編〕-Ⅲ, 1985. 3
- [11] C. Lewis, D. A. Norman: Designing Error, HUMAN-COMPUTER INTERACTION 1987, MORGAN KAUFMANN
- [12] 原田賢一: 構造エディタ, 情報処理, Vol.25, No.8 (1984)
- [13] 辻井潤一, 上原邦昭: ソフトウェア工学と自然言語処理, 情報処理, Vol.28, No.7 (1987)
- [14] 杉山, 秋山, 亀田, 牧之内: 対話型自然言語プログラミングシステムの試作, 電子通信学会論文誌, Vol. J67-D, No.3, pp 294-304 (1987)
- [15] Y. Tanaka, T. Imataki: Intelligent Pad, 情報処理学会アドバンスデータベース・シンポジウム, 1988.12
- [16] 伊藤絃二: 知的 CAI システム探訪, 情報処理, Vol.29, No.11 (1988)

6.2 集団レベルでの問題解決支援

6.2.1 はじめに

集団レベルでの問題解決支援の研究は、計算機支援による集団レベルの協調作業の研究すなわち、CSCW (Computer Supported Cooperative Work) の研究として、近年、アメリカで盛んに行われている。例えば ACM SIGCHI と SIGOIS の後援により 1988 年 9 月アメリカのオレゴン州ポートランドで CSCW '88 が開催された。また、本年 (1989 年) 4 月から 5 月にかけて ACM SIGCHI によって開催される CHI '89「計算システムにおける人間因子に関する会議」でも“Cooperative Work”が取り上げられようとしている。

この CSCW の例として以下のシステムがあろう。

- ・電子メールや電子掲示板システム
- ・コンピュータを活用した face-to-face の会議や遠隔地間の電子会議システム
- ・共同執筆を支援するシステムや設計デザインチームを支援するシステム
- ・組織の意思決定やプロジェクト管理システム
- ・事務所や会社などの OA (Office Automation) システムなど

これらの CSCW システムは個人だけでなく組織の行動にも関係するためコンピュータ科学分野だけでなく、特に、心理学、社会学、組織論、人類学なども関係する学際的研究を必要とする。また、この研究では実際にシステムを構築し、運用を試み、評価するという経験的研究手法を採らなければいけない。その上、組織は複雑な要因を持つため評価の困難さを伴う。

コンピュータ科学における CSCW システムの要素技術として、データベース、Hyper Text など構造化ドキュメント、アクセス制御、個人情報の保護及びネットワーク技術がある。また、この CSCW システムやソフトウェアのことを“groupware”と呼ぶ人もいる。

CSCW システムの実用化のために、人と計算機の 1 対 1 のヒューマンインタフェース研究の他に、集団と計算機のヒューマンインタフェース研究が重要となってきた。来るべき情報化社会のために今後活発に研究を行わなければいけない分野であろう。

そこで本報告では、まず CSCW システムの開発事例を紹介し (6.2.2 節)、次に集団と計算機のヒューマンインタフェースに関する研究論文として、CSCW システム構築に際して問題点を述べた論文と、組織の構造とその集団効果性 (例えば、組織の生産性) に関する研究をサーベイした論文を紹介する (6.2.3 節)。最後に今後の情報化社会に向けての社会変化について述べた論文を紹介する。(6.2.4 節)。

6.2.2 システム事例

本節では集団レベルでの問題解決支援システムの開発事例を 3 件紹介する。

6.2.2.1 会議での問題解決支援システム：黒板を越えて[1]

(1) はじめに

黒板の代わりに、LAN に接続したワークステーションと会議支援のソフトをそろえ、新しい会議室の形態を模索した研究である。これは研究所内での会議の効率化をねらって始められた研究である。

(2) Colab 構成

この計算機支援の会議室は“Colab”と呼ばれており、その様子を図6.2-1に示す。Colabは2人から6人用の会議室であり、各人はLANに接続したワークステーションの前に座る。各ワークステーションのディスプレイには会議参加者総てに共有なウィンドウ（public interactive window）と個人用ウィンドウ（private window）が見えている。この共有ウィンドウは情報交換のために使い、WYSIWIS（What You Is What I see の略、互いに同じものをみることができる環境）の考えを実現している。

日頃我々が使っている黒板に変わるものとして、この Colab 研究の中で以前、Boardnoter を開発した。図6.2-2に Boardnoter のスクリーン表示を示す。図中、一番大きな領域で自由なスケッチができる。そのスケッチの道具はその下の領域に並んでおり、前に描かれたスケッチが一番下に並べられている。

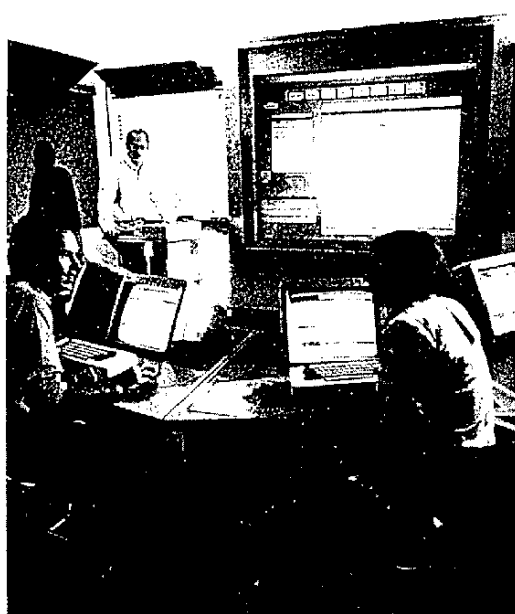


図 6.2-1 Colab の様子

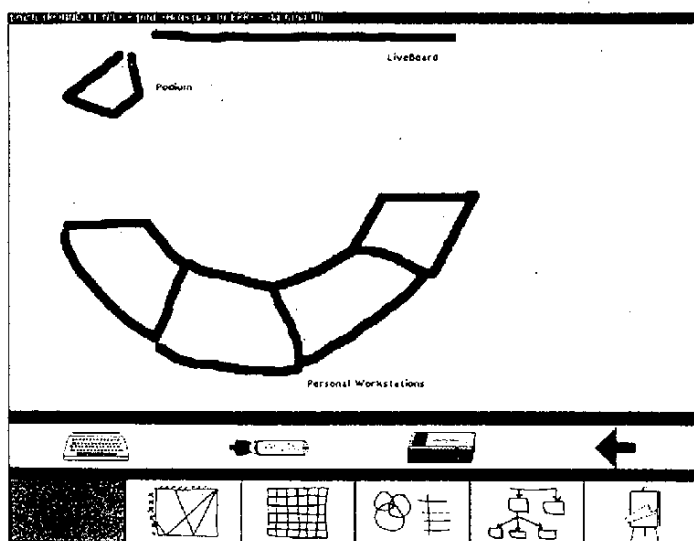


図 6.2-2 Boardnoter の表示

Boardnoter の他に、Colab のツールとして Cognoter と Argonter を開発した。Cognoter はプレゼンテーションのアイデアをまとめるためのツールであり、Argonter は企画提案（proposal）の検討と評価のツールである。

この論文では、特に、この Cognoter と Argonter に焦点を当てて報告している。

(3) Cognoter

Cognoter はプレゼンテーションのアイデアをまとめるためのツールである。Cognoter 上の作業は3つのフェーズすなわち brainstorming, organizing 及び evaluation に分けられる。

(ア) brainstorming

図 6.2-3 に brainstorming のディスプレイ表示を示す。brainstorming のフェーズで会議参加者はアイデアを提出する。ここでは、アイデアの追加移動はできるが、既に出されたアイデアの削除はできない。

(イ) organizing

organizing のフェーズでは、前の brainstorming のフェーズで出たアイデアを link と group の2種の関係付けで系統だてる。link 関係によりアイデアのプレゼンテーションの順番を明確にし、group 関係によりアイデアのグループ化を行う。

図 6.2-4 に link 関係付の例を図 6.2-5 に group 関係付の例を示す。

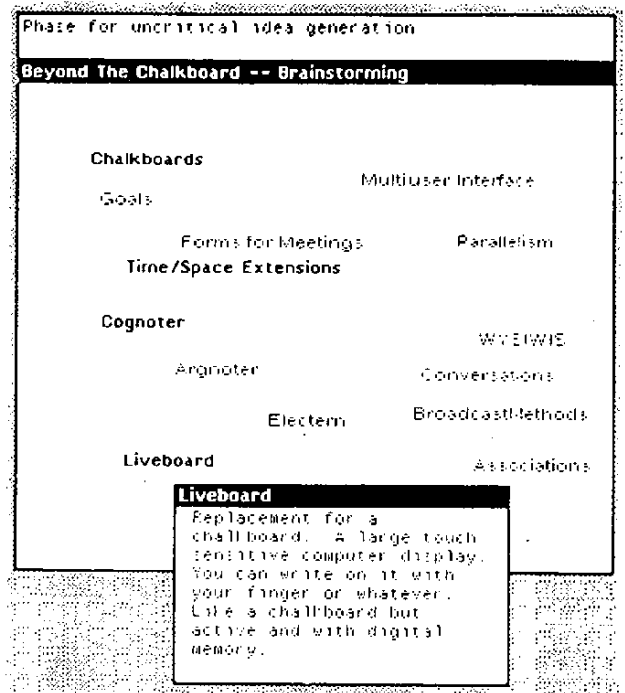


図 6.2-3 Brainstorming の場面

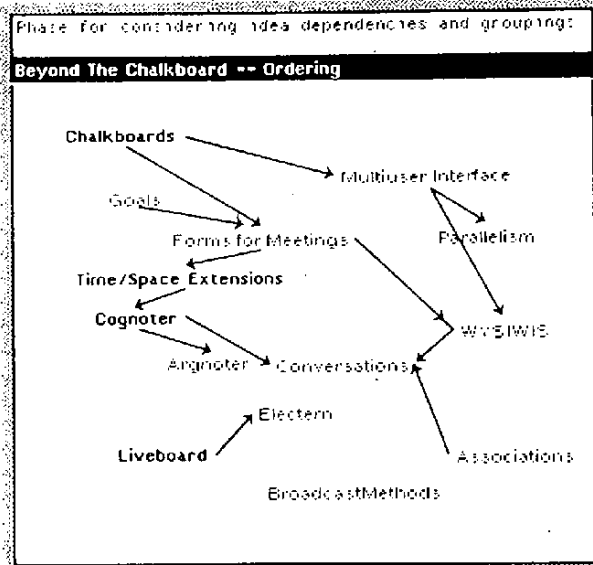


図 6.2-4 Organizing の場面 (link 関係付け)

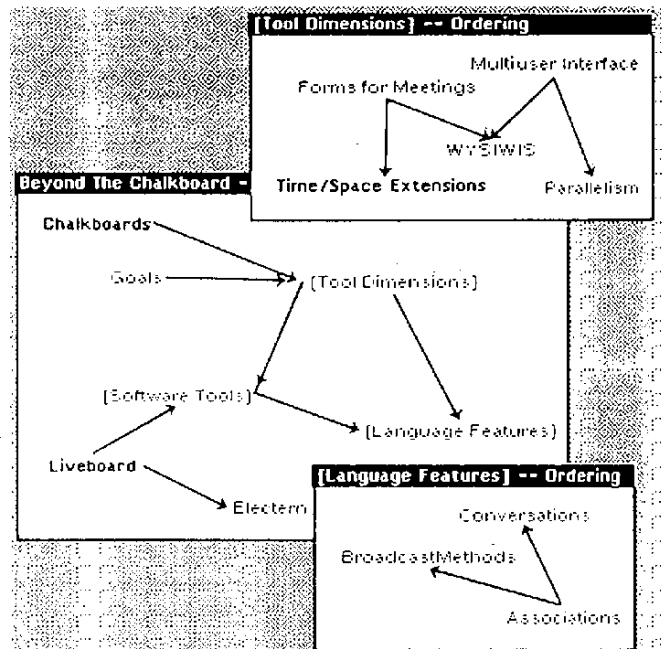


図 6.2-5 Organizing の場面 (group 関係付け)

(ウ) evaluation

evaluation のフェーズでプレゼンテーションの最終形を決定する。ここでは、必要ないアイデアを削除したり、重要なアイデアを詳細化を行う。

(4) Argnoter

Argnoterは企画提案の検討と評価のツールである。参加者は前もって案を作っており、それらを比較検討し、評価することで最もよい案を選択する。Argnoterでの作業は proposing, arguing 及び evaluating の3フェーズにより構成される。

(ア) proposing

参加者は proposing のフェーズで企画提案を会議に提出する。

(イ) arguing

各人は賛成又は反対を明確にし、その理由を記入する。この Argnoterには賛成意見と反対意見の比較表を作る機能があり、これらの意見を比較検討することができる。

ところで、plattによるとこのような会議参加者の多様な提案や意見による議論によってしばしば、大変生産的な意思決定がなされるそうである。

(ウ) evaluating

まず、個々の意見の仮定を検討する。各意見の仮定とは、各人の持つ信念に当たり、これらを明確にすると共に提案、意見および信念の論理的関係をはっきりさせる。

次に、評価基準を選び、提案をランク付ける。各人の評価基準の違いによって衝突がおこった場合にも、Argnoterを使えば、ある評価基準を採用してどうなるか試みることができる。

このように、議論の構造を明確にすることは会議参加者間のコンセンサス形成に役立つ。

(5) まとめ

以上述べたように計算機支援による会議室とツールを作成し実験を行った。それにより以下の結果が得られた。

- ・設計当初予想していた使い方と異なる使い方を会議参加者はする場合がある。例えば、Cognoterでは、グループ関係付けを brainstorming フェーズでの作業として設計した。しかし、Cognoterの organizing フェーズに入る前の brainstorming フェーズで、参加者が空間的にアイテムを配置することでグループ関係付けを始めてしまう場合があった。

- ・共有ウィンドウに同時に書き込めるのは一人だけであり、その書き込み中の別の人のアクティビティをどううまく運用するかが問題となる。
- ・他人が今何をやっており、自分は次に何をすべきなのか分からなくなる状態が参加者の間で観察された。
- ・“Scroll Wars” “Window Wars”と呼んでいる複数の人の表示域の取りあいが起こった。
- ・データの共有ウィンドウへの入力を個人ウィンドウで完全にデータを作成してから行うようにしたら、データ入力を待つ人々はデータ入力者が何をやっているか分からず、欲求不満になってしまった。活動の単位をどれだけにするかが問題となる。

現在、ビデオを設置し、会議参加者の挙動を観察している。今後は、Colabを会議のより深い理解と会議支援のツールの効果を試すための実験室と考え、各種の協調作業の研究を行う予定である。

6.2.2.2 電子会議システム[2]

(1) はじめに

64 Kbpsの通信回線を用いた動画も扱える遠隔地間の電子会議システムを開発した。その機能は、使用記

述などについて紹介する。

(2) 開発方針

以下のシステムを開発した。

- ・経済的なシステムとして、音声会議システムと描画会議システム
- ・導入しやすい動画システムとして 64 Kbps 動画会議システム

ところで、静止画会議システムは単独で開発する利点がないので動画会議システムに静止画機能を盛り込むことにし、あえて単独で開発することをしていない。

(3) システム概要

表 6.2-1 に各システム（モデル）の構成要素と使用回線を示す。

(ア) 音声会議システム

既存の電話回線を利用して音声会議システムを開発した。

(イ) 描画会議システム

描画会議システムでは描画パットあるいは電子白板を使うことができる。なお、このシステムは音声会議システムと同様、電話回線を用いているため経済的なシステムとなっている。

描画パットを用いた場合には片側 2 から 3 名程度の少人数の会議に有効であり、電子白板を用いた場合には片側 5 から 6 名の比較的多人数の会議を行うことができる。

(ウ) 動画会議システム

音声・描画・書画・動画という会議のために必要な要素を総べて扱える電子会議システムである。相手の顔を見ながら話をすることができるため、現在、行われている会議に最も近い形態の会議ができる。

ところで、動画の対象となる人の動きが口や頭の動き程度で、あまり動きがないという特質に注目すれば、画面と画面の差分信号だけを伝送することによって、伝送する動画の情報量を減少させることが可能となる。さらに、色彩と明るさに関する人間の視覚特性に着目して、色彩変化の大きい部分では明るさの変化情報を省略するといった記述を用いることによって、伝送する情報量を減少させることができる。これらの基本的考え方をもとにした「輝度・色差差分ベクトル符号化方式」という高能率符号化方式によって、64 Kbps を用いたビデオ通信装置を開発し、動画会議システムの基幹として使用している。

この装置は、図 6.2-7 に示すように、大別して(2)音声コーデック(1)ドキュメントコーデック(3)カラー静止画コーデック(4)カラー動画コーデック(5)システム及び伝送制御部(6)ディスプレイの 6 つの部から構成される。また、図 6.2-6 にディスプレイ画面構成を示す。

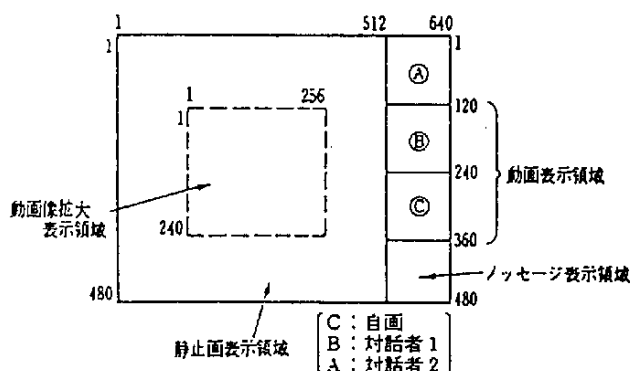


図 6.2-6 画面構成

表 6.2-1 各モデルの構成要素と使用回線

◎ 基本要素
○ オプション

構成要素 モデル名	音 声	描 画 パ ッド	電 子 白 板	書 画	動 画	フ ァ ク シ ミ リ	使 用 回 線
音 声 会 議 シ ス テ ム	◎	—	—	—	—	○	電話回線（1回線）、ファクシミリは別回線
描 画 パ ッ ド 会 議 シ ス テ ム	◎	◎	—	—	—	○	同 上
電 子 白 板 会 議 シ ス テ ム ・ タ イ プ A	◎	—	◎	—	—	○	同 上
電 子 白 板 会 議 シ ス テ ム ・ タ イ プ B	◎	—	◎	—	—	○	電話回線（2回線）、ファクシミリは別回線
動 画 会 議 シ ス テ ム ・ タ イ プ A	◎	○	○	◎	◎	○	64 Kbps デジタル回線
動 画 会 議 シ ス テ ム ・ タ イ プ A	◎	—	○	◎	◎	○	128 ～ 256 Kbps デジタル回線

(4) 導入効果

伝送路が 64 Kbps 回線 1 つで実現できるため以下のメリットがある。

- ・コスト、特にランニングコストの安い動画会議が可能となる。
- ・統合ネットワークシステムの中の 1 サブシステムとして、気軽に動画会議を導入することができる。
- ・将来の複合デジタル交換機への接続も容易である。
- ・タイムリーに会議を開催することができる。

また、経済的で導入し易いシステムであるため、多くの場所にそれぞれ最適な電子会議システムを「会議コーナー的」に設置することができる。

(5) 今後の課題

今後の課題として以下の点が上げられる。

- ・多地点間会議システムの開発
- ・会議予約システムなどの会議支援系システムとのリンク
- ・さらに経済化を追及した動画会議システムの開発

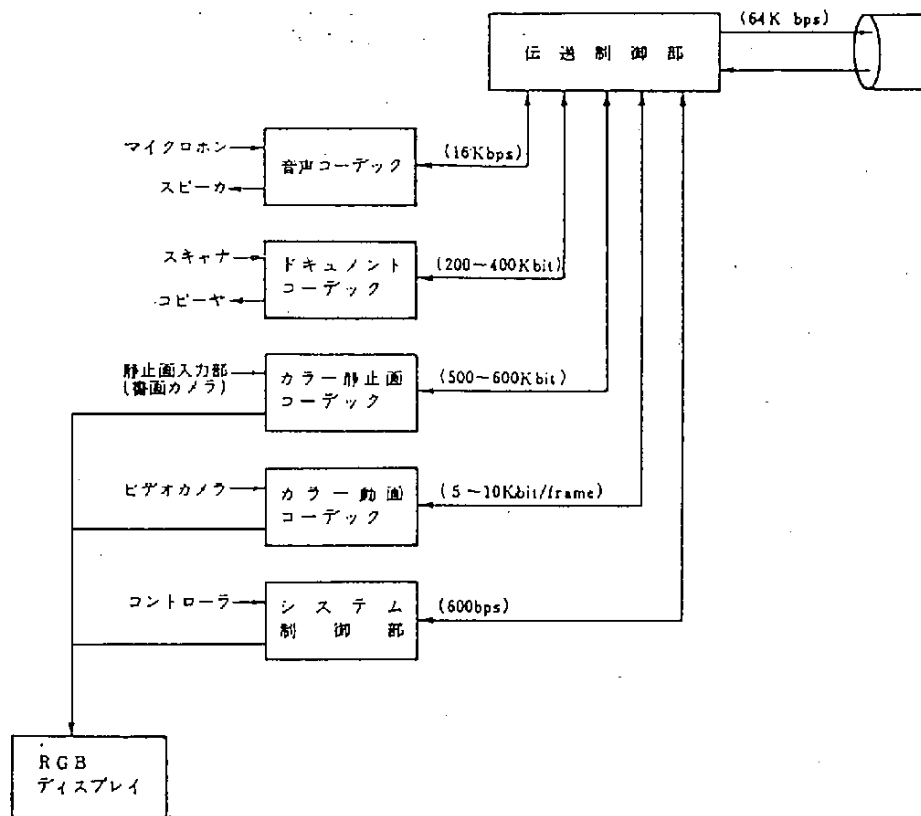


図 6.2-7 ビデオ通信装置の構成

6.2.2.3 「半構造メッセージ」は、連絡調整のために大変役立つ[3]。

(1) はじめに

高機能な電子メールシステムの研究である。この研究では、Semistructured (半構造) メッセージと言うアイデアに従い “Information Lens System” を開発し、その有効性を実証した。

半構造メッセージは定型的なメッセージと非定型的なメッセージの中間的なメッセージである。既存のメールシステムでの宛先フィールドのような定型フィールドの他に特に定められていないフィールドも許すところが特徴である。図 6.2-8、テンプレートを使ってメッセージを作っている例を示す。ここでは、各フィールドに入るものをメニューの中から選んでいる。また、フィールドにすでに内容が記入されているテンプレートもあり、その例を図 6.2-9に示す。

図 6.2-8 メッセージの作成 (その 1)

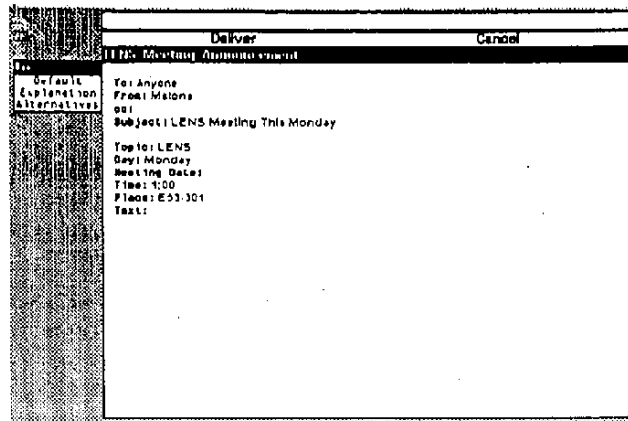


図 6.2-9 メッセージの作成 (その 2)

(2) 半構造メッセージの特長

この半構造メッセージには以下の利点がある。

- ・コンピュータの自動処理の可能性が大きい。

ユーザにフィールド毎に情報を記入させることにより、まだ実用的でない自然言語処理を用いなくてもメッセージに書かれた情報を自動処理できる可能性が大きい。

- ・厳しい制約に捕らわれずに人々はコミュニケーションできる。

これは、非定型な文書もメッセージの中に入れることができることによる。しかし、このためには柔軟なメッセージに対しても正常に動作する信頼性がシステムに要求される。

- ・既に、多くの人が半構造メッセージに慣れている。

既存の組織の中で、人々は自然に半構造メッセージを作り出し、仕事を行っており、導入がしやすい。

- ・メッセージのテンプレートを用意することは、ユーザにとって役に立つ機能である。
- ・システムの拡張が容易である。

(3) Information Lens System

半構造メッセージのアイデアを“Information Lens System”を開発して具現化した。このシステムでは情報の選別機能を実現している。つまり、届いたメッセージにフィルターを掛け、ソートし、優先順位をつけることで、役に立つメッセージを選び出すことができる。

(ア) 機能

情報選別機能のために以下の機能が用意されている。

- ・定型のテンプレートを使ってメッセージを作ることができる。
- ・各個人が記述したルールに従い自動的にメッセージの選別を行う。
- ・“Anyone”メッセージを出すことによって、不特定の人にメッセージを送ることができる。
- ・ルールに記述することによって、“Anyone”に送られたメールの内自分に必要なものを選別することができる。

また、返事作成の支援機能とメッセージの新しい種類を作成するための機能がある。

(イ) ルール

メッセージを作る時と同様なインタフェースでルールテンプレートを編集してルールを作る。その例を図 6.2-10 に示す。ルールには“local rule”と“central rule”の 2 種類がある。

- ・ local ルール

local ルールはメールサーバからユーザのワークステーションにメールが渡される時、起動される。図 6.2-11 の(a) から(b) が local ルールの例であり、各ルールの役割は次の通りである。

- (a) 適当なホルダーにメールを振り分けるルール
- (b) メールの削除のルール
- (c) 自動転送のルール
- (d) メッセージの“characteristics”を用いた複数ステップの処理を行うルール

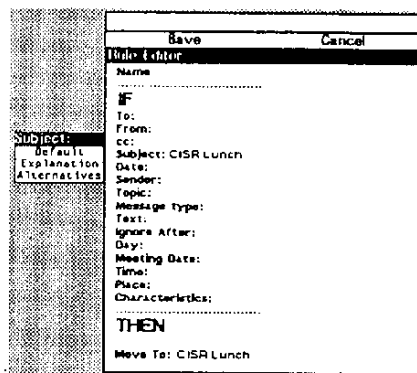


図 6.2-10 ルールの作成

- ・ central ルール

central ルールは“Anyone”に送られてきたメッセージに対するルールである。

図 6.2-10 e がその例である。この例の THEN 部“show”によりユーザにこのメッセージを転送する。central ルールを掛けた後 local ルールを掛ける。

- (イ) 返事作成の援助

既存のメールシステムの中で使われている“answer”や“forward”の一般化として次の機能を実現した。

- ・ 返事の示唆

返事として使えるようなメッセージの種類をメニューとして見せる。受け取ったメッセージとその返事メッセージの対応を図 6.2-12 に示す。

- ・ ユーザの行動の示唆

ユーザの取るべき行動の候補を示し、ユーザが選んだ行動にしたがって適切なアプリケーションを起動する。その上、必要な情報をメッセージから取り出してアプリケーションに引き渡すことができる。

- (ロ) メッセージの種類の作成

既成のメッセージを継承することで新たにメッセージの種類を作り出す。また、メッセージの編集ツールとして視覚的インタフェースを持つ。図 6.2-13 はメッセージテンプレートの継承関係を表示しているところである。

(4) 応用

Information Lens System の応用例を示す。

- (ア) 会議支援

Information Lens System に以下の拡張を行い、会議開催の支援機能を実現した。

- ・ “conference opener” というメッセージの種類を加えた。
- ・ “joint” という response option を加えた。これが選ばれると自動的に以下を行う。
 - i) 新しい、ホルダーを作る。
 - ii) Anyone のルールの中にこのメッセージを選別する機能を加え、
 - iii) このメッセージを新しいホルダーに移動する機能を加える。

- (イ) スケジュール調整

“meeting proposals” “meeting acceptances” というメッセージの種類を加え、会議の時間調整機能を実現した。

```

(a) IF      Message type : Action request
           Action deadline : Today, tomorrow
      THEN  Move to : Urgent
(b) IF      Message type : Meeting announcement
           Day : Not Tuesday
      THEN  Delete
(c) IF      Message type : Meeting proposal
           Sender : Not Axsom
      THEN  Resend : Axsom
(d) IF      From : Silk, Siegel
      THEN  Set Characteristic : VIP
      IF      Message type : Action request
           Characteristics : VIP
      THEN  Move to : Urgent
(e) IF      Message type : Request for information
           Subject : AI, LISP
      THEN  Show
(f) IF      Message type : NYT Article
           Subject : Computer
      THEN  Move to : Computers
      IF      Message type : NYT Article
           Subject : Movies
      THEN  Move to : Movies
      IF      Message type : NYT Article
           Article date : < Today
           Characteristics : Not MOVED
      THEN  Delete
      IF      Message type : NYT Article
           Characteristics : Not MOVED and not DELETED
      THEN  Move to : NYT Article

```

図 6.2-11 ルールの例

Original Message	Suggested Reply Types
Message	M e s s a g e
Action request	Commitment, Request for information, Action request, Message
Notice	Request for information, Action request, Message
Bug fix request	Bug fix Commitment, Bug fix announcement, Request for information
Meeting proposal	Meeting proposal, Meeting acceptance
Meeting announcement	Request for information

図 6.2-12 返信の種類

(5) まとめ

半構造メッセージを使うと以下の利点がある。

- ・通信の情報を形式化できる。
- ・自動的な情報の選別，種別，優先順位付けができる。
- ・ある種のメッセージに対して自動的に反応する。
- ・次に採るべきユーザの行動を示唆することができる。

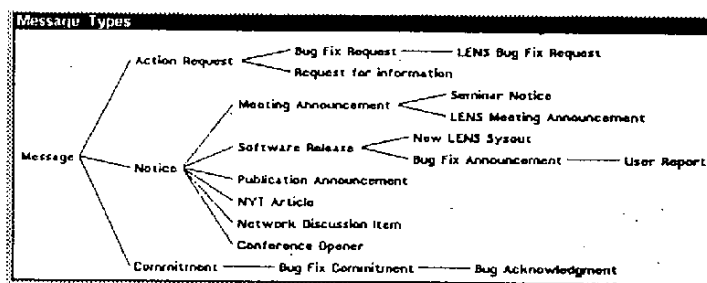


図 6.2-13 メッセージ間の継承関係

ところで，Information Lens Systemのよいところは，簡単にメッセージの種類やユーザを増やすことができることである。メッセージ種類の作成には，基本的には，メッセージの目的で分類し，新しいフィールドが必要になったり，デフォルトが必要になったら，新しいメッセージの種類を増やすのがよい。しかし，どんなメッセージ体系がよいかはその組織やアプリケーションにより，異なるため今後の課題となろう。

6.2.3 ヒューマンインタフェースの問題

本節では，集団レベルでの問題解決支援システムのヒューマンインタフェースに関する論文を紹介する。すなわち，CSCW システム構築に際しての問題点を述べた論文と組織の構造とその集団効果性（例えば，組織の生産性）に関する研究のサーベイ論文を紹介する。

6.2.3.1 なぜ，CSCW システムは失敗するのか[4]

(1) はじめに

本論文では CSCW システムを構築し組織に導入する際の問題点について述べている。

この論文ではまず初めに，会議開催のため，参加メンバーの時間調整を自動的に行う電子カレンダーシステムの以下の問題点を示している。

- ・全ての人が自分のカレンダーを管理しなくてはならない。
- ・個人の「空いてる時間」が本当に空いている時間ではない。
- ・このカレンダーシステムは会議を召集する人（普通は組織の管理者）にとって有り難いが，システムの導入により，カレンダーを管理しなければいけない人々（普通は部下）には余分な仕事が増えることになる。

(2) 設計と評価における問題

CSCW システムの設計と評価に関して以下の問題がある。

(ア) 苦勞する人と樂をする人の違い

電子メールシステムのように総べての人が平等に苦樂を共にするシステムがよい。しかし，グループ全体をサポートする電子カレンダーのようなシステムではグループ全員が使わなければならない。一人でも使わない人がいるとそのシステムは有効に機能しないからである。そのためには，全員が使うように仕事や組織を変更するのであろうか。しかし，それには構成員の再教育など多くの勞力が必要である。従って，システムが組織に適合するべきであり，総べての人がシステムを使うことで利益を得ることが必要である。

(i) 意思決定者（管理者）的発想の打破

CSCW システムを導入しようとする組織の管理者は単一ユーザを対象として開発したシステム、例えばワープロに対する発想で考えてしまう。「マルチユーザアプリケーション」についての理解がないのである。すなわち、管理者は本人と同類の人の便利さに関心が行くが、その他の余分な仕事をしなければいけない人々のことをきちんと考えないところがある。つまり、管理者は管理者の利益に敏感だが、他の種類の人々の負担には鈍感である。

(ii) CSCW システムに対する評価の困難さ

CSCW システムに対する評価は単一ユーザを対象としたアプリケーションに対する評価より困難である。それはグループメンバーの動機、経済及び政治的背景が互いに影響するからである。社会の情報化が進むに従い、評価のために人間工学や認知心理学などの分野で研究を行うようになってきた。しかし、現実のCSCW システムの評価には多くの金、時、人が必要となり難しい。

(3) まとめ

グループや組織がいかに機能し発展するかを研究しなければならない。同時に、開発する新しいシステムに対して各個人がどのように反応するかを知らなくてはならない。

また、CSCW システム導入にともない組織の構造を再編成してもよい時を以下に示す。

- ・導入するシステムがより group-friendly である時
- ・システムを導入することでより平等的な職場に変化していく時
- ・システム導入による全体の十分な利益が明らかな時
- ・新しいシステム導入に際しての組織成員間の不安を採り除けた時

6.2.3.2 組織に関する研究上の視点と人間工学[5]

(1) はじめに

この論文は集団の構造とその効果性（例えば組織の生産性）に関する研究の紹介論文である。歴史的な変遷に従い心理学的視点、組織論的視点、生物学的視点に分けてサーベイしている。

(2) 心理学的視点

心理学的研究は1950年代から1960年代にかけ、コミュニケーションネットワーク研究として行われた。しかし、1960年頃から、急速に衰退した。その理由は、研究の方法としてネットワークの型を刺激 (stimulus)、集団効果性を反応 (response) とする単純なS-R理論であったことによる。

(3) 組織論的視点

組織を開放システム（環境と物質、エネルギー、情報を交換するシステム）とみなして研究したコンティンジェンシー (contingency) 研究とネットワーク分析がある。

(ア) コンティンジェンシーの研究

組織を構成する諸要因の間の関係を明らかにする研究であり、Fiedlerのリーダーシップ効果に関する研究がある。

それによると、リーダーに有利な状況と不利な状況では課題動機型のリーダーが高い生産性を上げ、リーダーにとって中間的な状況では関係動機型のリーダーが高い生産性を上げていることが示された。ここで、課題動機型のリーダーとは、課題遂行に高く動機づけられたリーダーであり、関係動機型のリーダーとは他者との良好な関

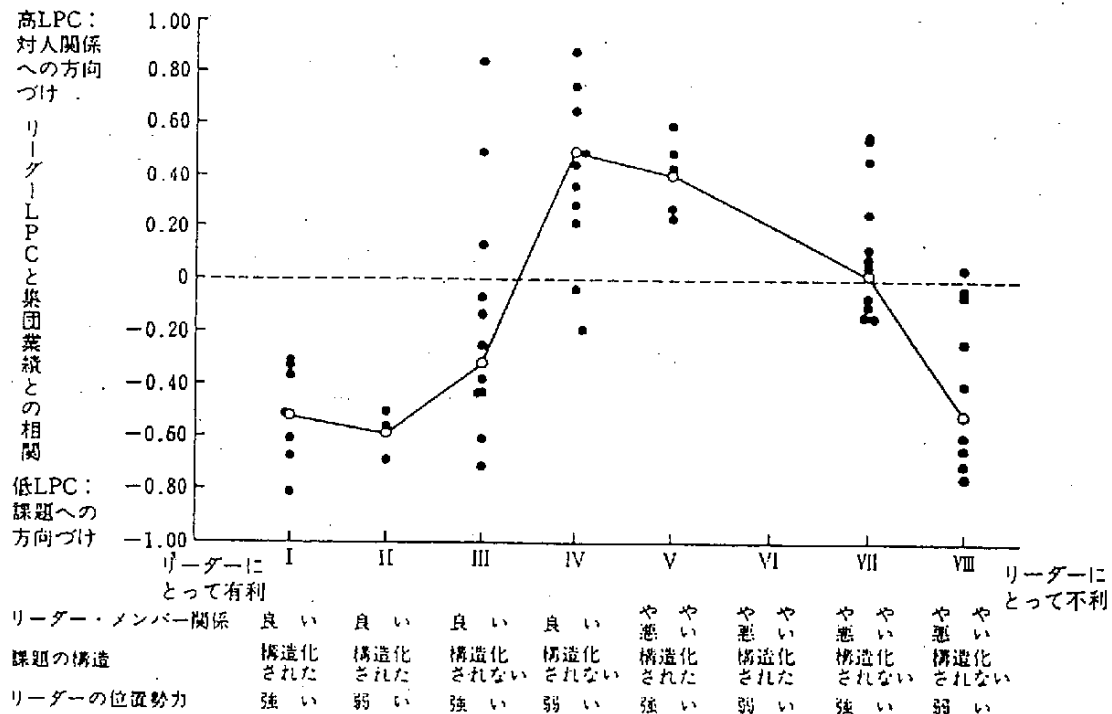


図 6.2-14 各状況ごとのリーダー LPC と集団業績との相関

係を望むリーダーである。(図 6.2-14 参照)

なお、図中 LPC (Least Preferred Coworker) 得点 (“これまで知り合った人の中で一緒に仕事をするのが最も難しかった人”を想定して16個の形容詞上で評定させた結果)の低いリーダーとは、最も働き難い人に対して好意的に評定した人であり、他者との良好な関係を望む(関係動機型)リーダーであるのに対し、LPC得点の低いリーダーとは逆に課題遂行に高く動機づけられた(課題動機型)リーダーである。

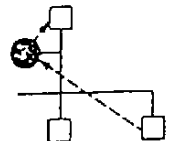
(イ) ネットワーク分析

ネットワーク分析は組織内のコミュニケーション構造を明らかにする研究である。この研究では組織の成員間の好悪の選択のデータばかりでなく、例えば、「仕事上の話を誰にしますか」といったコミュニケーションの流れを示すデータも集められている。データは個人を点に、関係を線(方向が決まっていれば矢印つき)で表した図に変換される。この図から組織成員の個人的役割特性(門番、連結者、意見先導者、世界主義者と呼ばれるもの。図 6.2-15)と、集団の全体的特性(連結度、開放度、支配度など)が分析される。

また、ネットワーク分析によって得られた構造と 図 6.2-15 組織内の個人的コミュニケーション役割

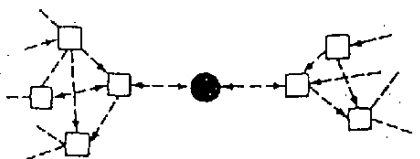
門番

コミュニケーション経路を流れるメッセージを統制するためにコミュニケーション構造のなかに位置する個人



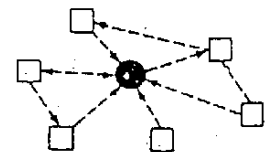
連結者

自分自身はどの仲間集団にも参加しないで2つのシステム内の2つ以上の仲間集団を結びつける個人



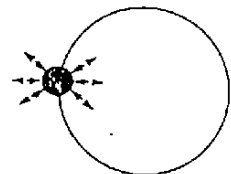
意見先導者

他人の態度や公然の行動をインフォーマルにかつ比較的頻繁に左右する影響力のある個人



世界主義者

システムの環境とコミュニケーションする場合の比較的高い個人



フォーマルな組織構造を照らしあわせ、その対応が調べられる。もし、両者のいずれが組織運営上の不都合を生じさせていると判断された場合、組織の変更を行ったり、作業にはネットワーク分析で見出された役割に合致した職務が与えられるように配置転換を行うこともできる。

(4) 生物学的視点

生物が持つ機能的特徴で、社会組織にもあてはまるものには以下の4点がある（吉村，1986；岩瀬1986より）。

- ・全体の統合的機能性：

細胞－組織－器官－器官系（消化，呼吸，内分泌，循環，神経，生殖，感覚，運動の各系）というレベルは異なっているが、そのそれぞれが調整しあい、統合的に働くこと

- ・合目的性：機能にムダがなく、目的に合致して働くこと

- ・自律性：他動的にではなく、自らが独立に活動し続けること

- ・適合性：外部環境が変動した場合でも複雑な適応活動により活動が維持できること

また、最近、自己組織化（self-organization）の機能も注目を集めている。Guetzkow and Simon (1955) は、各種コミュニケーションネットワークにおける作業手続きと効果性の関係を時間を追って調べた。その結果、不適切なネットワークが与えられると集団の持つ自己組織化の力は抑制されることが明らかとなった。

一方、浅井（1985，1987）は、緊急条件とそうでない条件下で集団に作業を課した場合の集団組織化を観察した。それによると緊急条件で作業を行った集団はそうでない集団にくらべ、求心的な集団構造を形成することによって、作業遂行を早めていたことがわかった。

(5) 今後

今後、人間工学でも生きている人間と生きている組織をさらに追及する必要がある。

6.2.4 今後の予感

集団レベルでの問題解決支援システムの研究開発は、今後の社会の情報化を反映して、盛んに行われるであろう。この情報化は社会構造も変化させる可能性があり、ヒューマンインタフェースもこの動向を踏まえて検討しなくてはならない。そこで本節では、今後の情報化による社会変化についての考察を述べた論文を一つ紹介する。

6.2.4.1 グループ通信：社会システムへの影響と課題[6]

(1) グループ通信の概念

マスメディア（新聞，テレビ），パーソナルメディア（電話，手紙）の他に新たに、グループ通信（電子メールや電子掲示板）が可能になってきた。それに伴い、今後、各種のグループ通信の機会のためのインフラストラクチャを提供するサービスや、その質の高度化をはかるサービス事態が重要なビジネスとなるであろう。

(2) グループ通信の普及と基礎的通信技法の錬磨

今後、書道塾，音楽教室，ライティング・スクール及び絵画学校が隆盛になるであろう。その理由は音声，文学，画像などさまざまな信号・記号がグループ通信で扱えるようになり、各個人にとって人間コミュニ

ニケーション能力の開発と錬磨が不可欠となるからである。

(3) グループ通信に並行して生じる社会変化

(ア) 分衆化

画一性を特徴とする大衆は、差異性を軸にうごめく細分された人々へと分化している。つまり、大衆化に対する分衆化が進んでいる。

(イ) ネットワーキング

1960年代の後半以降、アメリカで次第に拡がってきているネットワーキング運動は、当初は、パソコン通信ネットワークのようなグループ通信メディアとは何の関係も持っていなかった。しかし、パソコン通信の普及と共に両者の融合は急速に進んでいる。

(ウ) 情報権とネットワーク

今まで主権から派生した権利の一つである行為権とりわけ所有権ないし使用权の概念に支えられた個人や集団（企業）が広く出現し繁栄した。彼らは、直接には工業製品（商品）の大量生産・販売を、そして究極的には富み（取り引き力）の蓄積をめざす。産業及び商業という競争ゲーム（富のゲーム）のプレイヤーとして活躍した。このゲームの場となったのが市場であった。

そして、まさに今始まろうとしている近代社会の発展段階を形作るのが、これまた主権の派生的権利の一つである主体の「情報権」の概念に支えられた「ネットワーク」と呼ぶことができる集団であって、彼らは、直接にはさまざまな情報・知識（科学・技術知識だけでなく、宗教、芸術スポーツなどの知識を含む）の獲得を、そして究極的にはより高次の体系的な知（説得力）の獲得をめざす。研究及び普及の競争ゲーム（知のゲーム）のプレイヤーとなって活躍する。その場となる社会システムは、メタネットワークと呼ぶのが適切であろう。

(エ) 社会ゲームとしての知のゲーム

知のゲームの主たるプレイヤーは、ネットワークである。（或いは、知のゲームのプレイヤーとなっているネットワークのことを、特にスクールと呼ぶのも面白いかもしれない）各ネットワークの目標は、他の主体（とりわけ他のネットワーク）に対する「説得力の入手である。他の主体に対する説得力は、他の主体よりも優れた高度で体系的な知の保有によって、そして他の主体によってその事実が承認され評価されること（つまり、他の主体を心服させること）によって可能となる。

(4) まとめ

パソコン通信に代表されるグループ通信メディアは、知のゲームの普及を促進する強力な手段としての役割を果たすとともに、知のゲームの普及にともなって、それ自身もまた、さらなる発展の契機を与えられるであろう。

6.2.5 おわりに

本報告では、はじめに集団レベルでの問題解決支援システムの研究として主にアメリカで行われているCSCW研究を紹介し、それに続き、システム事例、ヒューマンインタフェース研究及び情報化社会の行方に関する研究論文を紹介した。

CSCW研究の適切な参考図書として、Computer-Supported Cooperative Work[7]を推薦する。この文献には以下の論文が掲載されている。

- ・初期の著名な研究
- ・CSCW システム開発報告
- ・CSCW システム設計のための理論研究

参考文献

- [1] Mark Stefik, Gregg Foster, Daniel G. Bobrow, Kenneth Kahn, Stan Lanning, and Lucy Suchman : Beyond the chalkboard : Computer support for collaboration and problem solving in meetings, Communication of the ACM Vol.30 No.1 pp.32-47 (Jan 1987)
- [2] 植田憲一, 水田 博 : 電子会議システム, 沖電気研究開発 第130号 Vol.53 No.2 pp.35-42 (昭和61年4月)
- [3] Thomas W. Malone, Kenneth R. Grant, Kum-Yew Lai, Ramana Rao and David Rosenblitt : Semistructured Messages Surprisingly Useful Computer-Supported Coordination, Computer-Supported Cooperative Work, Edited by Irene Greif, San Mateo, pp.311-331, California, Morgan Kaufmann Publishers Inc. (1988)
- [4] Jonathan Grudin : Why CSCW Applications Fail : Problems in the design and evaluation of organizational interfaces, Proc Conf CSCW'88 ACM, pp.85-93, (Sep 1988)
- [5] 浅井輝昭, 杉山貞夫 : 組織に関する研究上の視点と人間工学, 人間工学 Vol.24 No.1 pp.33-41, (1988)
- [6] 公文俊平 : グループ通信 : 社会システムへの影響と課題, 情報処理 Vol.28 No.8 pp.1044-1051, (Aug 1987)
- [7] Irene Greif : Computer-Supported Cooperative Work : A Book of Reading, San Mateo, California, Morgan Kaufmann Publishers Inc. (1988)

7. ヒューマンインタフェースにおける関連研究

7.1 音声入出力技術の動向

7.1.1 はじめに

この章の目的は、ヒューマンインタフェースに用いることのできる自然な入出力手段としての音声入力と音声合成技術について、アップトゥーデイトな情報を提供することにある。

音声合成技術は、聴取可能な合成音声の発声という意味においては、ほぼ完成に近づいている。8合目か9合目というところである。残された問題としては、音質の自然さ、韻律の自然さにかけることである。これは案外難問で、音響学的なアプローチが必要だが、どう合成した音声がどんな音質に聞こえるかは、未だに定量的には不明である。合成のために必要なテキストの構文解析は、通常意味まで踏み込まずに行うため、比較的容易である。ただし、韻律の自然さの付与は、構文解析の今一つの発展が望まれている。どのような音声が聞きやすいか（ヒューマンフレンドリーか）の研究は、あまり行われていないようである。

音声認識技術は、いまだ1合目あたりのようである。特定話者の単語認識をパターンマッチングで行うかぎりでは、かなり完成されているが、それでは音声認識として十分ではない。不特定話者の連続音声認識が本当の目的である。

現在の音声認識の研究の大きな流れとして、

- (1) Hidden Markov Modelに基づく確率的方法、
- (2) Neural Net を用いた方法、
- (3) 知識表現を用いた AI 的手法、
- (4) 音響的特徴の解析を進める方法、

が挙げられる。ただ今現在の日本、アメリカでは(1)と(2)が圧倒的に研究されており、特に(1)の Hidden Markov Model では、Carnegie-Melon 大学の Kai-Fu Lee が不特定話者連続音声認識システム (SPHINX) を構成したことが話題となっている。最終的なブレークスルーとして求められるものは、音響的特徴に関する新たな知見とされているが、一方現実的なシステムに求められているものは、そのような知識の欠如に対してローバストなモデルである。

本章では、まず音声認識の基本技術である音響分析の一般的方法と認識の代表的な方法について簡単に紹介したうえで、現在開発中の CMU と ATR の音声理解システムと SPHINX システムについて紹介する。音声合成については、DECTalk と NTT の日本語テキスト音声合成システムを紹介する。

7.1.2 音声認識の基本技術 [1],[2]

音声技術認識の一般的な方法 (図 7.1-1)

一般的には、マイクで受けた音声波形から特徴となる情報 (特徴パラメータ) を抽出し、それとあらかじめメモリに格納しておいた標準音声の特徴パラメータ (辞書) と比較し、どの音声に最も近いかを調べる (パターンマッチング) という方法がとられる。使用する前に話者の単語音声をコンピュータに記憶させておき、限定された単語だけを認識させるというのが、特定話者・限定単語方式である。現在市販されている

装置の多くはこの方法で、誰の音声でも認識できるようにしたのが不特定話者・限定単語方式である。

単語や連続的に発生した文章を、単音節 (CV)、音韻 (C, V) や母音-子音-母音 (VCV) 単位で認識させる研究、連続音声方式も進められている。しかし、同じ音韻でも前後の音声により著しく変形 (調音結合) してしまい、この調音結合をいかに処理するかが問題となっている。

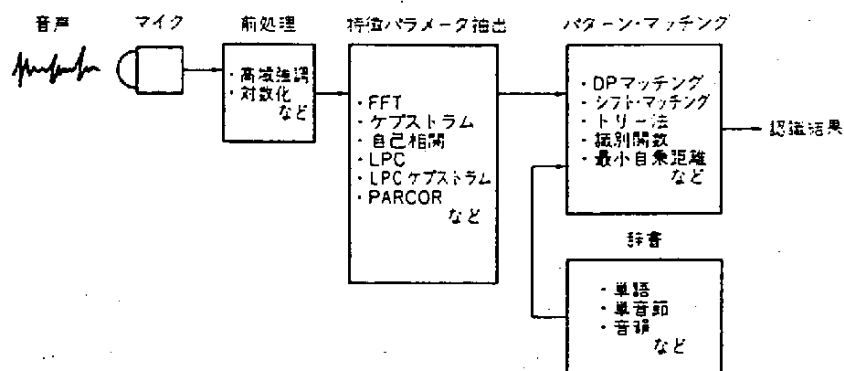


図 7.1-1 音声認識の一般的な構成 [1]

7.1.2.1 音響分析

(1) 音声の特徴パラメータ抽出

特徴パラメータ抽出法は、1)フーリエ変換を基礎とする周波数領域におけるパラメータ、および、2)自己相関を基礎とする時間領域におけるパラメータに大きく分かれる。ただし、一般に、2)の場合にはパワー・スペクトルを求めてからケプストラムに変換するという方法が、また、2)の場合には相関係数から線形予測係数 (LPC) を算出し、さらに、LPC からスペクトル包絡を求める方法がとられる。このように、いずれも最終的にはケプストラムのようなスペクトル包絡パターンに変換する場合が多い。また、これらのパラメータの間には図 7.1.2.1-1 に示したような関係がある。もちろんバンド・パス・フィルタ群を利用してハードウェアでパラメータ抽出を行う方法もあり、高速処理が可能である。

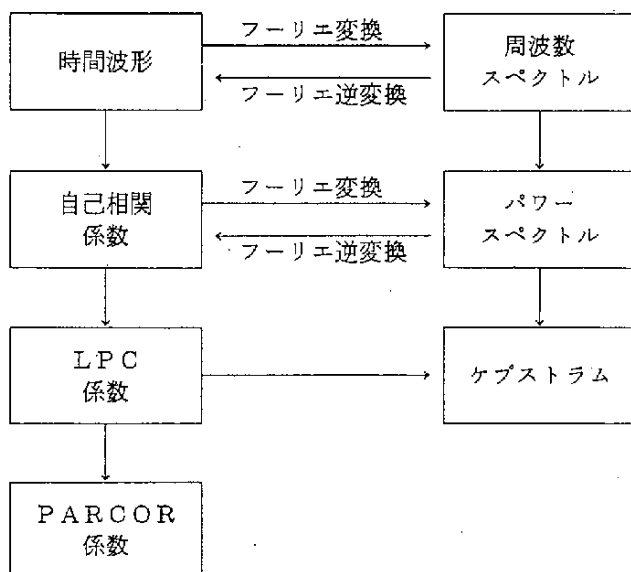


図 7.1-2 各種特徴パラメータ間の関係 [1]

(2) ケプストラム

図 7.1-3(a)に示すように、有声音のパワースペクトルには、声帯振動によるギザギザしたパターンが現れる。ギザギザしたパターンからスペクトル包絡を抽出する方法もいろいろ提案されているが、その代表的なものにケプストラム処理がある。

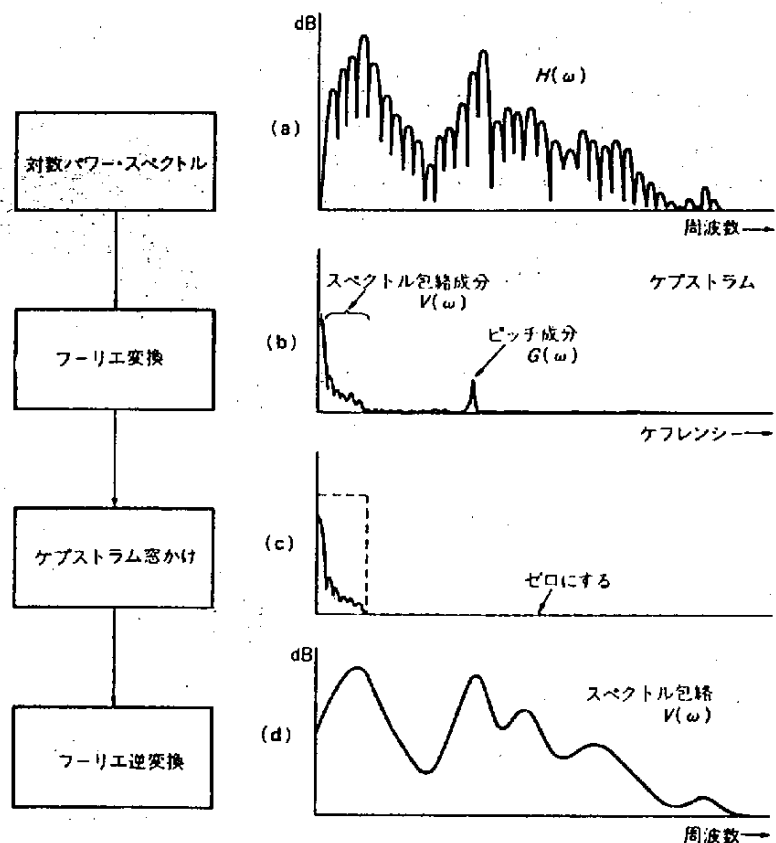


図 7.1-3 ケプストラムの求め方 [1]

(3) 短時間スペクトル (図 7.1-4)

子音のようにスペクトルが時間とともに激しく変化している場合には、波形を時間軸方向に細かく区切って、その1区切り内のスペクトルを次々に求めるという短時間スペクトル分析をしなければならない。短時間スペクトルを求めるときに問題になるのは、1区間をあまり短くすると周波数分解能が悪くなり、逆に、あまり長くすると時間分解能が悪くなるということである。

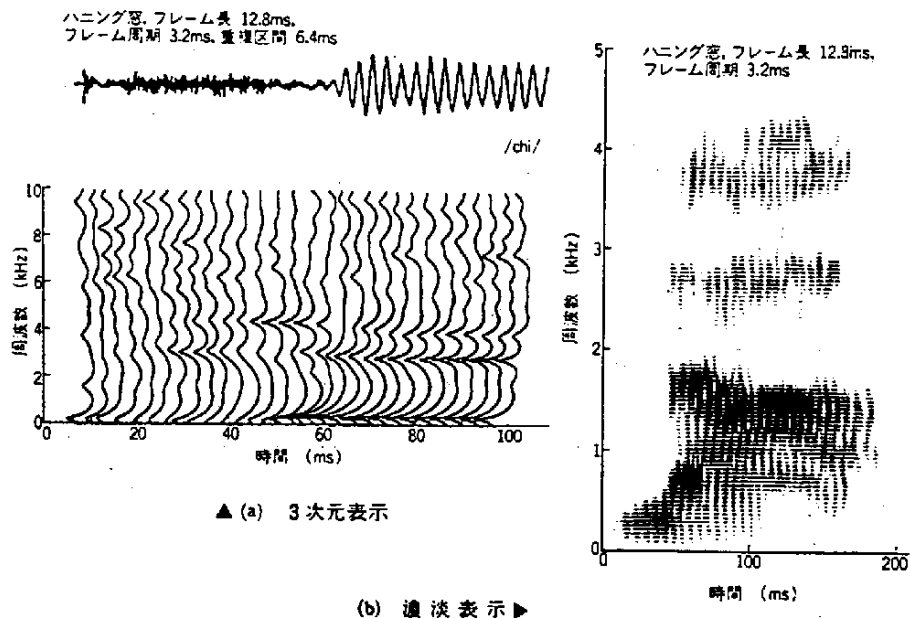


図 7.1-4 単音節音声の短時間スペクトルの例 [1]

(4) 線形予測法

線形予測法をデジタル・フィルタで実現すると、図 7.1-5 のようになり、自己回帰過程というモデルで表される。このシステムの伝達関数は、

$$H(z) = 1 / (1 + a_1 z^{-1} + \dots + a_p z^{-p})$$

になり、極のみ有することから、全極モデルという。ただし、 z^{-1} は 1 サンプル時間の遅れ要素に相当し、上式は Z 変換で表現した伝達関数である。

サンプリングの最高周波数を $f_{\max}$ とすれば $z^{-1} = \exp(-i\pi f/f_{\max})$ となる。

これを上の式に代入して、 $|H|^2$ を求めると次式になり、

$$|H(f)|^2 = 1 / |1 + \sum_{k=1}^p a_k \exp(-i\pi k f/f_{\max})|^2$$

任意の周波数 f に対するパワー・スペクトルが得られることになる。

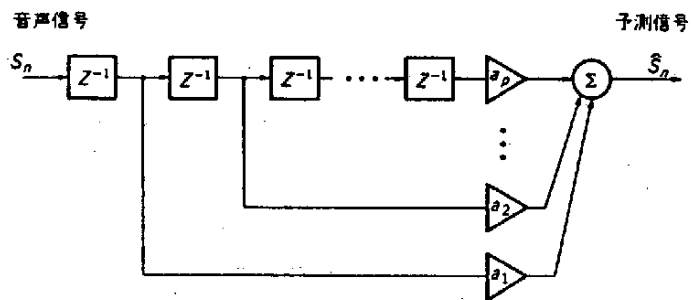


図 7.1-5 線形予測法のブロック図 [1]

7.1.2.2 認識

(1) テンプレートマッチング [1],[2],[3]

パターンマッチング部では、入力音声の特徴パラメータと標準音声の特徴パラメータとの距離を計算し、その距離が最小となる標準音声を探すという方法（テンプレートマッチングの代表）が最も多く、その際重要なことは、入力音声と標準音声の時系列特徴パターンをどのように対応させてマッチングを行うかという時間軸整合の問題である。

DP マッチング

DP（ダイナミック・プログラミング）法が代表的手法で、DP マッチングは、標準パターンと入力パター

ンの距離を計算する際に、両者の時系列情報を1対1に対応づけることなく、2つのパターン間の距離が最も小さくなるように入力側を部分的にずらしながら対応づける方式である。(図7.1-6)

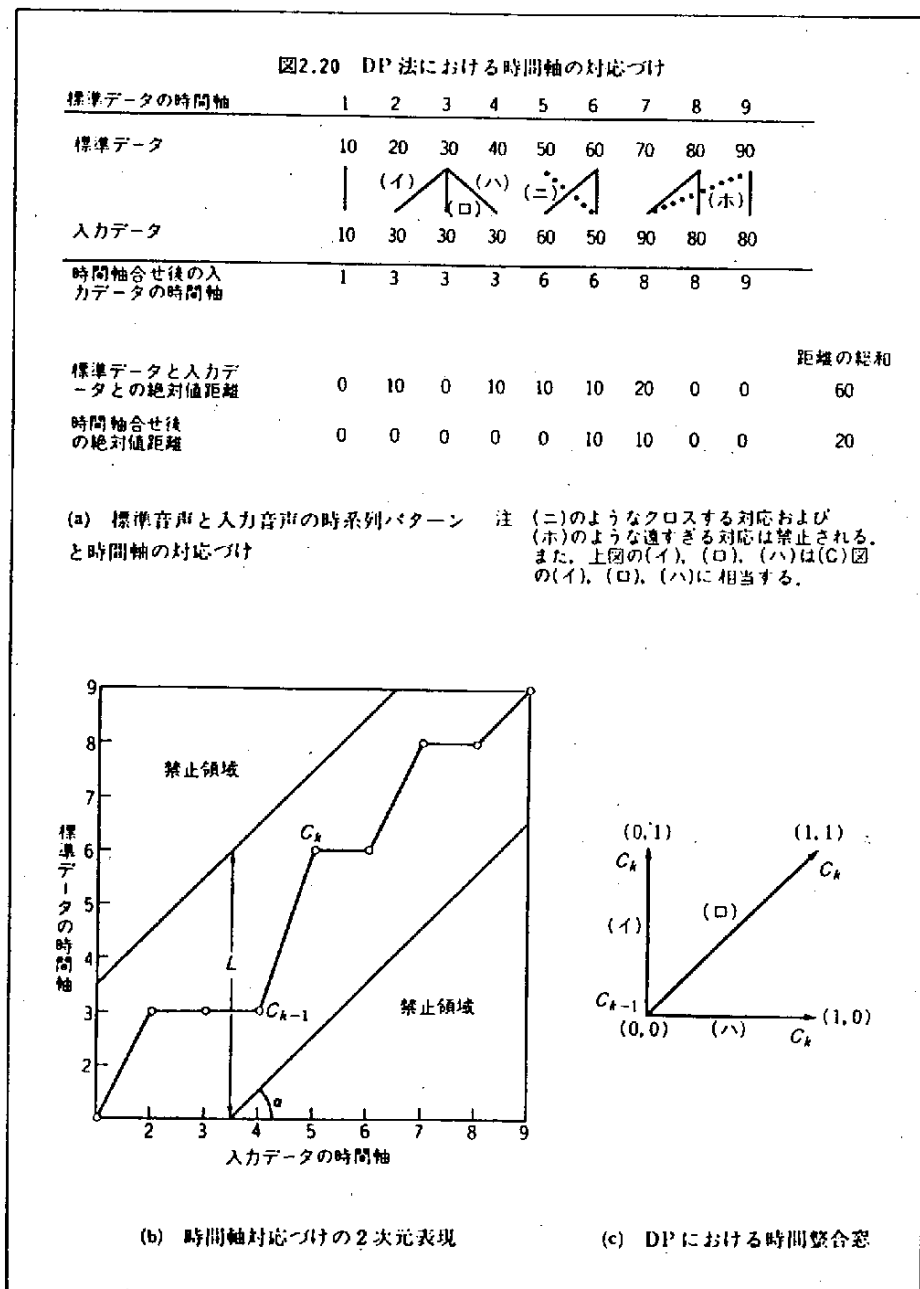


図7.1-6 DP法における時間軸の対応づけ [1]

(2) AI手法 [4],[5]

連続音声は自然言語との関係が深いことから、人工知能技術の導入が試みられており、特にエキスパートシステム研究の発展に伴い音声研究との接近が進行し、音韻認識レベルにおける知識工学的手法の適用が試みられている。

注目されているのは、スペクトログラムに代表される特徴パラメータの図形表示の読み取り知識である。

現在開発中の連続音声認識エキスパートシステムはすべて、グラフィック表示された特徴パラメータの変化の様子を読み取る知識をプロダクションシステムを用いて表現することにより実現されている。

(3) ニューラルネット [6]

神経回路網を、音声に内在する構造を解明するために用いた例としては、USCD の Elman と Zipser による研究が有名である。有声破裂音の単音節を対象とした実験では、6 音声周波数と時間を軸とする二次元パターンとして表現し、3 層構造のネットワークに /ba//da//ga//bi//di//gi//bu//du//gu/ の 9 種類の音節に識別を学習させている。学習の完了したネットワークを用いて、さまざまな入力に対応する隠れユニットの応答を調べたところ、教師信号に母音、子音の概念に基づく区別を含ませなかった場合にも、母音の識別を行うユニットと子音の識別を行うユニットが形成されていることが見いだされている。これは、母音と子音の区別が恣意的なものではなく、物理的にも実体を伴った自然なカテゴリを分けてあることを示唆するものと解釈できよう。

入野らは、神経回路網の特徴抽出能力を用いて話者に依存しない音声の特徴量を探ることを目的に、3 層構造のネットワークに BP を適用して不特定話者の母音の識別実験をおこなっている。音波形の自己相関関数を入力として用いた実験では、母音の識別には 4 kHz 以下の低域のホルマント構造が重要であるという、従来の知見に対応した回路網が形成されることを見いだしており、また、基底膜の振動パターンを入力して用いた実験では、基底膜上での位置を入力ユニットに対応させた場合と空間周波数成分に対応させた場合とを比較して、振動パターンの大域的形状が母音識別に重要な役割を果たしていることを示唆している。

神経回路網の音声認識への適用例は急速に増加しつつあるが、ほとんどは /a//i//u//e//o/ や /b//d//g/ の識別など、問題の一部分に試験的に用いているにとどまり、大規模に適用した例は限られている。このような状況が生じている理由の一つは、入出力ベクトルの次元数を増すと、急速に計算量が増加するという規模依存性の悪さにある。音声認識への大規模な適用が困難であるもう一つの理由は、時間的特徴の表現方法にある。

実用的な規模の音声認識システムを神経回路網を用いて実現するためには、全体を適切なサブシステムに分解する方法を明らかにすることが課題となろう。

(4) Hidden Markov Model [3],[7],[8]

最近、HMM (Hidden Markov Model) による統計的手法が、

- ・実際の音声の揺らぎをより正確に反映し易い
- ・統計理論や情報理論等による理論的展開がし易い
- ・確率の概念を用いて言語処理等を統合し易い
- ・DP マッチングを特殊な場合として含み拡張性がある

などの特徴から着目され、盛んに研究されている。

HMM の場合は、通常、各特徴ベクトルをベクトル量子化したラベルの時系列として音声表現する。訓練時には、入力音声のラベルの系列を出力する確率（尤度）が最大になる HMM を探す。（図 7.1.2.2-2）

HMM は、 N 個の状態 $S_1, S_2, \dots, S_N$ を持ち、初期状態がこれらの状態上に確率的に分布していて、一定周期ごとに、ある確率（遷移確率）で状態を次々にぎ遷移すると共に、その遷移の際に、ある確率（出力確率）でラベルを一つずつ出力するというマルコフ・モデルである。

出力ラベル系列は観測できるが、状態そのものは直接観測できないという意味で Hidden Markov Model (HMM) と呼ばれる。

音声の HMM としては、一般に状態 S_i から前に通過した状態 S_{i-1} に戻るような遷移は時間前後関係を乱すので許さない。このような HMM の構造としては、図 7.2.-7 のようなものがある。

HMM に関連して重要な基本問題として次の三つがあり、それぞれにアルゴリズムが開発されている。

- 1) モデルの評価：モデルMがラベル系列を出力する確率（尤度） $\Pr(O|M)$ をもとめる。
- 2) モデルの訓練：ラベル系列Oが与えられて、 $\Pr(O|M)$ を最大にするようにモデルMのパラメータを推定する。
- 3) 状態系列推定：モデルMがラベル系列Oを出力するときの最も可能性の高い状態系列を推定する。

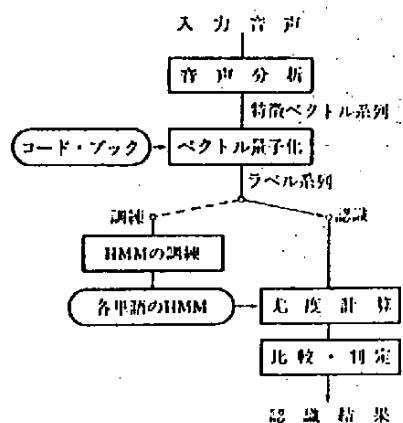


図 7.1-7 HMM による単語音声認識
の基本構成 [7]

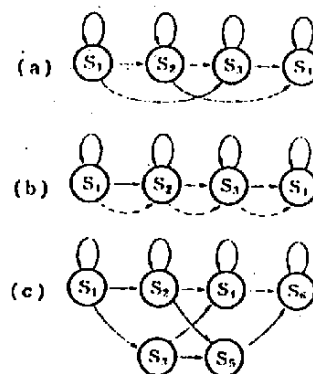


図 7.1-8 状態が半順序になっている
HMM の例 [7]

7.1.3 音声理解システム [9],[10]

人間は、頭の中で考えた事柄を感じたことを、概念から単語列に変換し、更に、発声器官を駆動して声にだして、他人に伝える。聞き手は、会話状況、文脈、意味、構文、単語等の言語情報を駆使して、話し手の音声を理解し、意図を把握している。

このように、音声認識の過程を、人間の言語活動の一環として総合的に捕らえようという研究が音声理解システムである。

1985 年から、米国では、DARPA の Strategic Computing Project の一つとして、CMU, BBN (Bolt Beranek and Newman Inc.), MIT (Massachusetts Institute of Technology) を中心に第 2 次の音声理解系の研究が開始されている。

ヨーロッパでも、英国のアルペイ計画、EC 諸国のエスプリ計画等の新世代計算機プロジェクトが活発に行われており、我が国では、国際電気通信基礎技術研究所 (ATR International) の ATR 自動翻訳電話研究所で音声理解を含む自動翻訳電話のプロジェクトが開始された。

7.1.3.1 CMU と ATR のシステム

(1) CMU のシステム

1985 年 1 月から DARPA の戦略計算プロジェクトの一環として、CMU を中心として、4 年間の「不特定話者大語彙連続音声認識」プロジェクトが開始された、1 千万ドル/年の全米的な大プロジェクトである。目標は、不特定話者の 10,000 単語からなる連続音声の認識で、システムを実時間に近い速さで動作させることを目指し、様々な知識源を分散統合して、スーパーコンピュータも含む並列分散システム上で実現しようとするものである。

システム構成は、図 7.1-9 に示すように、知識源 (Knowledge source) と参照可能なメモリ (shared

memory) としてのブラックボード (blackboard) からなり、知識源は、並列に独立に動作するように考えられており、音韻認識部、単語認識部、言語処理部、分散システム用ツールからなる。

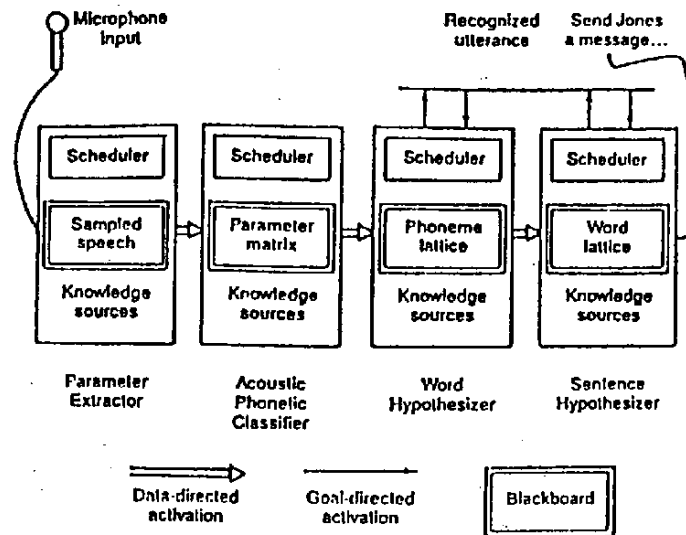


図 7.1-9 CMU 音声理解システム (ANGELS) [9]

音韻認識部では、ピッチ同期で DFT にかける、スペクトラム情報を計算する。その上で、音韻のクラスごとに Locator を動作させ、音韻のおおよその位置の候補を見付け出し、classifier で細かい音韻の認識を行う。音韻の認識結果は、音韻ラティスとして、確率と共に単語認識部に送られる。

単語認識部 (word hypothesizer) の開発は、Alex Rudnieky をグループリーダーとする約 7 名のワードグループで行われている。そこでは、音韻ラティスを基にして、単語スポッティングが行われる。スポッティング (matcher) は、音韻ラティスのほかに、ストレス情報、音節 (syllable) の核の情報なども用い、スポッティングを効率よく実行する。

言語処理部 (sentence hypothesizer) は、格フレーム (caseframe) による構文解析を用いて正しい文章を見付け出すことを目的としている。音韻認識部と単語認識部は、どちらかと言えば、ボトムアップ的に構成されていたが、この言語処理部は、トップダウン的な処理を中心としている。

分散システム (distributed system) グループは、Roberto Bisiani をリーダーとし、分散 AI システムを実現するツールの整備を目指している。

(2) ATR のシステム

ATR の自動翻訳電話研究所における音声認識の目標としては、4 年目に、文節ごとに区切って発声した約 1,000 単語からなる日本語音声声を認識する実験システムを作成する。更に 7 年目には、単語数を 3,000 単語に増やし、文節ごとに区切って発声した日本語及び英語の会話音声声を認識する実験システムを構築する。100 単語以下の少量の音声データを用いた発声者への話者適応が可能なシステムを目指す。

音声認識システムは、オブジェクト指向的概念による構成法を取り入れ、できる限り各処理部をモジュール化することを試みる。更に、各処理部は、共通に参照できる shared memory としての分散ブラックボード (distributed blackboard) を持つ。システム構成の一例を図 7.1-10 に示す。

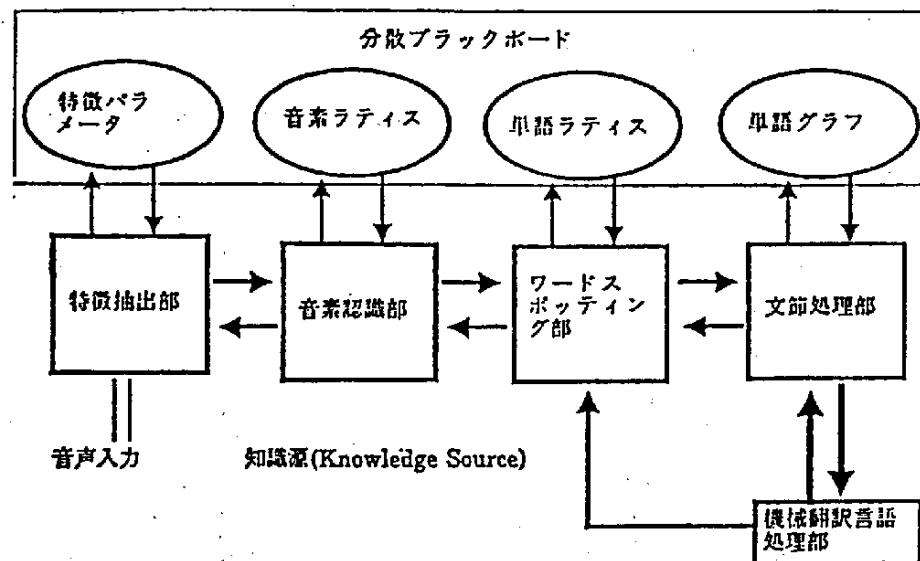


図 7.1-10 ATR 音声理解システム構成例 [9]

特徴抽出部は、スペクトログラム情報及び LPC 分析情報をベースとしてホルマント軌跡等の動的なパラメータも含む。

音素認識部ではスペクトログラムリーディング情報に基づく音素認識アルゴリズムと、LPC スペクトル情報を用いたパターンマッチングに基づく音素認識アルゴリズムを組み合わせる。

文節処理部はワードスポッティング部からの単語認識結果に基づき、構文、意味、分散ブラックボードの韻律情報を駆使して、後続の単語を予測して、その情報をワードスポッティング部に送り返す。

7.1.3.2 SPHINX システム [10]

(1) 概要

これまでいくつかの音声認識システムが良い性能を示しているが、それは次のような条件の下でのみ可能であった。すなわち、(1)特定話者、(2)離散単語、(3)小語い、(4)文法的制約。

これらの条件は、次の四つの欠陥に起因する。

1. 複雑ではあるが論理的な音声のモデルの不在、
2. 音響的、音韻学的、文法的な知識を適切に使用していないこと、
3. 学習可能であり、かつコンテキストに敏感でないような音声の単位が用いられていないこと、
4. 話者間の変動や話者の特性を扱えないこと。

この文献は、これらの問題に取り組んだ、大語い、不特定話者、連続音声認識システム、SPHINX について述べている。

不特定話者

不特定話者の達成は、最も困難な問題と考えられてきた。ほとんどの音声のパラメータ表示が強く話者に依存しているため、ある一人に合わせた標準パターンは他の人には合わないのである。

この問題に対する第一のアプローチは、AI的手法を用いて話者の変動に対して比較的不変なパラメータを見い出すことである。しかしながら、現在のところこの方法による実際のシステムは成功していない。

第二の方法は、一つの音韻や単語に対して、複数の標準パターンを用意することである。この方法は、大語いでは成功していない。

最後の方法は、新しい話者に対してシステムを適応させる方法である。しかしながら、これまでのシステムは真に不特定話者認識とはいえないものであった。

SPHINX では、前記の四つの問題点を解決するために、以下の四つの方法をそれぞれに適用した。

1) Hidden Markov Model

Hidden Markov Model (HMM) は、音声のローバストなモデルを提供する強力な方法である。SPHINX は、HMMを音韻から単語、さらに文章までのすべての知識の表現に用いる(図7.1-11)。標準的なアルゴリズムの他に、HMM の色々なバリエーションを実験的に用いた。

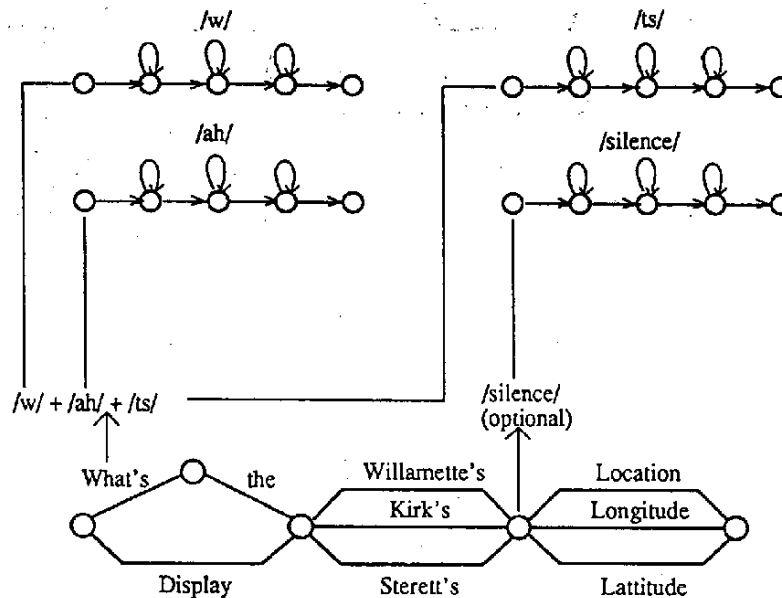


図 7.1-11 有限状態文法を用いた連続音声認識のための HMM [10]

2) 知識の導入

人間の聴覚の特性や音響学的知見をシステムに導入することを試みた。ケプストラム係数、パワー、微係数などはベクトル量子化のコードブックに新たなパラメータを追加することで容易に導入できる。音韻の継続時間や、母音前後の不規則な波形部分の継続時間、あるいはフォルマントの周波数などのパラメータを HMM に導入する方法についても幾つかの方法を試みた。

3) 音声を表示する単位

基本的には、phone モデルを用いた(図 7.1-12,13)。さらに function-word-dependence phone モデルを導入して“the, a”などの function words を明示的にモデル化した。これら function words は最も誤認識の多いものであるが、非常に多く出現するので学習しやすいものである。

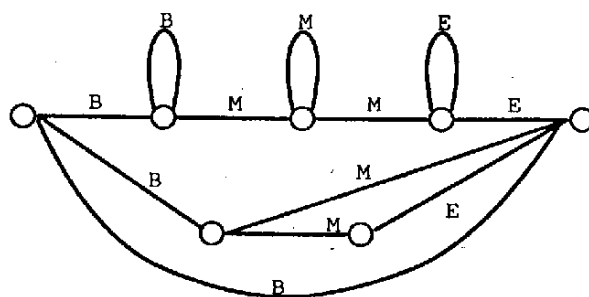


図 7.1-12 SPHINX の基本構成に用いた音韻の HMM [10]

Phone	Example	Allophones	Phone	Example	Allophones
/iy/	beat		/ng/	sing	/eng/
/ih/	bit		/ch/	church	
/eh/	bet		/jh/	judge	
/ae/	bat		/dh/	they	
/ix/	roses		/b/	bob	
/ax/	the		/d/	dad	
/ah/	butt		/g/	gag	
/uw/	boot	/ux/	/p/	pop	
/uh/	book		/t/	tot	
/ao/	bought		/k/	kick	
/aa/	cot		/z/	zoo	
/ey/	bait		/zh/	measure	
/ay/	bite		/v/	very	
/oy/	boy		/f/	fief	
/aw/	bough		/th/	thief	
/ow/	boat		/s/	sis	
/l/	led	/el/	/sh/	shoe	
/r/	red		/hh/	hay	/hv/
/y/	yet		/cl/	(nuv.clos.)	/p,t,k,q-cl/
/w/	wet		/vcl/	(voi.clos.)	/b,d,g-cl/
/er/	bird	/axr/	/epi/	(epin.clos.)	
/m/	mom	/em/	/sil/	(silence)	/h#/,/h#/,/pau/
/n/	non	/en/,/nx/			

図 7.1-13 SPHINX の基本構成に用いた音韻の表 [10]

4) 話者適応

HMM を話者に適応させるために、二つのアルゴリズムを提案した。第一番目は、話者クラスター選択

の方法である (図 7.1-14)。話者クラスター選択は多数の話者が幾つかのクラスターに分類されうると仮定する。したがって、話者を適応させるとき、まずどのクラスターに最も近いかを判定し、そのクラスターのコードブックを認識に使用する。

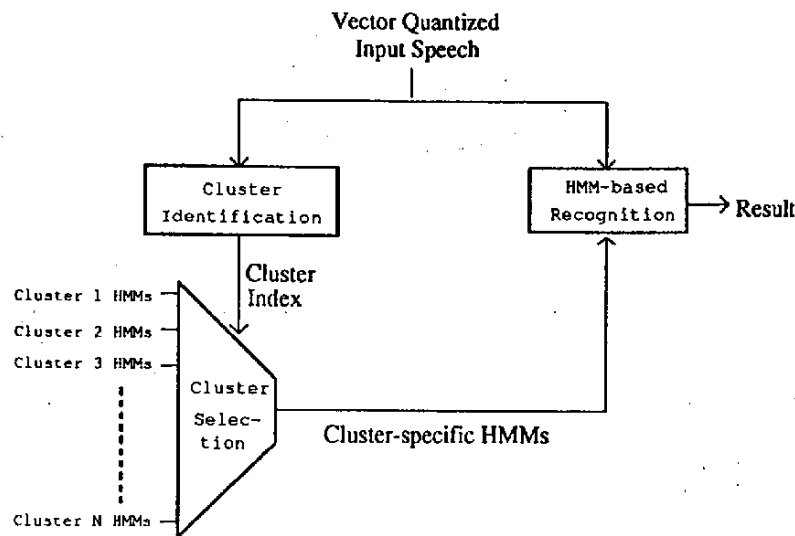


図 7.1-14 クラスター選択による話者適応 [10]

もう一つの方法は、interpolated re-estimation アルゴリズムである (図 7.1-15)。まず話者に依存しないコードブックと HMM から出発し、コードブックは変えずに HMM のパラメータを話者に合わせて変更する。

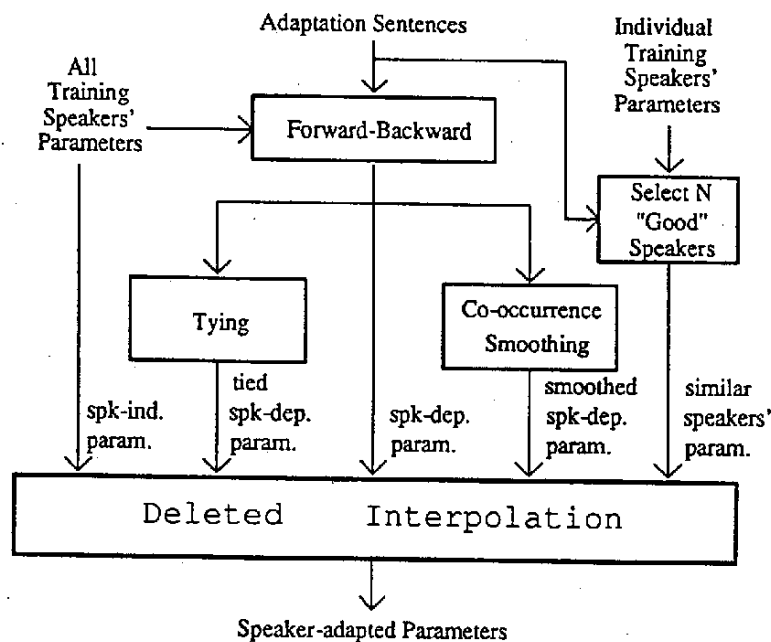


図 7.1-15 Interpolated re-estimation による適応 [10]

(2) まとめ

上記の方法を SPHINX システムに導入した。SPHINX は DARPA の 997 語からなる resource management task に用いて、不特定話者で、文法の perplexity が 997, 60, および 20 のそれぞれの場合、

71%, 94%, 及び 96% の単語正解率を得た。話者適応を用いた場合、誤認識率が凡そ 5-10% 減少した。これらの結果は、同様のシステムにおいて報告されたうち最高の結果である。

この研究において得られた最も重要で喜ぶべき教訓は、大話し不定話者連続音声認識は可能性があるということである。

7.1.4 音声合成

7.1.4.1 DECTalk [11]

本システムは、MIT の、Dennis Klatt によって開発された 3 層構造のソフトウェア (図 7.1.4.1-1) を用いており、ハードウェアとしては二つのプロセッサをワンボード化した。Klatt の援助の元に DEC で開発されたハードは、一般的なチップ (Motorola の 68000) と TI の DSP チップ (TM 32010) を用いている。ソフトウェアはすべて C 言語で書かれており移植性がよい。

第一段目のソフトは、アスキーコードで書かれたテキストから、音声のデジタル表示を発生する。次段では、その表示から、音響パラメータを生成する。第三段目において、音声信号を発生する。音声のデジタル表示は、音素に対応している。英語においては、約 40 の音素で十分である。“The old man sat in a rocker” は、音素表示で書くと、

DH IY' OW LD M' AE N) S' AE T IH N, AX R' AA K RR.

となる。

注意すべき点の一例として、1920 は “one thousand nine hundred and twenty” と読み、めし “nineteen twenty” と読み。同様に、数の読み方に関する特別の注意を払わないと、“\$ 12.40” を “dollars twelve point forty” と読んでしまう。

高い品質を得るためには、大きな語いの辞書と大きな規則集を必要とする。開発システ

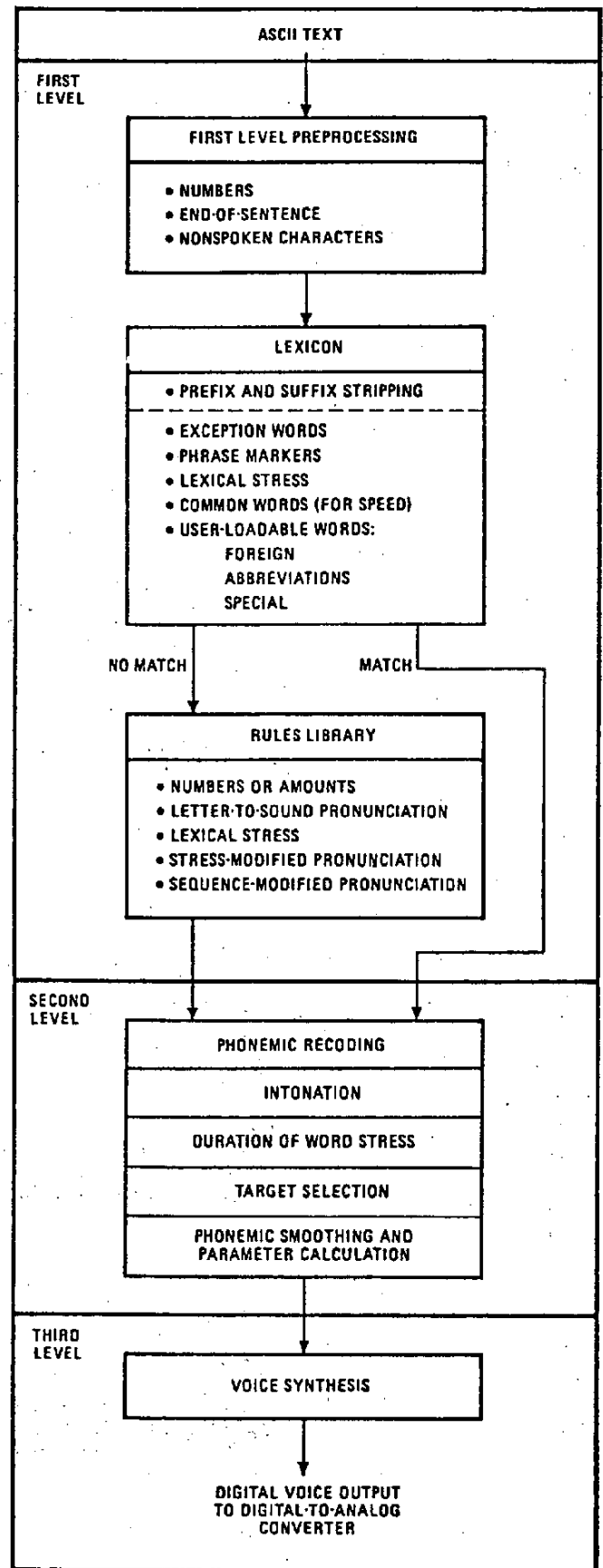


図 7.1-16 Dectalk におけるテキスト音声合成のための 3 段階の処理 [11]

ムにおいては、500 近い「文字から音への規則」を用意した。

辞書は約 7000 語からなる。1 語 20 バイトで 140 kbyte が ROM に書き込まれている。正しく選ばれたこの程度の大きさの辞書を用いて、「文字から音への規則集」と共に用いることによって、標準的なポケット辞書に見られるおよそ 20,000 単語を正しく発音できる。

人間の音声を作成するとき、各フォルマントの周波数を滑らかにつなぐことが大事である (図 7.1-17)。DECTalk は、子音-母音遷移の軌跡の理論とターゲットへの漸近に基礎を置く複雑なスムージング処理を採用している。スムージング処理は、各音韻のそれぞれのフォルマン周波数に対して行われる。

全部で 18 個の制御パラメータが第二段階の処理部 (TM32010) で計算され、音韻系列の順番通りに合成装置へ送られる。

システムのハードウェア構成を図 7.1-18 に示す。6800 には 256 kbyte の ROM がつながっており、そのうち約半分が辞書で占められる。RAM は 64 kbyte である。このシステムは、RS 232 C により端末およびホストから呼ぶことが可能である。発声速度は一分間 120 語から 300 語まで可変である。男性、女性、子供などの声を選ぶことができる。

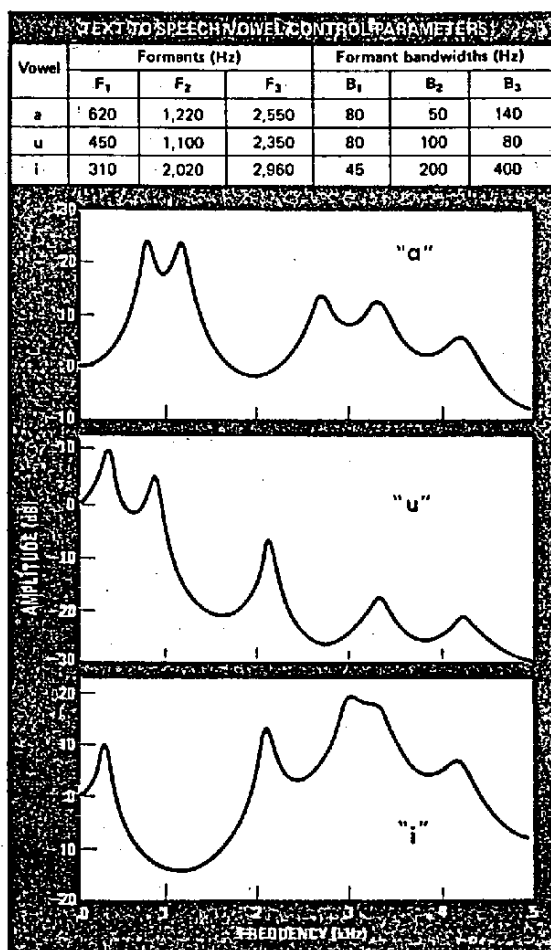


図 7.1-17 音質のためのキイになるパラメータ [11]

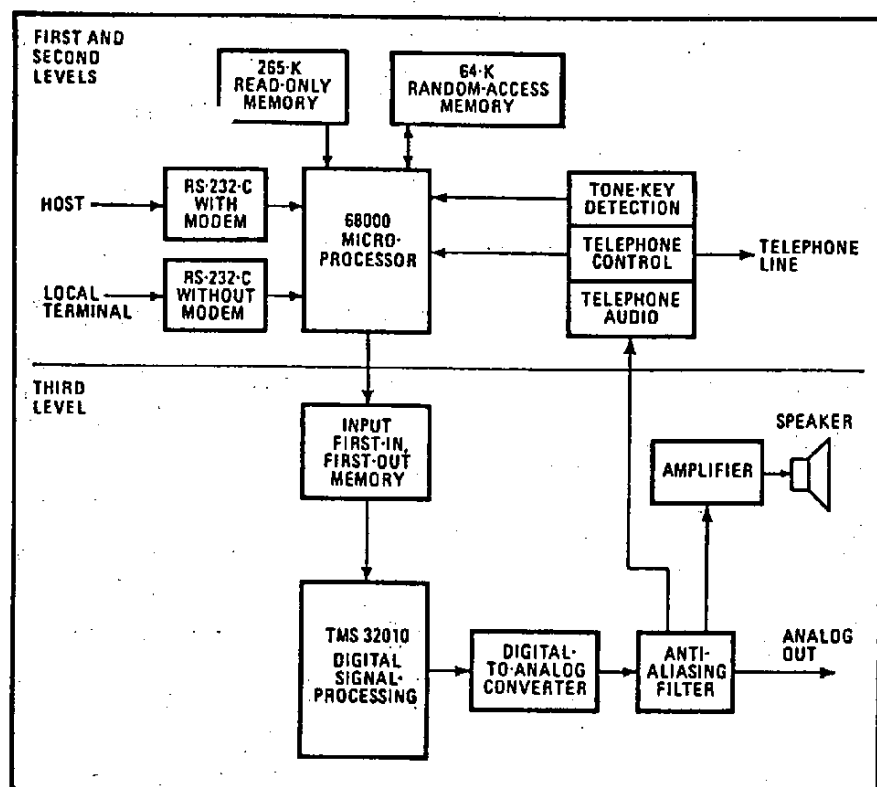


図 7.1-18 Dectalk のハードウェア構成 [11]

7.1.4.2 NTT の日本語テキスト音声合成システム [12]

日本語の漢字・仮名混りの文書（テキスト）を音声に変換する日本語テキスト音声合成システムを作成した（図 7.1-19）。入力となる文章は、約 58,000 単語の辞書と単語の接続を規定する単語検索テーブルとを用いて解析され、読み、アクセント、文法情報を持つ単語・形態素の系列として表現される。次に、近接した語の相互依存関係を定めることにより、複合語、文節、アクセント句、呼吸段落などより高次の言語単位への再構成を行い、複合した語のアクセント、イントネーションなどの韻律パラメータを決定する。最後に、CV を単位とする音声合成法によりテキストに対応する音声波形を合成する。

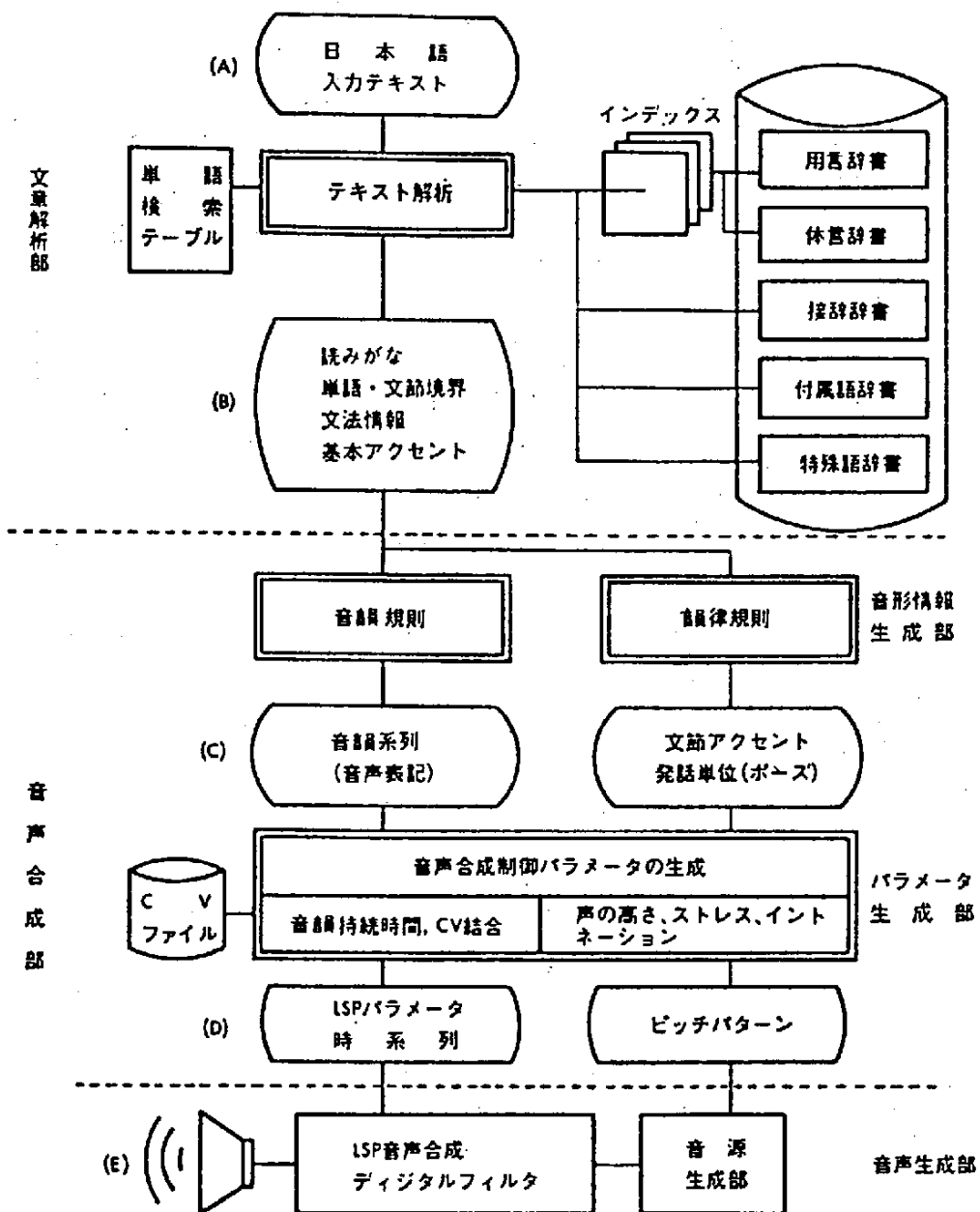


図 7.1-19 NTT の日本語テキスト合成システム構成 [12]

(1) テキスト音声合成における日本語処理の問題点

1) 分かち書き化

たとえば,

(I) ○東京/都/の/産業/別/人口/は

(II) ×東/京都/の/産業/別人/口/は

2) 漢字の読み

〔家〕: 政治家, 将来軍家, ○○家新館

人気 (にんき, ひとけ), 今日 (きょう, こんにち)

3) 表記のゆれの呼吸

(i) 漢字表記のゆれ…… (終生-終世), (双葉-二葉)

(iii) 送り仮名のゆれ

申し込み書-申込み書-申込書

4) アクセント

移動……かえる (帰る) + せる → かえらせる (帰らせる)

消失……けいざい (経済) + てき (的) → けいざいてき (経済的)

(2) 日本語解析用辞書

文章解析用辞書として, 三省堂「新明解国語辞典 (第二版)」から, 自立語について, 見出し語, 漢字表記, 文法情報, アクセント型を抽出して, 新たに辞書を構成した。単語数は約 58,000 語である。

参考文献

- [1] 伊福部達, “音声タイプライタ”, CQ 出版録, 1983
- [2] 千葉成美, “音声の自動認識”, スペクトラム, Vol. 01, No. 09, 1988-09, pp. 40-49
- [3] 中川聖一, “音声認識における時系列パターン照合アルゴリズムの展開”, 人工知能学会誌, Vol. 3, No. 4, 1988, pp. 414-423
- [4] Victor W. Zue, “The Use of Speech Knowledge in Automatic Speech Recognition”, Proc. IEEE, Vol. 73, No. 11, 1985, pp. 1602-1615
- [5] 溝口理一郎, 角所収, “知識工学を応用した音声認識”, 日本音響学会誌, Vol. 42, No. 12, 1986, pp. 942-947
- [6] 河原英紀, “神経回路網モデルと音声認識”, 人工知能学会誌, Vol. 3, No. 4, 1988, pp. 453-460
- [7] 大河内正明, “Hidden Markov Model に基づいた音声認識”, 日本音響学会誌, Vol. 42, No. 12, 1986, pp. 936-941
- [8] Frederick Jelinek, “The Development of an Experimental Discrete Dictation Recognizer”, Proc. IEEE, Vol. 73, No. 11, 1985, pp. 1616-1624
- [9] 鹿野清宏, 樽松明, “音声理解研究の動向”, 日本音響学会誌, Vol. 42, No. 12, 1986, pp. 948-952
- [10] Kai-Fu Lee, “Automatic Speech Recognition”, Kluwer Academic Publishers, Boston, 1988
- [11] E. Bruckert, M. Minow, and W. Tetschner, “Three-tiered software and VLSI aid developmental system to read text aloud”, Electronics, April 21, 1988, pp. 133-138
- [12] 佐藤大和, 匂坂芳典, 小暮繁, 嵯峨山茂樹, “日本語テキストからの音声合成”, 通研実用化報告, Vol. 32, No. 11, 1983, pp. 2243-2252

7.2 画像処理技術

画像処理技術をヒューマン・インタフェースという立場から分類すると、画像入力、画像作成、画像蓄積・検索、画像出力、に分類できる。しかし本節では、3章の物理的入出力と重複しないように、(1)画像認識(入力関連)、(2)コンピュータ・グラフィックス(特にアニメーション、作成・表示関連)、(3)画像符号化及びデータ構造(蓄積・検索関連)、の3分野に分類する。そして、今後のヒューマン・インタフェースにとって必要となるであろう画像処理の要素技術に関し、特に重要なものや興味を引かれるものについて、各々3～4件の論文を紹介する。

以下に分野ごとに選出した論文の概要を示す。

(1) 画像認識(文献[1]～[3])

まず、パターン認識関連の学会の中でも最も権威のある認識国際会議の最新の子稿集から2件の論文を紹介する。

文献[1]はこの国際会議で最も多い23件の論文発表があった「3次元画像処理と物体認識」のセッションから選んだもので、カメラ(range finder)と物体との距離データを用いた物体認識の最近の進歩に関するものである。この他のセッションでは「ロボットの目」に関する論文発表が多く、この2つのテーマが画像認識研究の中心になってきている。

文献[2]はヒューマン・インタフェースの入力系にとって不可欠である文字認識に関するもので、漢字認識研究の動向を述べたものである。

文献[3]は最近研究が増えつつある視線感知に関する一手法である。視線感知は将来のヒューマン・インタフェースに於いて重要な役割を果たす技術であると考えられる。

(2) コンピュータ・グラフィックス(文献[4]～[7])

コンピュータ・グラフィックスの中でも、ヒューマン・インタフェースにおいて視覚的效果の大きな意図伝達メディアであるコンピュータ・アニメーションに関する文献を取り上げる。

文献[4]～[6]は、計算機上で定義する物体の形状や動きのモデルに物理的法則に関する知識を取り入れる物理ベース・モデリングに関する文献である。このモデリングは、より現実世界に近い物体や環境を計算機の中に視覚的に構築するという視点から、今後盛んに研究されていくテーマであり、将来のグラフィックスのモデリングの主流の一つとなろう。

文献[7]は、自然言語から自動的にアニメーションを生成するシステムに関する文献であり、誰でもが容易にアニメーションを作成できるための重要な技術である。

これらの技術をヒューマン・インタフェースの視点からみると、①リアリスティックな視覚的対話、②高品質なグラフィックスやアニメーションの容易な作成、等を実現するものであり、一層のビジュアル・インタフェースの普及が進むと考えられる。

(3) 画像符号化とデータ構造(文献[8]～[11])

画像符号化に関しては、蓄積・検索を効果的に行う符号化及びデータ構造に加え、通信分野で特にホットな話題となっている知的画像符号化を紹介する。

文献[8]は、画像をピラミッド式に階層化し、縮小画像の高速アクセス及び高能率データ圧縮を可能とする符号化方式に関する文献である。これは、本のページを速くめくって目的とする画像を捜し出す方法を、計算機上でユーザフレンドリーな検索インタフェースとして実現するために必要な技術である。

文献[9]及び[10]は、送信端及び受信端が互いの知識を共有し、通信路において真に本質的なやりとりを

行い、ユーザにはその情報処理様式に最も整合したインタフェース環境を提供する知的通信を実現するための知的画像符号化に関する文献である。この符号化は、画像認識・理解・生成といった画像処理技術をベースにしており、通信のみならず、知的なヒューマン・インタフェースを実現する一つのアプローチと考えられる。

文献[11]は、ハイパーメディア・システムを構築するためのデータ構造とシステム構築法に関する文献である。特に、動画像等を扱うための時間軸の概念の導入に関する研究は、今後ヒューマン・インタフェースの基礎となるマルチメディア・データベース・システムを構築する上で不可欠であろう。

尚、各論文の紹介においては、(1)全体概要（論文の概要、ヒューマン・インタフェースの視点からの報告者のコメント）、(2)技術的なポイント、に分けて述べている。

7.2.1 Recent Progress in the Recognition of Objects from Range Data

(1) 全体概要

本論文では、距離データによる物体認識について、まず①距離データの取得について簡単に述べた後、②距離データからの3D物体認識のパラダイムを示し、③近年の主要な研究について簡単に紹介するとともに、それぞれについて問題点を指摘している。

距離データを用いた3Dコンピュータビジョンは大いに進歩してきたが、依然として基本的な問題がいくつか残っている。まず、物体記述が難しい問題である。多面体記述はモデリングには十分な能力はないと思われ、曲率記述は雑音を含みやすい。(距離データの)平滑化はある場合には幾何学的形状を変えてしまうし、時間がかかることもある。さらに、マッチングの問題は物体記述が十分な能力を持つまで解決はできないと思われる。ただ、これらの処理時間がかかる点については、並列処理アーキテクチャを持つコンピュータに期待がかけられる。(扱われるデータの局所性が高く並列処理に向いている)

(2) ポイント

1. 距離データの取得

レーザーを用いて縞模様や格子点状のパターンを物体に照射し、それらをカメラで計測する方法（三角測量の原理）と、超音波や平行光線を投射して、物体から反射してくるまでの時間を計測（変調波の位相のずれ、または、パルス発振による直接計測縞する方法の2種類に大別される。前者はレーザスキャナとカメラの角度によって、エッジの周辺のデータが欠落したり、必ずカメラから隠される部分が存在するという欠点がある。後者はより正確な距離データを得ることができるが、解像度がビームスポットの大きさに依存することや、ビームの出力をあまり大きくできないという欠点がある。

2. 3D物体認識のパラダイム

3D物体認識の一般的なパラダイムについて概説している。すなわち、距離画像から物体のシンボリックな記述をつくる処理は、(i)有意な特徴を決定し、(ii)この特徴を抽出し、(iii)これらを組み合わせ物体を記述する、という3段階からなる。その後、認識の処理として、(iv)モデルのマッチングをとり、(v)場合によっては物体の配置や向きまで決定する、という2段階の処理を行う。

3. 近年の主要な研究

プライマルスケッチにもとづくもの、曲率にもとづくものなど、いくつかの流れが存在し、それぞれ1980年代初頭から継続して行われているが、いずれも決定打とはなりえず、なお今後の進展が望まれている。

7.2.2 Chinese Character Recognition : A Twenty-Five-Year Retrospective

(1) 全体概要

本論文では、漢字認識研究の草分けである著者が研究を行った約25年前と比較しつつ、漢字認識技術および漢字 OCR の現状を参考文献を紹介しながら概説している。漢字認識では、従来の英数字認識とは異なり、2000 種類以上の文字を認識せねばならない。従って、類似文字が非常に多いため、英数字認識の手法を延長では漢字は認識できない。漢字認識はパターン認識の分野における大きな挑戦であった。しかし、主としてアジアの多くの研究者の努力による認識アルゴリズムの開発と、低価格な半導体スキャナ、プロセッサ、メモリ等の関連技術の進歩により、現在では印刷漢字や手書き漢字が認識できる OCR が商品化されるに至っている。これらの認識率は99%以上と宣伝されているが、各種の応用分野で受け入れられるためには誤読率が少くとも0.1%未満になる必要がある。

(認識率+誤読率+リジェクト率=100%)

文字認識によるデータの入力にはヒューマン・インタフェース改善には不可欠であり、今後は認識率向上とともに各種のドキュメント(伝票、書籍、新聞、名刺など)を読み取る技術および装置の研究開発が続けられるであろう。

(2) ポイント

1. 前処理

前処理技術には、2値化、行および文字切り出し、フォーマット分析などがある。特に文字切り出しでは、文脈を利用して切り出し精度を上げる方法が発表されている。郵便物の宛名認識の実験では、一度文字切り出しをして認識したあと、住所として意味を成さない部分に切り出しエラーがあると推定して、条件を変えて再度切り出すことにより、切り出し精度を80%から90%に改善した。

フォーマット分析は文書画像を文字列と図表とノイズの各領域に分ける技術で、射像、線密度、連結成分の外接四辺形などの特徴が利用されている。新聞の見出しなどの様に、縦書きと横書きが混在する場合などは今後の課題である。

2. 階層識別

漢字認識では認識対象文字が多いので、入力文字パターンを全ての文字の標準パターンと照合していると非常に時間がかかる。このため、認識対象文字をいくつかの類似文字グループに分けておき、第一段階でどのグループに属するかを調べてから第二段階でどの文字であるかを識別する。階層識別は多くの場合2階層が採用されているが、多階層の判定木が用いられている場合もある。

3. 後処理

英文用 OCR では単語辞書との照合によるスペル・チェックが行われる。漢字 OCR でも住所や姓名などの様に文脈がわかっている場合には、単語チェックが認識率向上に有効である。更に、文章の読み取りでは文法を利用した誤読文字チェックも研究されている。

7.2.3 顔画像計測に基づく視線感知法

(1) 全体概要

人間と計算機がスムーズに対話するための一手段として、計算機に目を持たせ人間の視線を感知する研究が行われている。具体的な応用例としては、作業員が視線でコントロール・パネル上のメータ等を指し示したり、逆に人間がメータを監視しているか否かを確認したりすることが考えられる。従来から視線検出装置

としては、眼鏡に取り付けて「何処を見つめているか」を検出するアイカメラなどが研究されている。本論文は、考え方を逆転させ、「何処が見つめられているか」を顔画像から検出することにより、完全非接触型でユーザに負担のない視線感知法を実現する画像処理アルゴリズムについて述べたものである。

従来、使用条件を選ばない高精度な視線感知が完全非接触型で実現できれば、ヒューマン・インタフェースの大幅な改善が期待できる。

(2) ポイント

1. 視線感知の原理

顔画像中の視線の向きは、目の画像だけから判定することは難しく、顔の向きの情報が必要となる。図7.2-1に示すように、カメラに対する視線の確度 (θ) は、カメラに対する顔の確度 (θf) と、顔に対する眼球の回転角 (θe) との和である。

眼球の顔に対する回転角 (θe) は図7.2-2に示すように、左右の眼頭間の中点Eから左右の黒目間の中点Bへの距離 αl と、黒目径 (= 目球半径) a の比から計算される。

顔のカメラに対する回転角 (θf) は、図7.2-3に示すように、左右の目頭間の中点をE、顔の左右の端点をF₁、F₂として、F₁E と F₂E の比から計算される。

2. 実験結果と今後の課題

髪の毛など頬に掛かっていず、かつ耳の出ている顔画像について、視線がこちらを向いているか否かを実験した結果は、人間による感知実験とほぼ同じ精度で検出することができた。今後は、顔の両縁が髪の毛で覆われて見えない場合や、眼鏡使用の場合への適用や、動画から高精度に視線感知するための本手法の高速化が課題である。

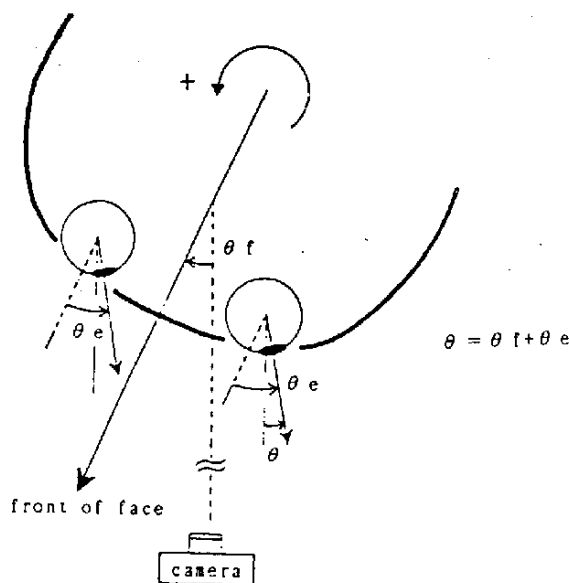


図 7.2-1 視線感知の原理

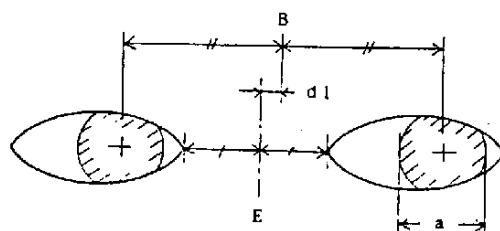


図 7.2-2 眼球の回転角の求め方

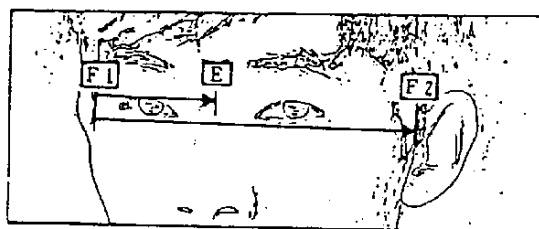


図 7.2-3 顔の回転角の求め方

7.2.4 “Physically Based Modeling for Vision and Graphics”

(1) 全体概要

本論文では、物体のグラフィックス・モデルに自然界の物理法則を導入した“物理ベース・モデリング”

の特徴と著者のグループが開発したいいくつかのモデルについて概説している。このモデリングは、①形状モデリングとモーション・デザインの統合、②直感的に把握しやすい対話型の形状修正、③リアリスティックな動きを容易に作成、④画像生成だけでなく、コンピュータ・ビジョンへの適用、等の特長を有する。

この論文は、Pixar社のLasseter J.氏が伝統的アニメーション技法を用いて作成したアニメーション・フィルム“Luxo Jr.”での主人公の動きを物理ベース・モデルで実現したということでも話題となった。これはアニメーションを用いたヒューマン・インタフェースの視点から見ると、視覚的効果のあるアニメーションを自動的に生成できるという点で大きな意味を持つと考えられる。また、コンピュータ・ビジョンにおける形状抽出及び連続画像から物体の動き追跡の能力を有するため、撮像画像からの物理ベース・モデルに基づいたアニメーションの自動生成の可能性も考えられる。

(2) ポイント

1. 物理ベース・モデル

エネルギー関数とその極小化をうまく利用することにより、比較的少ない計算処理で、物理的なアナロジーを確立・維持できる物理ベース・モデリングである。これにより、①で述べた特長を実現できる。

2. 個々のモデル

著者のグループが開発したモデルとして、Parameterized Model, Elastic Model, Snake Model, Symmetry-Seeking Modelを概説している。

Elastic Modelは薄板を張るエネルギーから導出されたエネルギー関数を用いたモデルであり、モデルの任意の部分に力を加えることにより、そのモデルの形状を変化させることができる。これにより、対話的な形状モデリングが直感的に、容易に行える。

Snake ModelはElastic Modelに、画像のグラジエントの局所的最大点に向かって形状を引っ張るようなエネルギー関数を加えたものである。このモデルを用いることにより、TVカメラ等から入力した画像から物体の形状抽出を対話的に行ったり、連続画像から物体を追跡するというようなコンピュータ・ビジョンの処理も可能である。

3. Spacetime Constraints

アニメーションによりダイナミック・シミュレーションを行う際の難点である制御の問題（ユーザによる初期設定とアクティブな力の指定）を解決するアルゴリズムである。動きのパスを時間軸上でサンプリングした後、各サンプルを集めた一つの大きな状態ベクトルを作り、そのベクトルに対してエネルギー制限を与えることにより、動きを自動的に生成できる。

7.2.5 Flocks, Herds, and Schools: A Distributed Behavioral Model

(1) 全体概要

本論文では、鳥や魚の群れの行動をリアルにアニメーションするためのモデリングについて、アルゴリズムの内容と実験結果を述べている。群れにおいては、個々の個体が各々の意志に従って行動しながら、群れ全体として統一的な動きをする。群れを構成する個体数が多くなると、キーフレーム法のような従来手法では複雑すぎてアニメーションを生成できないため、雲の動き等を表現するのに用いられる“Particle Model”（対象を微粒子とみなし、各粒子は確率的に発生・移動・消滅する）を発展させたモデリングを研究している。

単一のキャラクターだけでなく、非常に多数の個体からなる群れをリアルにシミュレーションすること

は、計算機上でよりリアリスティックな世界を実現するというヒューマン・インタフェースの重要な課題を解決する一つのアプローチであろう。

(2) ポイント

1. アニメーション・モデル

単に形状や物理的性質をモデリングするのではなく、キャラクターの行動（振舞い）を作る“Behavioral Animation”を実現するため、Particle SystemとActor Systemをベースにしている。このモデルにより、個体が群れの中に留まるように求心的に動くと同時に、他の個体と衝突しないように遠心的に同期して動くことができる。また、群れが二つに分離したり、二つの群れが合体するといった群れ特有の行動もシミュレートできる。

2. 幾何学的飛行

グローバル座標系と個体自身のローカル座標系を有する。ローカル座標系にて個体が受ける物理的力として、進行方向の軸の回りの回転を利用した“Banking”を採用している。

3. 群れのシミュレーション

自然界での群行動の研究を基にして、①近隣の個体との衝突回避、②近隣の個体との速度マッチング、③近隣の個体に接近することによる群れの求心、の3種の独立した行動をシミュレートする。

3種の行動を仲裁する方式として、衝突発生等の緊急事態において最も要求の高い行動が選択されるように、優先権方式を採用している。

よりリアルな動きを実現するために、近隣の局所的な視野範囲の距離に基づいた方式を採用して、個体の知覚をシミュレートしている。また、他の個体だけでなく、周囲の障害物との衝突回避も行っている。

4. 群れの動きの指定

群れの動きとして、①初期設定のみ与えて、目的なく勝手に振舞う行動のシミュレーション、②動きのパスを位置と時間で指定して、ある目的を有する行動のシミュレーション、が可能である。

7.2.6 Using Dynamic Analysis for Realistic Animation of Articulated Bodies

(1) 全体概要

本論文では、人間等の関節のあるキャラクターのリアルなアニメーションを実現するための力学モデリングについて、アルゴリズムの内容とアニメーション作成システムを詳細に説明している。この力学モデルは、身体各部分の質量と外部及び内部の力を考慮しており、自然な動作が自動的に生成できる。特に、外部力としては重力と床の反発力及び摩擦を、内部力として関節の部分にバネとダンパを仮定している。

これも物理ベース・モデリングの一つであり、他のモデリングと同様な特長を有するが、特に、リアルな人物をディスプレイの中に召使として登場させるというヒューマン・インタフェースを実現するために必要な要素技術の一つである。

(2) ポイント

1. 力学を用いたアニメーション

運動学に基づくキーフレーム法では、ユーザが動きのパスと時間を全て指定する必要があり、人間のモデルをリアルに動かす場合などは非常に大きな負荷となる。それに対し力学に基づく方法では、リアルな動きを生成するための制限が自動的に与えられるので、詳細な指定を行うことなく容易に作成できる。問題とし

ては、計算時間、複雑な形状の際の数値解析の不安定性、動きの制御があげられる。

このモデルでは、計算時間はかかるが、幾つかの要素をうまく記述できる“Gibbs-Appel dynamics formulation”を用いている。

2. 力学を用いた動きの制御

このモデルでは、重力、結合部の制約、床からの反作用力といった力及びトルクが自動的に計算される。結合部の制約は、各関節の自由度に対してバネとダンパを仮定して、自然な動きが生成できる。また、床からの反作用力としては、バネとダンパにより表される垂直方向の力と、水平方向に働く摩擦力をシミュレートする。

力やトルクを制御する上で最も難しいことは、指定した動作を生じさせる内的な力やトルク（筋肉の動き等）を見いだすことである。しかし現状では、筋肉等の動きを考慮できる高度なモデルではなく、個々の自由度に対してのみ働く理想的なアクチュエータとして、制御の力やトルクを扱っている。そして、各自由度は、①直接的力学制御：力やトルクを指定、②弛緩：いかなる力やトルクも働いていない、③凍結：ローカルな位置を保持、④方向付け：ワールド座標での方向を維持、⑤ハイブリッドな位置制御：ユーザが指定したキーフレームから、それを達成する力やトルクを求める、の5つの状態の一つにある。

3. 動き制御のためのグラフィックス・エディタ

動き記述を容易にするために、ユーザフレンドリなエディタがインプリメントされている。主な機能としては、①ユーザが自由度に対する制御関数（力とトルクの時間の関数）を設計する、②ユーザが時間軸上で自由度の状態を5つの状態から選択する、がある。

7.2.7 自然言語を解釈し動画を自動生成するアニメーション・システム

(1) 全体概要

本論文では、自然言語（日本語）で書かれた物語からアニメーションを自動的に生成するストーリー駆動型アニメーションにおける、台本作成、シーン演出、動き生成の各機能について概説している。これらの処理部を推論エンジンを有するエキスパート・システムで実現し、簡単なアニメーションではあるが自動生成できることを実証している。

自然言語でアニメーションを自動生成できれば、誰でもが自分の頭の中にあるイメージを映像として表現できることになり、言語等のバックグラウンドが異なる人間同士の意図伝達のためのヒューマン・コミュニケーション・メディアとして、アニメーションが広く利用されることになる。

(2) ポイント

1. アニメーション自動生成システムに必要な機能

- ① 物語理解部：自然言語で書かれた物語から台本を作る機能
- ② 演出部：舞台装置等の演出項目を決定して詳細な台本を作る機能
- ③ 動作生成部：詳細な台本から実際の動きを生成する機能

2. 物語理解部

物語の内容を深く理解する必要はないが、動作に関しては十分な理解が必要である。そのため、構文解析知識と動作に関する知識からなる辞書を有するエキスパート・システムとして実現されている。特徴としては、①物語を時間的に変化する状況を並べたものととらえる、②間接的表現から生成に必要な事項を推測（前向き推論）、③常識を用いて適当に決定（仮説設定）、④不足した記述を前後関係から推測（後向き推

論), ⑤時間推移による矛盾の除去(首尾一貫性保持機構), がある。

3. 演出部

まず, 登場人物と位置及び時間関係についてのルールを用いて, 物語理解部で作成された台本をシーンに分割する。次に, 演出に関する人間の経験的知識を使って, シーン単位に登場人物, 動作, 位置などの演出項目を詳細に記述する。

4. 動作生成部

演出の加えられた台本から, アニメーションの具体的な時刻の入った動作スケジュールを作成した後, 各コマの画像を生成する。台本からは具体的な時刻等の詳細な指示は期待できないので, 物語中の定性的時間記述から実際の動作の実行を行うイベント駆動型スケジューラ及び状況判断メカニズムが必要となる。ここでは, 動作物体及び動作をオブジェクトとして定義し, 各物体動作の待機リストと動作リストを用いたスケジューリングを行うとともに, 動作に埋め込まれた条件文による条件判断を行っている。

尚, 関節のあるキャラクタである動作物体は, オブジェクト指向言語をベースに木構造で記述され, 関節の回転を組み合わせて複合的な動作を生成する。

7.2.8 画像データベース向き, 2階層符号化方式

(1) 全体概要

2階層符号化方式は画像検索への適合性を重視した符号化方式であり, 縮小サイズ画像(索引画像)を高速に表示するための索引画像データと, それを補間して通常サイズの画像を十分な品質で得るための補間データとに分離したデータ構造をとる。画像検索ではキーワードなどで選ばれた複数の候補画像を一覧し, 目的とする画像を絞り込む必要がある。2階層符号化方式では複数の索引画像を高速表示することにより, 候補画像を一覧することができる。そして, 目的とする画像の再生では索引画像データと補間データとによって通常サイズの画像が再生される。

補間データの作成においてはブロック符号化をベースにしているため, 従来のアダプティブブロック符号化との効率比較を行っている。その結果では, 2階層符号化方式の方が符号量, S/N評価値, 目視評価において改善がみられる。

この符号化方式は, 検索マシンとの組合せにより, 静止画検索システム FORKS として試作されている。このシステムでは, 索引画像が1画面に16枚表示され, 1枚の索引画像当り0.3秒程度で表示可能になっている。

(2) ポイント

1. 索引画像データの作成

原画像を縦横4画素間隔でサンプルして索引画像データを作成する(図7.2-4)。索引画像の画素値にはサンプル点近傍の平均輝度を用いる。これは, 折り返し雑音除去と原画像に含まれる雑音除去が目的である。索引画像は無圧縮で格納されるため, ディスクからのDMA転送により高速表示が可能である。

2. 補間データの作成

索引画像に基づいて 4×4 ブロック内画素の輝度値を2次元線形予測する(図7.2-5)。そして, 原画像との予測誤差の程度に応じて, ①補間データを作成しないモード, ②予測誤差をブロック符号化するモード, ③予測誤差をダイナミックに多階調符号化するモードに分けて補間データが作成される。

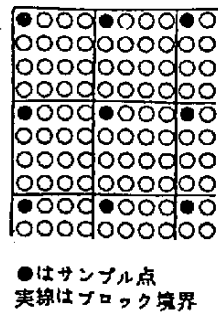


図 7.2-4 索引画像データ

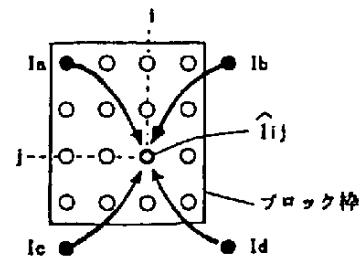


図 7.2-5 2次元線形予測

7.2.9 顔画像の分析合成符号化方式の動向

(1) 全体概要

次世代の画像通信を目指して知的画像符号化の研究が開始されている。本論文では従来の波形符号化とは全く異なった考え方に基づいた顔画像の分析合成符号化の研究動向について述べている。即ち、顔画像の分析合成符号化では図 7.2-6 に示すように送信側と受信側において顔の 3 次元モデルを共通の知識として持ち、送信側では構造モデルに基づいて被者画像の分析を行い、構造情報、動き情報、表情情報、更新情報が受信側に伝送される。受信側では、受け取った各種情報と顔の 3 次元モデルから顔画像の合成を行う。

ここで、画像分析には「画像処理技術」が利用され、画像の合成においては「コンピュータグラフィックス技術」が利用されることになり、知的符号化、特に分析合成符号化は従来独立的に発展してきた両技術を「符号化」によって統合しようとする試みと捉えることができる。その結果、「符号化」本来の目的である情報圧縮以外に、文字→画像変換等のメディア変換、知的インタフェース等の中核技術になると考えられる。

(2) ポイント

3 次元モデルの構成法、顔画像の分析を行う上で必要となる特徴点の抽出、動き追跡、表情の分析と合成について概説している。

1. 3 次元モデルの構成法

標準的な顔の 3 次元ワイヤフレームモデルを用意しておき、これを対象とする人物の顔画像に応じて変形する。その後、これに顔画像を貼付けるというテクスチャマッピング技術を用いる。正確な 3 次元形状のモデルは作成できないが、簡単でかつ自然な顔画像が得られる。

2. 表情の分析と合成

表情の分析は問題が山積みしており、決め手となる技術が無い。合成については、3 次元モデルの形状変形によるアプローチが研究されている。これには、表情の変化が引き起こす特徴点の変位を求め、それにモデルの形状を合わせるという従来技術を用いた方法、表情は基本となる表情の合成で表現できるという心理学分野での研究を利用し、基本となる表情に対応する 3 次元ワイヤフレームモデルの変形規則を定めるという全く新しいアプローチとがある。

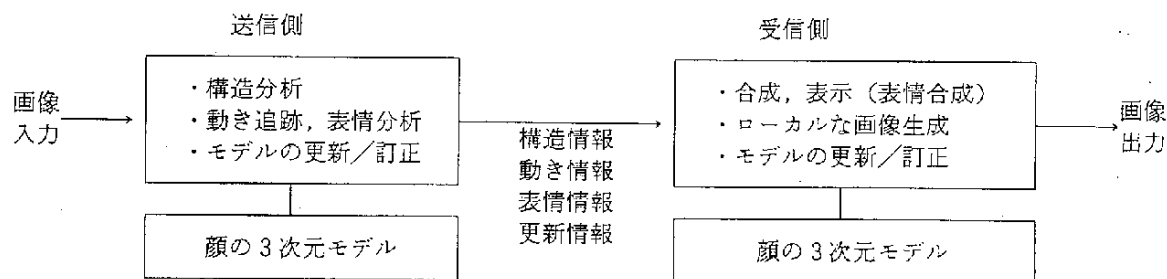


図 7.2-6 顔画像の分析合成符号化

7.2.10 画像と音声の知的インタラクティブ符号化の一検討

(1) 全体概要

画像と音声の知的インタラクティブ符号化は、TV 電話や TV 会議の様に顔画像が出現する会話のシーンにおいて、音声情報と画像情報との関連を考慮して情報を送信することによりデータ圧縮を図ろうとする符号化方式である。ここでは、受信側で顔の 3 次元ワイヤフレームモデルを保有していることを前提にして、テキストからの表情合成と、音声からの表情合成について述べている。即ち、テキストや音声情報に基づいて前記 3 次元モデルを変形することにより表情を合成する。これは、顔画像の分析合成符号化と共通の技術である。このとき、画像伝送のための情報量は不必要であり、大幅なデータ圧縮が可能となる。

又、この手法はデータ圧縮以外にメディア交換技法とも捉えることができ、ユーザフレンドリな知的インタフェースの実現ということが応用分野として揚げられる。より具体的には、送信側に機械を置き、人間と接する受信側で顔画像を合成することにより、あたかも人間同士が対話しているかのような環境を提供できる。

(2) ポイント

1. テキストからの表情合成

テキストの各音韻記号に応じて、予め用意された表情変形ルールに基づいて 3 次元ワイヤフレームモデルを変形する。表情変形ルールは様々な音韻連続を含む単語を学習データとし、これから各音韻の口形と音韻から音韻への過渡状態の口形変化とを視察することによって経験的に定められている。

この方式を用いて知的符号化を実現するためには、音声あるいは画像から音韻記号を抽出できる必要がある。しかし、知的インタフェースの手段としては、即応用が可能である。

2. 音声からの表情合成

3 次元ワイヤフレームモデルを音声信号だけで駆動して自動的に表情を合成する。即ち、音声信号をベクトル量子化し、そのインデックスを決定すると、受信側ではこのインデックスに対応した 3 次元モデル変形パラメータが用意されており、これに従って顔画像を変形する。これは、音声認識等の複雑な処理を必要としないので、容易に実時間処理が可能な方式である。

3. その他

上記の表情合成では、口周辺以外の動きについてはテキストや音声の内容とは無関係に合成されている。より自然な動きを合成するため、感情情報の抽出とそれによる表情合成が課題である。

7.2.11 ハイパーメディアシステム構成のためのデータ構造とその応用に関する一考察

(1) 全体概要

本論文では、様々なメディアに対処できるハイパーメディア・システムのためのデータ構造、汎用性のあるシステム・アーキテクチャを作ることを目的として、3種類のデータ構造、時間軸の概念、システム・アーキテクチャとしての階層構造を提案し、さらに映像編集システムへの適用例を示している。

今後のヒューマン・インタフェースではマルチメディアを容易に扱えることが必須であり、そのためにはハイパーメディア・システム等のマルチメディア・データベースを基本とするシステムが要求される。その中で、動画像等の時間情報を有するデータを計算機の中で如何に扱うかが重要な課題であり、本論文でとりあげているような時間軸の概念の導入に関する研究が益々盛んになろう。

(2) ポイント

1. 3種類のデータ構造

汎用性のあるシステム構築のためには、独立に扱える情報を分離し、それらを有機的に結び付けることができるデータ構造が必要である。そのため3種類のデータ構造、①コンテンツ構造：扱われるターゲット情報であるコンテンツを表現する構造、②ロジカル構造：ハイパーメディア・システムのノードに対応する情報の表現や処理とは無関係に存在する論理構造、③プレゼンテーション構造：具体的処理に関わる情報がマッピングされてできる構造であり、ロジカル構造からマッピングされる、を提案している。

2. 時間軸の概念

ダイナミックな構造規定や動画像を扱うためには、データ構造中に時間軸の概念を導入することが必要である。本論文での特徴は、①基準となる絶対時間軸を考え、どの要素もそれによって定義される、②グローバルとローカルな時間軸、③時間改像度の概念、④時間の概念は主としてプレゼンテーション構造で扱う、である。

3. システム・アーキテクチャー

汎用性（拡張性、柔軟性、可搬性、応用性）のあるシステム構成とするために、上述のデータ構造に基づいた階層的なシステムの積み上げを考えている。基本構成として、

データファイル部：データ入出力に関する一番低いレベルの処理

データ構造部：データ間の関係を記述したデータ構造（システムの機能を決定）

応用部：データ構造部を介してデータにアクセスするためのインタフェース

からなり、さらにデータ構造部を3階層に分けて、システムを構築するようになっている。

4. 応用システム例としての映像編集システム

ハイパーメディアの特徴を活かした映像編集システムを、上述のシステム構築法を用いて作成している。映画の基本ストーリーに従って撮られた素材のデータベース格納がロジカル構造に、編集後の表現手段も含めて得られる構造がプレゼンテーション構造に対応する。

参考文献

- [1] "Recent Progress in the Recognition of Objects from Range Data"
J. P. Brady, N. Nandhakumar and J. K. Aggarwal. The University of Texas at Austin
9th International Conference on Pattern Recognition, Vol. 1, pp.85 ~ 92, Nov. '88
- [2] "Chinese Character Recognition: A Twenty-Five-Year Retrospective"

- G. Nagy Rensselaer Polytechnic Institute
9th International Conference on Pattern Recognition, Vol. 1, pp. 163 ~ 167, Nov. '88
- [3] “顔画像計測に基づく視線感知法”
青山宏, 河越正弘 電子技術総合研究所
計測自動制御学会, 第5回パターン計測部会研究会資料, pp. 17 ~ 24, Sep. '88
- [4] “Physically Based Modeling for Vision and Graphics”
Michael Kass Apple Computer
NICOGRAPH '88 Proc., pp. S 3-1 ~ S-15, Nov. '88
- [5] “Flocks, Herds, and Schools: A Distributed Behavioral Model”
Craig W. Reynolds Symbolics Graphics Division
Computer Graphics, Vol. 21, No. 4, pp. 25 ~ 34, July '87
- [6] “Using Dynamic Analysis for Realistic Animation of Articulated Bodies”
Jane Wilhelms University of California
IEEE CG & A, June '87, pp. 12 ~ 27
- [7] “自然言語を解説し動画を自動生成するアニメーション・システム”
高島洋典, 島津秀雄, 友納正裕 日本電気 C & C情報研究所
日経CG, 1987年8月号, pp. 110 ~ 117
- [8] “画像データベース向き, 2階層符号化方式”
吉良健二, 井上誠喜 NHK放送技術研究所
テレビジョン学会技術報告 Vol. 8, No. 35, pp. 47-52, TEBS 100-9 (Dec. 1984)
- [9] “顔画像の分析合成符号化方式の動向”
原島 博 東京大学工学部
テレビジョン学会技術報告 Vol. 12, No. 9, pp. 19-24, ICS 88-17, IPA 88-11 (Feb. '88)
- [10] “画像と音声の知的インタラクティブ符号化の一検討”
森島 繁生 成蹊大学工学部
相沢 清春, 原島 博 東京大学工学部
1988年度画像符号化シンポジウム (PCSJ 88), 5-3, pp. 67-68 (1988)
- [11] “ハイパーメディアシステム構成のためのデータ構造とその応用に関する一考察”
外村 佳伸
信学技報, Vol. 88, No. 14, pp. 17 ~ 24 (IE 88-3), '88

7.3 自然言語

7.3.1 はじめに

人とマシンとが円滑にコミュニケーションをする上で、自然言語処理は大切なヒューマンインタフェースとなる。

ここでいう自然語源とは、プログラミング言語や形式言語と区別した、我々が普通に話している言葉を指している。

さて、自然言語の研究対象が何であるか、この言葉から得るイメージをはっきりさせるために、JICSTで、「自然言語」で始まるキーワードを抽出してみると、

自然言語理解、自然言語解析、自然言語処理、自然言語インターフェース
等と続く。また、自然言語に関連する研究分野は、

機械翻訳、人工知能、計算言語学、認知心理学、応答システム、知的情報処理
等である。このことから、自然言語の研究は、コンピュータの世界では、かなり古い研究対象であることが分かる[15]。

自然言語の研究は、文法理解の解明をスタートにして、意味理解の研究へと進んだ。'60年代の関心と期待は、統語解析にあったが、その限界に気が付いた'70年代は、意味解析の研究へとウェイトが移った。'70年代後半から、'80年代は、知識型の新しい言語理論が数多く提唱された[11]。

言語理論は、文の構文的規則を扱う文法理論〈文法性の問題〉と言語記号の解釈（付随した意味）を扱う意味論〈意味解釈の問題〉に分けて考える。

文法の目的は、文の統語と意味の特性から導いた文法規則を用いて、文字列に対して解釈を加えずに、文の構造を明らかにするところにある。実際には、後に述べる理由により、文法だけでは、文構造に曖昧さが残る。

意味論は、単語の意味に関する知識と文の背景である一般世界の知識を使って、文を解釈することを扱う。つまり、第一点は、品詞の並びとしては問題のない構文でも、意味の通じない構造があるので、文を正しく解釈することにある。二つ目は、発話する人の時空間により、言語的に等しくても、意味的には異なることを示すことにある。

ところで、構文解析と意味解析は、独立して順にこなすのではなく、両者を融合した解析が人間の方法に近いと考えられる。自然言語の解析で問題になる曖昧さは、多品詞性、係り受けの曖昧性、不明瞭さに起因するために、一つの表現が何通りにも解釈される。そこで、意味の情報を使って曖昧性を解消しようとする「統合型」の解析が提案されている。

以上の問題について、85年以降の文献をもとに、次節から順次のべる。

7.3.2 自然言語理解の仕組み

自然言語処理の流れを図示すると、単一文の理解過程は、図7.3-1のようになる。入力文が与えられると、形態素解析、構文解析、意味解析という一連の処理を行い、入力文の意味を抽出する。抽出した意味は、意味表現と呼ばれるデータ形式で記述される。この意味表現から得ると、理解・問題解決部は、目的に応じた処理をする。たとえば、応答システムであれば解答文を、機械翻訳であれば翻訳文を応答文生成部で処理する。この処理の各段階では、それぞれの処理に固有の知識を利用する[12]。

形態素分析では、文字列から形態素を抽出するが、構文情報や意味情報なしで、完全に行うことは不可能

である。一般には、形態素解析の結果は、曖昧さを含んだまま、次の解析に移る。

構文解析では、形態素列から句構造や依存構造を抽出する。構造解析の考え方は、採用する文法理論によって左右される。句構造とは、単語まとまりを示す構造である。依存構造とは、語句の修飾・被修飾関係を示す文法情報である。句の形や修飾関係は、単語の意味や文全体の意味を無視しては決定しえないので、句構造の曖昧性や多品詞性・多義性による曖昧性を排除できない。

意味解析の目的は、構造解析中に生じた曖昧性を単語の持つ意味を手がかりに解消することと、意味表現を作成することである。構文的には正しくても、意味的に整合性がとれなければ、正しい文といえない。意味解析が難しいのは意味を形成し難いことや、十分な解析をするためには大量の意味情報が必要であるのに、ケースバイケースで必要な情報のタイプが異なり、一貫したアルゴリズムがたて難いことにある[12]。

ところで、言語理解を行うに当たり、構文解析、意味解析、更には、文脈の解析を独立して処理することが、果たして妥当か問題がある。おそらく、人間は、これらを融合して言語解析を行っていると考えられる[11]。

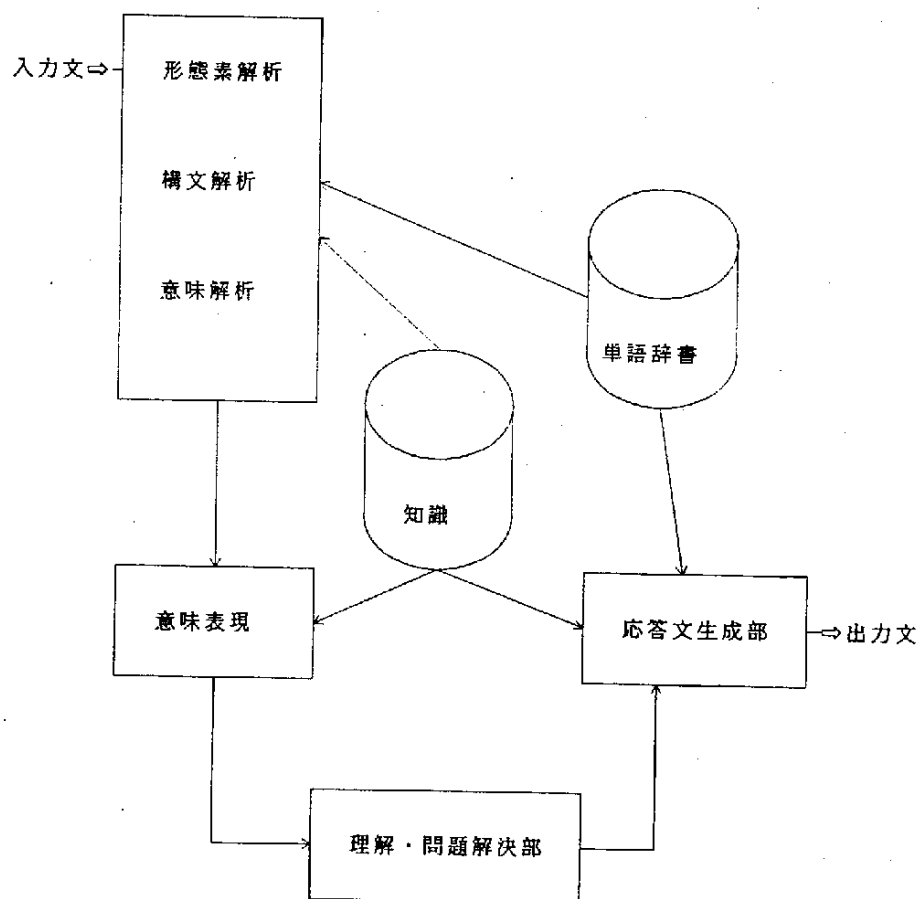


図 7.3-1 自然言語処理の流れ [12]

7.3.3 文法理論

文法は、自然言語の多様な表現を統合的に見出した規則である。言い換えると、文法とは現実の言語現象から科学的に抽出した規則の体系である。したがって、文法理論は、自然言語の一部をカバーしているに過ぎない。

現代の文法理論は、Chomskyの提案による生成文法を基礎にしている。この生成文法は、「再帰的」な構造

を内在しており、有限の単語と規則から文法的には完璧な文を無限に生成できるという考え方である。

この考え方にしがった文法理論は、言語学の立場から提唱された普遍文法と計算機処理の立場から提唱された単一化文法（句構造文法）の二つの流れがある。この二つの理論の相違は、「変形」を認めるか否かにある[11]。

変形文法理論では、深層構造で示している意味は、変形のやり方によって、多様な文の表現（表層構造）ができるとしている。つまり、同じ意味に対して変形のかけ方の違いにより、構文の構造は異なるというものである。チョムスキーが常にその中心にいる変形文法理論は、変化、発展して、現在、GB理論に結実している[8]。GB理論は、それまでの変形文法理論と比較して、言語の普遍性を重視している[10]。

非変形文法は、変形を用いずに言語現象を記述する。つまり、自然言語の文法とは、表層構造での部分構造の規則であると考えて、文全体は部分の積み重ねであるとする。この手法では、統語解析の計算量がかなり少なくてすむ。このことは、計算機処理の立場からは、かなり重要である。最近では、この考え方に基づいた単一化文法（ユニフィケーション）で総称される理論が、計算言語学の中心研究テーマである[27]。このうち、日本語句構造文法を基礎にした句構造文法が代表的である[8]。

句構造文法がGBと一線を画している点は、主辞と他の範疇との関係である。句構造文法は、主辞と他の範疇とを単一化するという概念である。単一化文法といわれる所似である。句構造文法での句構造木は、二分岐で記述することを前提にしている。単一化原理は、部分木の親子の間にのみ働く局所的制約を記述している。この局所性という文法記述は、句構造文法の特徴の一つである[7]。

二つに文法理論は、以上に述べた通りに大きな違いを持つ。句構造文法理論は、人間の言語は本質的に計算可能であり、その計算は、それほど複雑でないという立場をとる。

変形文法理論の立場は、言語というものは極度に複雑なものとする[5]。

(1) GB 理論

GB理論は、変形文法の伝統に則り、発話される文の構造を表す表層構造（S構造）と抽象構造である深層構造（D構造）を区別して説明する。S構造は、文の構造を表すが、言語表現そのものではない。時制や移動の痕跡を含む構造である。言語表現そのものは、音声形式と呼ばれている。また、意味表現に対応する構造は、論理形式と呼ぶ。（図7.3-2参照）

ところで、GB理論の語源である統率（Government）と束縛（Binding）がこの理論の中心である。

統率理論は、句構造における主辞と補語の関係を規定する。

束縛理論は、名詞や名詞語句が指示するものを決定するための理論である。名詞句を用いるときに、同じ個体を指示している先行詞の必要性や先行詞の位置を明らかにしなければならない。名詞句は、

照応表現　：代名詞との対応、名詞句の痕跡

代名詞類　：

指示的表現：固有名詞、明示的代名詞

に分類する[8]。

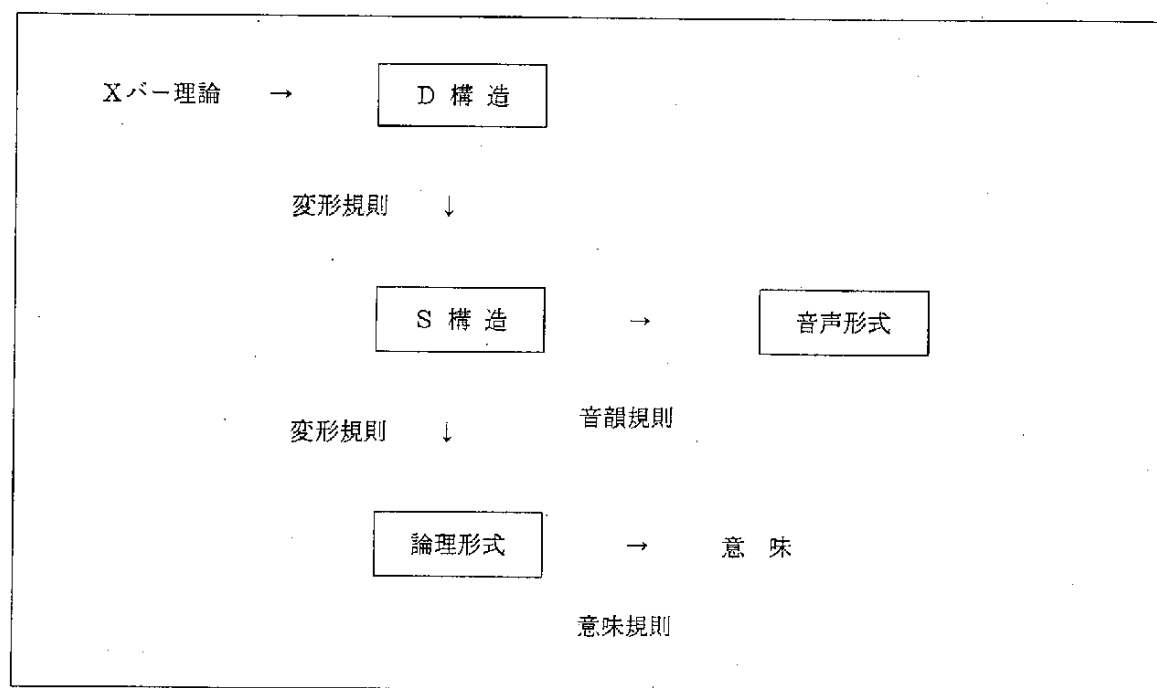
GB理論の重要な特徴の一つは、特定の構文のためだけの規則はないことである。たとえば、受動文、疑問

文といった規則はない。受動化は、目的語を主語にする移動操作をする。

この受動化は、普遍文法の諸特性から、規定される。

このような変形は、一応、すべて可能と想定した上で、うまくいかないのは何らかの「制約」が働いていると考える。制約とは、論理的には可能なはずの事を不可能にする文法上の概念である[10]。

GB理論の普遍文法の構成は、図7.3-2の通りになる[6]。



統率理論 境界理論 シータ基準
束縛理論 投射の原理

図7.3-2 GB理論と普遍文法の構成[6]

(2) 句構造文法

句構造文法では、出発記号と書き換え規則を用いて、出発記号に書き換え規則を有限回適用して得られる終端記号（単語や句）の並びを規定している。句構造文法の書き換え規則を制約すると表現力が低下するが、文解析の計算量はかなり少なくて済む[16]。つまり、句構造文法は、自然言語の表層構造での部分的構造の規則性として捉える。たとえば、図7.3-3において、枠で囲まれた部分構造の文法的適格性を規則にして、これらの部分の適格性を積み重ねて、句構造木全体の適格性とする[8]。

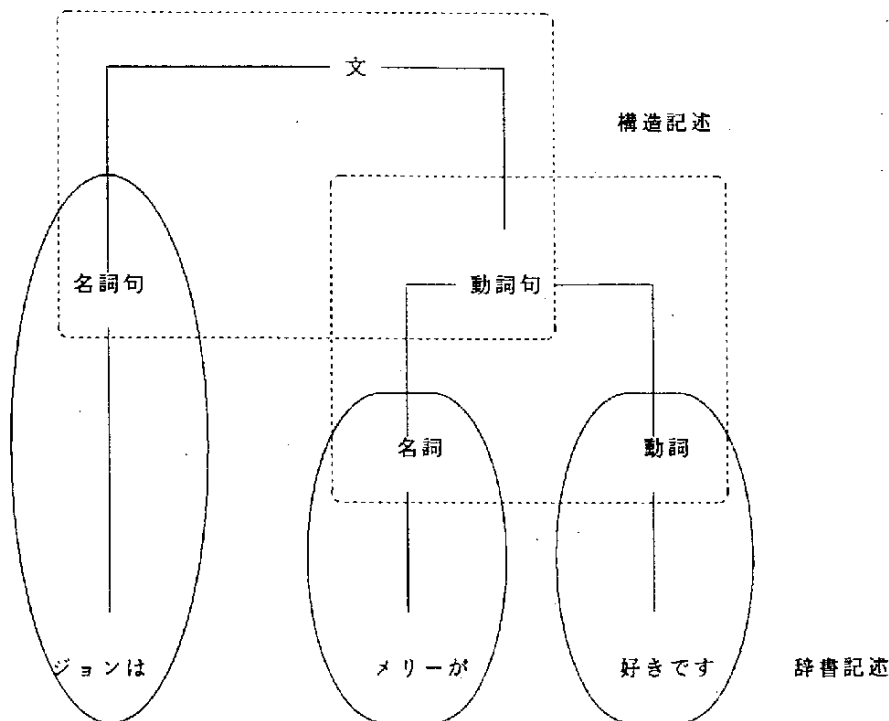


図 7.3-3 局所的構造 [8]

句構造規則による構文木は、二分岐で記述することを前提にしている。構文解析の進め方は、「トップダウン解析」と「ボトムアップ解析」に分けられる。また、各々の解析の途中で複数の句構造規則が適用できるとき、全ての可能な句構造規則を適用して解析を続ける方法を「模型」と呼び、その内の一つのみを適用して解析を続ける方法を「縦型」と呼ぶ。縦型では、適用しない他の句構造規則は一時保留しておき、後で、保留地点までバックトラックして戻る [8]。

この文法の一つである文脈自由文法では、句の整合性のみを判断する。図 7.3-4 に示すように、文脈自由文法は、句構造規則によって示される句の内部構成は、その句の外にある文要素に依存しないとする。つまり、文脈に関して制約がなく自由な規則記述であるという意味で、文脈自由文法と言う。この文脈自由文法は、句を単位として、句の組み合わせがより大きな句になるという書き換え規則で成り立つ [16]。

たとえば、

「私は、図書館で借りた本を、無くしてしまった。」

という文中の「図書館で借りた本」という名詞句の内部構成は、連体修飾句「図書館で借りた」と名詞句「本」とから構成されており、この事実、この句の外から構成がいかなるものであっても影響は受けない [18]。

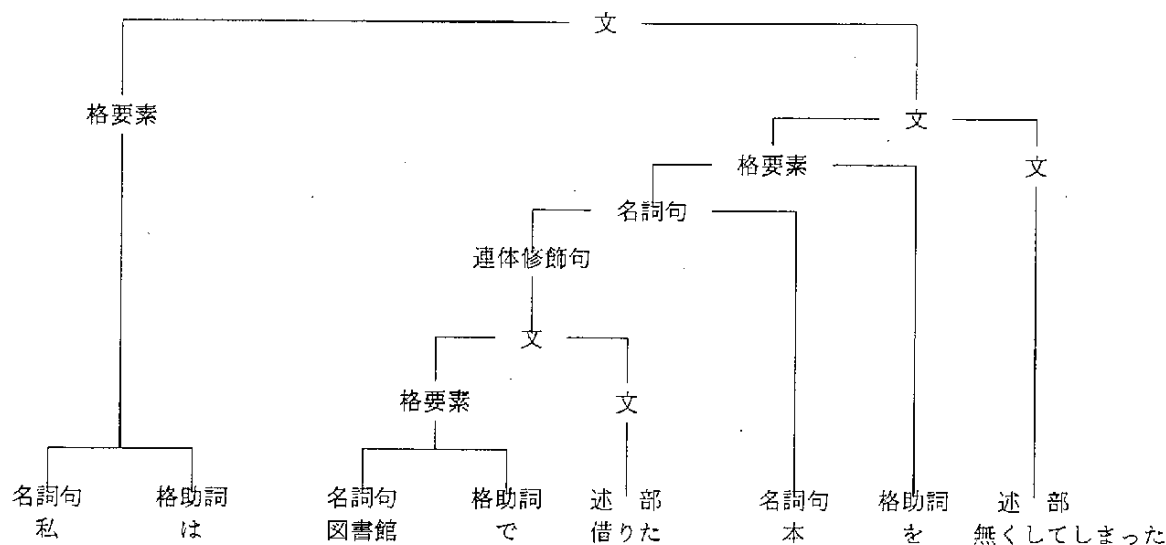


図 7.3-4 文脈自由文法による文解析 [16]

7.3.4 意味論

コンピュータで意味論を扱うには、意味を計算の対象にしなければならない。

また、文の理解においては、文脈情報や対象とする世界の知識を上手に利用することが大切であるかが分かっている。

ところで、従来の代表的な意味論には、論理文法的アプローチによるモンタギュー意味論がある [17]。モンタギュー意味論は、伝統的集合論に基礎を置いているために、実用と合わない面がある。これは、モンタギュー意味論でいう可能世界意味論が、全ての命題の真偽を分かっていることを要求しているからである。このモンタギュー意味論は、フレーゲの述語論理の計算言語学への拡張であるが、その限界が、以上のように指摘された [6], [25]。その批判の上に、状況意味論とメンタルスペース理論の二つの有力なアプローチが試みられている [3]。

さて、誰が、何時、何処で発話したか——その時空間により、文の解釈は異なる。そこで、図 7.3-5 にあるような概念の世界を、言葉の世界と切り離して考えてみる。

言語理解は、文脈を辿りながら、言語世界から概念世界へマッピングして、曖昧性のない概念表現を得ることにある。ここでいう言語世界での文脈依存性は、現実世界では状況に対応する。現実世界にある実在の解釈の仕方は、状況により変わる。たとえば、煙りは、「一つの状況」では火事の煙であり、「別の状況」では焚き火の煙であるというように、解釈は変わる。このように、語の多義性は、文脈により幾つもの意味を持つことができるので、言語表現の豊かさの源泉につながると共に、文の効率を高めている [3], [9]。

ちなみに、モンタギュー意味論は、一階述語論理を用いて言語表現の真理値を計算する。この方法は、ある対象世界の全てを考慮しておかなければ真理値を割り当てられない。従って、実用的な自然言語処理システムを作れない [5]。

また、フレーゲは、文の意味とは思想命題を与えることで、文の外延（指示対象）が真理値を与えるとした。しかも、意味と指示対象は、一対一の対応を示さなかった。フレーゲの示した例は、

「明けの明星」と「宵の明星」は意味は異なるが、指示対象は金星で同じだ。
 というものである [21]。

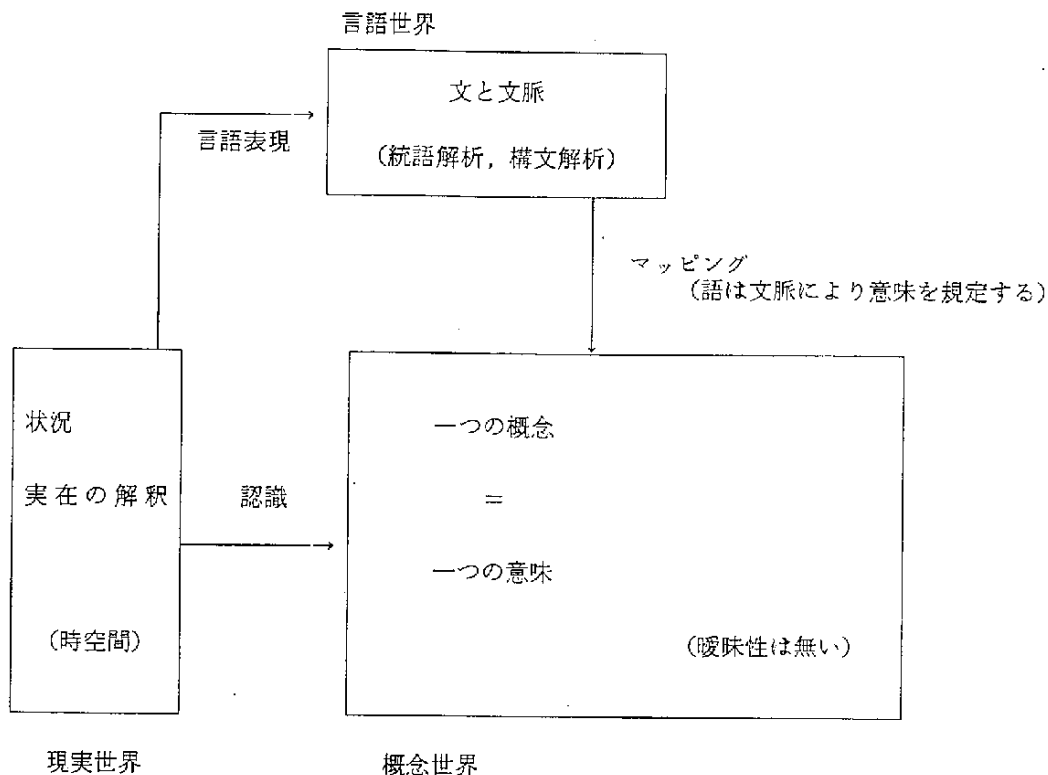


図 7.3-5 言葉を理解するための三つの世界 [3]

(1) 状況意味論

状況意味論は、スタンフォード大学の Barwise と Pelly らによって提唱された新しい意味論である。Barwise は、集合論や帰納的関数論に貢献のある数学者であり、Pelly は、人格同一性や指示詞の問題の哲学者である。スタンフォード大学の言語情報研究センター (CSLI) を中心に研究が進んでいる [9]。

状況意味論のポイントの第一は、「言葉の効果」という概念である。誰が、何時、何処で、発話したかにより、文の解釈が異なる。このような文を、「効果のある文」という。自然言語のほとんどは、効果のある文なので、状況意味論は、強力な方法である [2]。

このように、状況意味論は、自然言語の持つ文脈依存性を重視し、言語の持つ情報に重点を置く。状況意味論では、文の意味は、言語表現の情報内容と言語外の要因（環境）との関係と考えられる [3]。

別の言葉でいうと、状況意味論における文の意味は状況間の関係であるという。

この状況意味論のアイデアの内容は、

- ・ 談話の状況
パラメータ〈話し手、聞き手、発話文、時空間〉
- ・ 結合の状況
表現の指示の対応〈指示代名詞の照応、時制表現の時間〉
- ・ 記述の状況
話し手の記述の状況

状況意味論ポイントの第二は、「状況の理論」という概念である。

状況の理論では、言語によって記述できる状況は、現実世界の一部でしかないが、その一部の命題の真偽が分かれば、充分であるとする [6]。

つまり、状況意味論は、一つの実在（可能世界とも言う）しか認めない立場をとる。そして、図7.3-6の様に、一つの状況は、情報が少なくても真偽が分からなくても許されるとする。この長所は、

「私は、一匹の犬をみた。」

を意味解析するとき、部分的状況に一匹の犬が存在するので、真であり、より広い状況に犬が二匹存在しても一向に構わない[25]。

このように、状況意味論では、モンタギュー意味論と異なり、部分的な意味を計算することができる[3]。

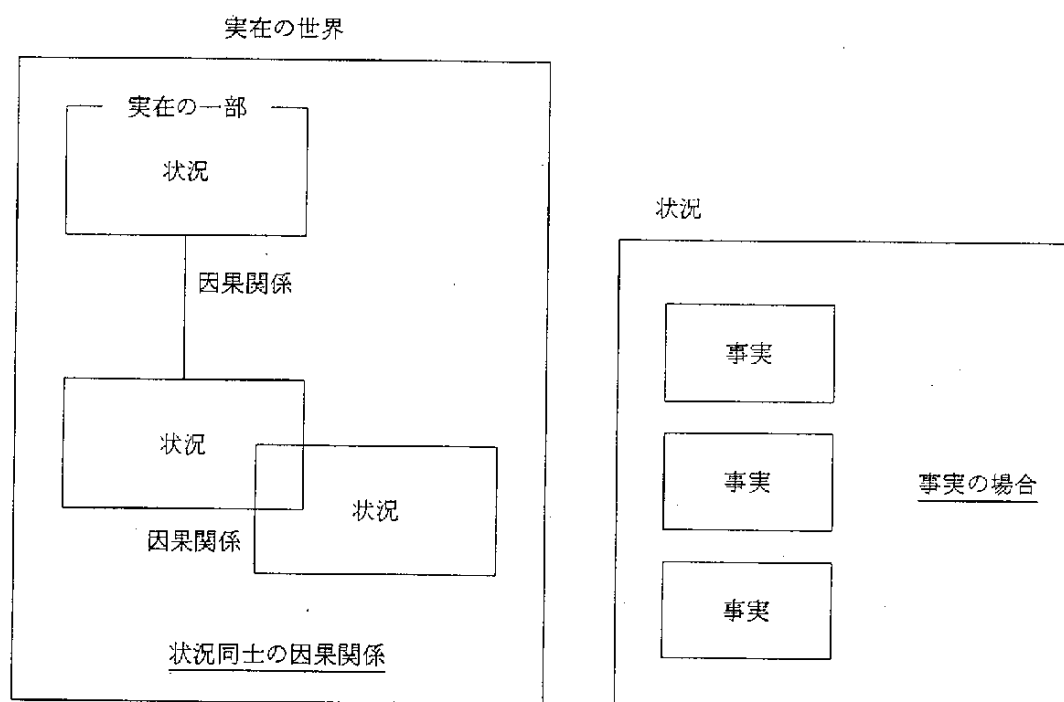


図 7.3-6 状況意味論の世界観 [25]

(2) メンタル・スペース理論

メンタル・スペース理論は、ジル・フォコニエ (Gilles Fauconnier) により提唱された[21]。これは、「言語によって心の中に作り出されるものは何か」を問題にしており、これをメンタル・スペースと名付けた。言語表現は、意味解釈において直接、世界と対応づけられるものではなくて、まず、メンタル・スペースを作り出すと考える。つまり、言語は、自分が作り出す構築物に関係づけられる[14]。メンタル・スペースは、言語理解において文法構造と外的世界の間を仲介する知的インタフェースである。

メンタル・スペースは、要素と、要素間の関係とを持つ増加可能集合である。このメンタル・スペースは、一般には話し手の「信念」を起点スペースとする。談話の進行に従って、メンタル・スペースは変化する[21]。

メンタル・スペース理論の構造は、二つのスペースを基本に仮定する。これを、トリガー・スペースとターゲット・スペースと呼びコネクタで結合する。この一方を信念スペースとして、他方を現実スペースとする例を考える。

「メリーの絵の花の色は黄色であると、マックスは信じている。」

ここで、「信じている」は、信念スペースを導入するキーワードとなる。このとき、図7.3-7に示す通りに、メンタル・スペース理論は、現実の色と信じている色が異なっても問題ないことを合理的に説明す

る [3]。

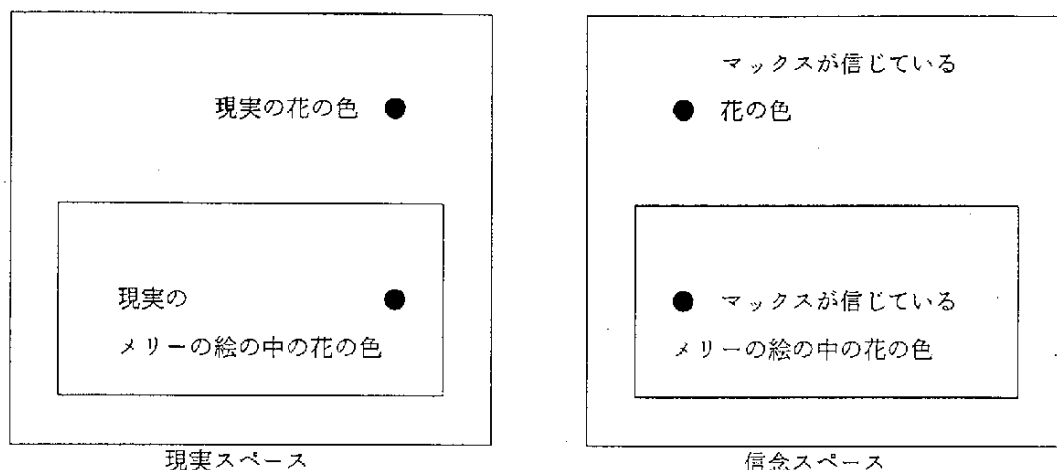


図 7.3-7 メンタル・スペース理論の構造 [3]

8.3.5 曖昧性の解決

文の役割と機能は、一言でいうと、単語の意味と単語の係り受けの関係を伝達することである [12]。さて、自然言語の解析で問題になる曖昧さは、構文構造の曖昧性や意思の曖昧性、不明瞭さに起因する。とくに、係り受けの曖昧性のために、文は一つの表現が何通りにも解釈される [16]。

構文構造における曖昧さは、意味や文脈の情報を使わないと解消できない。構文と意味の二つの解析を並列的に使う協調型解析手法の提案がある。たとえば、構文構造や意味構造を、ある時点で唯一に確定できない時、幾つかの可能性を残しながら処理を続けて、徐々に絞り込んでゆく [1]。

(1) 係り受けの曖昧性

文の聞き手は、単語の意味を推定しながら、係り受けにしたがって接続することで、文全体の意味を再現できる。つまり、係り受けの曖昧性のために、文は一つの表現が何通りにも解釈されることになる。たとえば、「東京と千葉の一部」というとき、曖昧さがある。「東京と（千葉の一部）」なのか、「（東京と千葉）の一部」なのか判然としない。

「東京と千葉の人口」のときは、意味的な情報が曖昧性を解釈できる。これは、「（東京の人口）と（千葉の人口）」という意味であることは明確である。

「太陽と地球の表面」の時は、後に続く言葉により解釈が異なる。通常は、「（太陽の表面）と（地球の表面）」と解釈できる。しかし、この後に「の関係」と続けると、むしろ、「（太陽）と（地球の表面）」とし解釈するほうが、自然である [16]。

その他の曖昧に関する問題も含めて整理すると、

- ・ 多義語の意味の推定
- ・ 多品詞の品詞の推定
- ・ 係り受けの推定
- ・ 代名詞照応
- ・ 省略語句の推定

・同音意義語の推定

等があげられる [12]。

英語の自然言語の構造的曖昧性のタイプを G. Hirst は表7.3-1 のように分類している [16]。

表 7.3-1 Hirst による英語の自然言語の構造的曖昧性のタイプ [16]

a. 付加の曖昧性	
前置詞句	—— 名詞か動詞か
前置詞句	—— どの名詞にかかるのか
前置詞句	—— 動詞に対してか形容詞に対して
前置詞句	—— どの動詞にかかるのか
関係節	—— どの名詞にかかるのか
副詞	—— 動詞に対してか、それとも文にしてか
分詞	—— 表層の主語か文か
b. 解析的の曖昧性	
関係節か補語か	
前置詞句か複合動詞句の一部か	
前置詞句か形容詞の一部か	
現在分詞か形容詞か	
現在分詞か動名詞か	
名詞句がどこで終るか	
名詞句の内部構造	
補語の主語	
述語の内部構造	

(2) 文脈的曖昧性

ところで、文を孤立状況で使用することは非現実的であり、文より大きな単位の中で前後の文脈を考慮しなければならない。文脈理解では、照応解析、新旧情報、語順、焦点解析などの文脈情報を得るためには、推論が重要である。つまり、自然言語理解における推論の役割は重要である。知識ベースを用いて自然言語が含む曖昧性を解消し、更に、「行間を読む」推論をする [3]。

文脈的曖昧性に関する有名な例を次にあげる。

「ぼくは、うなぎだ。」

食堂で喋れば、「ぼくが食べたいのは、うなぎだ。」という意味になる。

この曖昧性を解消するには、前後の文からの情報を用いる必要がある [4]。

(3) 統合的解析

言語解析では、統語（構文）解析・意味解析を、独立して、順にこなしてゆくのではなく、二者を融合した解析が、人間に近いと考えられる。このことは、どの程度に統語解析が済んでから、意味解析を行うということになる。

この統語と意味の解析処理の二者を組み合わせる方法として、

- ・一つの規則記述で表して、全てを同時に検査する「ユニフィケーション文法」 [22]。
- ・共通記憶域を介して更新する「black board [19]。

- ・統語解析の段階では曖昧な構造を残す「Muの文法」[13]。
- ・文の断片に意味解析を施す「早期意味解析」[26]。
- ・構文構造や意味構造を、ある時点で唯一に確定できない時、幾つかの可能性を残しながら処理を続けて、徐々に絞り込んでゆく「並列型」[1]，[24]。

等がある。

7.3.6 あとがき

自然言語の課題は多様である。ここで触れなかった課題について、簡単に触れておく。

- ・自然言語処理の形態素解析においては、未知語の存在を前提にしなければならない。特に、専門用語は未知語（未定義語）であることが少なくない。毎年、数十万語という規模で作られているという。いかに辞書を整備しても未知語を無くすことはできないので、未知語処理の不十分な自然言語処理システムは実用にならない[28]。
- ・文の意味表現を理解するためには、知識が必要である。知識表現法は、人工知能関連文献に詳しい。これは、論理表現とグラフ表現に分けられ、それぞれ特徴がある[23]。
- ・出力文の合成では、文脈、環境に応じた適切な文を合成する必要がある。敬体、代名詞化、省略語など、出力文でこなすべき課題は多い[11]。
- ・電子辞書に関しては、特定の文法理論に依存しない辞書を作成できるか、また単語の言語特性と単語が示す物の情報とをどのように記述するかなどの課題がある[19]。

参考文献

- [1] 雨宮真人他：文脈自由文法の並列構文解析アルゴリズム，情報処理 全国大会（前期）（1989）
- [2] 安食俊宏，溝口文雄：句構造文法と状況意味論との融合の試み，in 22）。
- [3] 石崎 俊：コンピュータで文間を補う，数理科学，3月，サイエンス社（1989）
- [4] 石崎 俊：文脈処理技術，情報処理，Vol.27, No.8（1986）
- [5] 郡司隆男：言語理論の最近の動向，数理科学，3月，サイエンス社（1989）
- [6] 郡司隆男：自然言語の文法理論，産業図書（1987）
- [7] 郡司隆男：句構造文法，情報処理，Vol.27, No.8（1986）
- [8] 白井英俊：自然言語の文法理論，in 11）。
- [9] 白井英俊：状況意味論の立場から，情報処理，Vol.27, No.8（1986）
- [10] Sells：Lectures on Contemporary Syntactic Theories, CSLI, Stanford Univ. (1985).
- 郡司隆男（他）：現代の文法論，産業図書（1986）。
- [11] 田中穂積，辻井潤一（共編）：自然言語理解，オーム社（1988）。
- [12] 高木 朗，伊東幸宏：自然言語の処理，丸善（1987）。
- [13] 辻井潤一：機械翻訳のための文法とその問題点，第一回「大学と科学」シンポジウム，「日本語の特性と機械翻訳」予稿集（1987）
- [14] 長尾 真：まえがき，in 21）。
- [15] 長尾 真：計算言語学の歴史と展望，情報処理，Vol.27, No.8（1986）
- [16] 西田豊明：自然言語処理入門，オーム社（1988）。
- [17] 西田豊明：意味の表現と処理，in 11）。
- [18] 野村浩郷：自然言語処理の基礎技術，コロナ社（1988）。
- [19] 辻井潤一：今後展望，in 11）。
- [20] Fauconnier：Espaces Mentaux

- 坂原 茂（他訳）：メンタル・スペース，白水社（1987）
- [21] 古川康一，溝口文雄（共編）：自然言語の基礎理論，共立出版（1986）
- [22] 松本裕治：知識表現，情報処理，Vol.27, No.8（1986）.
- [23] 溝口文雄他：並列論理型言語による意味解析，情報処理全国大会（前期）（1989）.
- [24] 向井国昭：状況意味論，in 22）.
- [25] Mellish：Computer Interpretation of Natural Language Descriptions, Ellis Horwood, England（1985）.
田中穂積訳：自然言語意味理解の基礎，サイエンス社（1987）
- [26] ユニフィケーション文法，in 22）.
- [27] 吉田 将：自然言語処理の必然と課題，電子情報通信学会，Vol.70, No.9（1987）.

本書は、日本自転車振興会から競輪収益の一部である機械工業
振興資金の補助を受けて作成したものである。

—— 禁 無 断 転 載 ——

情報機器のパーソナル化の技術動向
に関する調査研究資料

発 行 平成元年3月
発行者 財団法人 日本情報処理開発協会
〒105 東京都港区芝公園3丁目5番8号
電話 (03) 432-9371
財団法人 パーソナル情報環境協会
〒105 東京都港区虎ノ門1丁目17番1号
電話 (03) 507-0791
印 刷 日青工業株式会社
〒105 東京都港区西新橋2丁目5番10号
電話 (03) 501-5151

