

03-R 010

# 新情報処理技術に関する総合的調査研究 報告書

平成 4 年 3 月

JIPDEC

財団法人 日本情報処理開発協会



この報告書は、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平成3年度に実施した「新情  
報処理技術に関する総合的調査研究」の成果をとりまとめたも  
のであります。





## 序

過去30余年にわたる情報処理技術の進歩は著しく、その成果を支えに社会の情報化も急速に進展し、コンピュータは今やわれわれの社会活動にとって不可欠のものとなった。

しかしながら、その利用範囲の拡大や多様化にともない、情報処理機能の一層の高度化が求められ、特に最近では、固定化した逐次処理型の情報処理機能から、人間にとって親和性に富む、より柔軟性の高い知的情報処理機能への変革が、強く期待されるようになった。

このため、従来のコンピュータが不得手とする、人間の脳が行っているような、高度かつ多彩な情報処理の実現に向けて、必要とされる機能、計算原理及び実装技術、さらには、革新的な情報処理技術と社会の係わり等を含めた総合的な調査研究を、平成元年度から開始した。

3年度は、通商産業省における新情報処理技術開発（四次元コンピュータ）フィージビリティ調査研究委員会の下部機構として、当協会に「NIP Tワークショップ実行委員会」を設置して、21世紀を目指す革新的な情報処理技術のイメージを具体化するための調査研究を行うとともに、新情報処理技術推進のための基本計画の策定、新機構の組織や運営方針等についても検討を行った。

最後に、本調査研究に多大なご協力を頂いた委員各位に厚く御礼申し上げる次第である。

平成4年3月

財団法人 日本情報処理開発協会



## N I P T ワークショップ実行委員会名簿

(順不同)

委員長	甘利 俊一	東京大学工学部計数工学科教授
副委員長	弓場 敏嗣	電子技術総合研究所情報7-キタチ部部長
委員	上坂 吉則	東京理科大学理工学部情報科学科教授
	大津 展之	電子技術総合研究所知能情報部長
	岡部 洋一	東京大学先端科学技術研究センター教授
	神谷 武志	東京大学工学部電子工学科教授
	佐藤 知正	東京大学先端科学技術研究センター教授
	島田 潤一	電子技術総合研究所光技術部長
	島田 俊夫	電子技術総合研究所情報7-キタチ部計算機方式研究室長
	古谷 立美	電子技術総合研究所情報7-キタチ部計算機構研究室長
	矢嶋 弘義	電子技術総合研究所光技術部光情報研究室長
	米澤 明憲	東京大学理学部情報科学科教授





## 本書の構成

第1部 リアルワールドコンピューティングプログラム基本計画書

第2部 資料編

A. 基本計画書概要

B. 基本計画書（英語版）

The Master Plan for the Real-World Computing Program



## 第1部

リアルワールドコンピューティングプログラム

基本計画書



## 目次

### 第1部 基本理念

1 研究開発の目的	2
2 背景	3
3 リアルワールドコンピューティングパラダイム	5
4 研究開発の内容	9
5 基本方針	11

### 第2部 研究開発計画

1 理論基盤	13
1.1 研究の枠組み	13
1.2 研究課題	14
1.2.1 情報の柔軟な表現	14
1.2.2 情報および処理モジュールの評価	15
1.2.3 柔軟な記憶と想起	15
1.2.4 情報の統合および処理モジュールの統合	16
1.2.5 学習と自己組織化	16
1.2.6 最適化の方法	17
1.3 研究スケジュール	17
2 応用のための新機能	18
2.1 研究の枠組み	18
2.2 研究課題	20
2.2.1 柔らかな認識と理解	20
2.2.2 柔らかな推論と問題解決	22
2.2.3 柔らかなヒューマンインタフェースとシミュレーション	23
2.2.4 柔らかな自律制御	24
2.3 研究スケジュール	24

<b>3</b>	<b>超並列システム</b>	<b>26</b>
3.1	研究の枠組み	26
3.2	研究課題	26
3.2.1	超並列アーキテクチャ	26
3.2.2	オペレーティングシステム	28
3.2.3	超並列システムにおける言語	29
3.2.4	システム開発・プログラミング環境	30
3.2.5	システムの評価	31
3.3	研究スケジュール	31
<b>4</b>	<b>ニューラルシステム</b>	<b>34</b>
4.1	研究の枠組み	34
4.2	研究課題	35
4.2.1	ニューラルモデル	35
4.2.2	ニューラルシステムのハードウェア	37
4.2.3	ニューラルシステムのソフトウェア	38
4.2.4	超並列システムとの統合	39
4.2.5	システムの評価	40
4.3	研究スケジュール	40
<b>5</b>	<b>光システム</b>	<b>41</b>
5.1	研究の枠組み	41
5.2	研究課題	41
5.2.1	光インタコネクション技術	41
5.2.2	光ニューラルシステム	42
5.2.3	光デジタルシステム	43
5.2.4	システム開発の環境	44
5.3	研究スケジュール	45

第 1 部：  
基本理念

## 1 研究開発の目的

リアルワールドコンピューティングプログラムは、21世紀の高度情報化社会において必要とされる情報技術の基盤を確立することを目的とする。

コンピュータは急速に発達し、大きな計算パワーを提供するようになった。しかしながら、例えばパターン認識、不完全情報に基づいた問題解決、学習能力といった多くの分野において、コンピュータの情報処理は依然として柔軟性に欠け、人間の足元にも及ばない。

21世紀の高度情報化社会において人間とコンピュータの高度な協調関係を生みだし、革新的かつ包括的な技術の創造を支援するためには、情報処理技術は、人間の情報処理能力により近く、実世界の多種多様な情報を処理できるようなものでなければならない。本研究プログラムのめざすものは、そのような先進的で柔軟な情報処理技術の基盤の確立である。

21世紀の情報システムは、単一の技術ではなく、多くのキー・テクノロジーに支えられてはじめて実現されるであろう。超並列コンピューティング、光コンピューティング、ニューラルコンピューティング、論理をベースとするコンピューティングなどが、そのような技術に含まれよう。これらの技術を柔軟に統合し、情報システムとして結実させてゆくことが望まれる。本研究プログラムは、これらの重要でありながら依然実現されていない技術への挑戦である。我々は、これらの先進的技術の理論基盤を確立し、その可能性を探るとともに、それらを統合するためのスタイルを与えることを試みる。また、技術の可能性と有用性を確かめるためにいくつかの実世界問題に対しての挑戦も試みる。我々は、ここで得られた知見を広く世界に公開することにより、人類の共有財産としての知識と富の発展に貢献することを目指す。

この基礎的かつ挑戦的な目標を達成するために、研究における国際的および学際的連携を促進することが極めて重要である。また、企業、国立研究機関、大学の間の共同研究を支援していくことも不可欠である。



## 2 背景

コンピュータと通信技術の発達に支えられて、情報化社会は、新たな財・サービスの提供、生産流通の合理化などの産業活動のみならず、国民生活水準の質的向上、地域の振興、さらには教育や文化の面でも革新的な変化をもたらすであろう。

特に、情報のネットワーク化が、企業における組織内ネットワークから、組織間さらに家庭・個人を結ぶネットワークへと進展するにつれて、情報処理の主体が企業から家庭・個人へと広がってゆく。その結果、処理されるべき情報が量的に増加するのみならず、情報の質すなわち、種類も多種多様なものになってゆき、このような高度な情報ネットワーク社会の基盤として、あらゆる人がネットワークの多様な情報資源をより高度な次元で容易に利用できるようになるための新しい技術が強く求められる。

こうした社会的背景、そしてまた様々な応用分野での技術的要求として、現在のコンピュータをより人間に近く、実世界環境において人と共存し協調できる能力を持ったものにしてゆくことが不可欠である。

歴史的な経緯を振り返ると、コンピュータは、技術的進歩に支えられて驚異的な発展を遂げてきた。まず、コンピュータは伝統的なフォン・ノイマン方式の路線に沿って、数値計算、文書処理、そしてデータベースの蓄積と検索の管理といった分野で発達した。というのは、それらの応用分野は明確なアルゴリズムを持ち、伝統的なコンピュータで処理するのに向いていたからである。

次の発展は、論理的（演繹的）推論といった人間の知的思考過程に光を当て、論理と記号を操作する方向であった。AIや知識工学の研究分野は、コンピュータに記号で表現された知識や推論規則を取り扱う能力を与えるのに貢献した。1981年、日本は新世代コンピュータへの新しいアプローチとして、第五世代コンピュータの概念を打ち出したが、それはコンピュータに強力な論理的推論能力を付与し、大規模知識情報処理への扉を開こうとするものであった。

今日、コンピュータは、解法のためのアルゴリズムが存在しプログラム言語で明確に記述可能な良設定問題を解く能力においては、人間をはるかに上回る。しかしながら、解くべき多くの問題が不良設定でありアルゴリズムとして記述困難な実世界において人間の行なっている柔軟な情報処理に比べると、コンピュータは依然として柔軟ではない。

従って、そうした実世界問題に対処し情報処理技術の新たな地平を開くためには、人間の直観あるいは記号下（subsymbolic）レベルの情報処理に光を当てて、人間のような柔らかな情報処理の基本的な原理を追究し、発達するハードウェア技術の基盤の上にそれらを新しい情報処理技術として具現化してゆくことが重要かつ不可欠である。

近年、人間の脳における情報処理の仕組みに関する知識の蓄積は著しく、また、21世紀初頭には、超並列的なシステムを可能にするハードウェア技術も利用可能になると

期待されている。こうした状況を踏まえて、我々は新たな情報処理のパラダイムとして「リアルワールドコンピューティング」のコンセプトを提唱し、人間の情報処理の持つ「実世界性」を情報処理システムに与えることを目指す（図 1.1 も参照）。

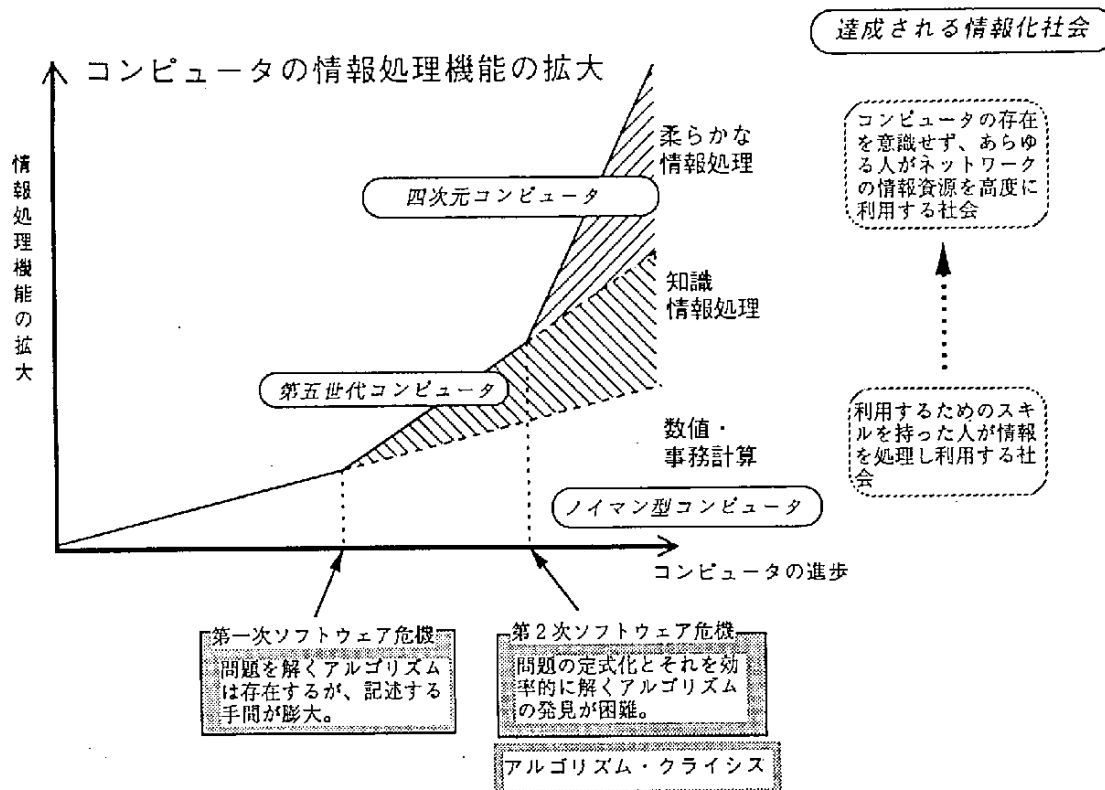


図 1.1: コンピューティングシステムの機能の拡張

21 世紀を、人間と情報処理システムがより密接かつ柔軟に協調するような世界規模で繁栄した情報ネットワーク社会にするために、新しい基礎的で包括的な技術の実現をめざすリアルワールドコンピューティングプログラムの実施が喫緊の課題である。

### 3 リアルワールドコンピューティングパラダイム

情報処理は、変化する実世界環境において生命が生存し続け、適応するための進化の過程を通じて獲得した機能である。また、周囲の状況を認識し、予測を行い、その予測に基づいて行動するといった、多面的な機能でもある。ここでは、このような機能を次のような2つのカテゴリーに分けて考える。

**A: 論理的情報処理**

**B: 直観的情報処理**

論理的情報処理は、意識的、分析的、直列集約的、デジタル的といった言葉で特徴づけられる。これに対して、直観的情報処理は、無意識的、総合的、並列分散的、そしてアナログ的といった言葉で特徴づけることができる。

基本的に、情報処理技術は人間のこうした2つの情報処理機能の両面を全体として、補完し、代替していくことを意図したものである。しかしながら、歴史的には、側面 A の機能において優れた機械がコンピュータとして実現され、ハードウェア技術の驚異的な発展に支えられて飛躍的な進歩を遂げてきた。一方、いわゆる「直観的情報処理能力」については、従来のコンピュータは人間の能力にはるかに劣っている。コンピュータの柔軟性の欠如もこの2つの側面の発展がバランスを欠いたことに起因する。

日常生活において、人間は曖昧性や不確定性を含む種々の情報を評価し、予測、計画、あるいは判断に必要な情報を変容する環境の中で獲得している。柔軟な情報処理機能を有したシステムを開発することは、より高度の情報処理機能を目指すパターン情報処理、知識情報処理、知能ロボットなど、今日の先端的情報処理の諸分野に共通した要請となっている。

こうしたシステムを開発するには、人間が行う情報処理能力の直観的側面の機械による実現を探究するとともに、前述の2つの側面を統合することが最も重要となる。統合の結果として、より人間的で、柔らかな情報システムの基盤となり得るような新しい情報処理のパラダイムに到達できる。そうした柔軟な情報処理を「リアルワールドコンピューティング」と呼ぶことにする（図 1.2 参照）。

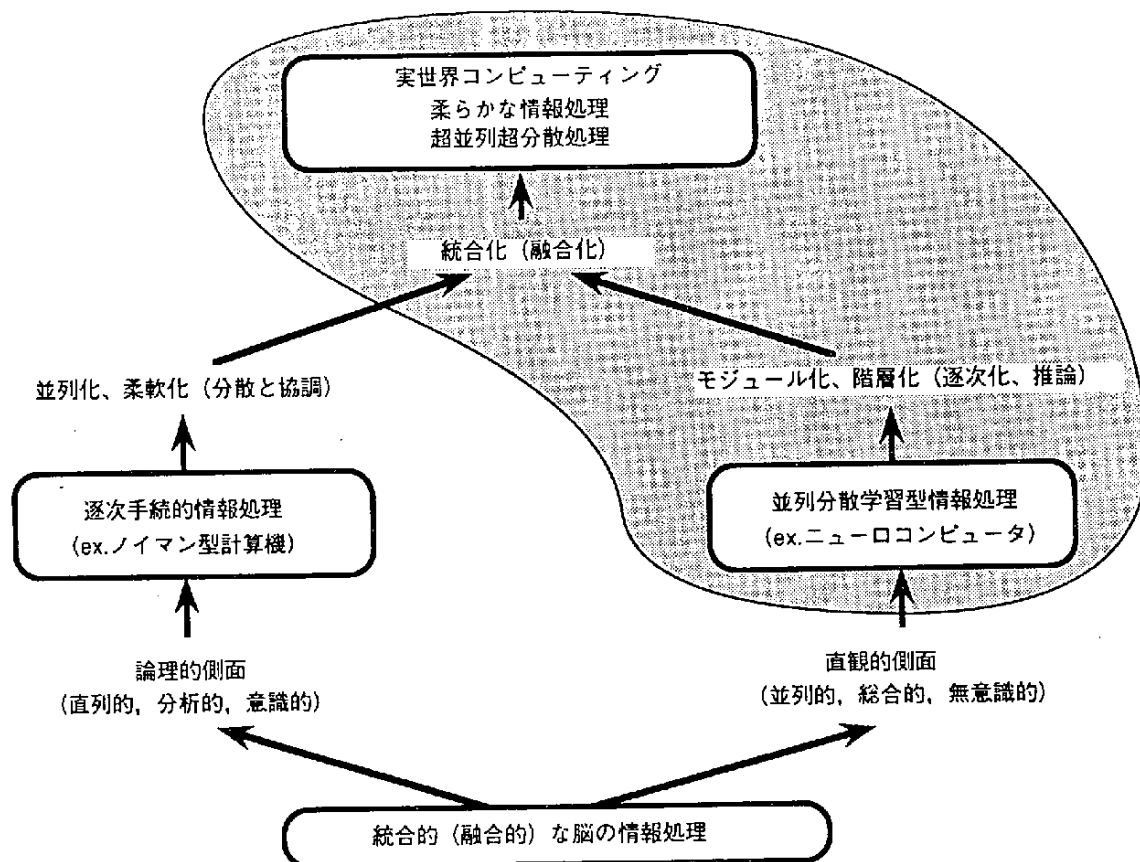


図 1.2：情報処理技術の2分化と統合

リアルワールドコンピューティングのシステムは、次のような特性によって特徴づけることができる：

- 開放性 (Openness)
- 頑健性 (Robustness)
- 実時間性 (Real-time)

「開放性」とは、予期していない状況に遭遇した場合にも、システムが自律的かつ適応的に自己を変化させたり拡張できることを意味する。「頑健性」は、入力情報が雑音で汚れていたり、環境が変化したりしてもシステムの挙動が頑強かつ安定していることを意味している。「実時間性」とは、システムが現実的な時間内に解を求めることを意味する。

人間がこのような実世界性を有した情報処理能力を維持できる理由は、我々の脳が、情報を分散的に表現し超並列的に処理する能力、学習および自己組織化能力、および情報統合能力を持つことに由来する。従って、上記のような特性を持つリアルワールドコンピューティングシステムを実現するためのキーコンセプトとして、次の二つがあげられるであろう。

- 柔らかな情報処理

情報の曖昧さや不確実さを許容するとともに、学習、適応能力を持った 情報処理

- 超並列超分散情報処理

相互に複雑に絡み合った多様な情報を超並列超分散的に処理する 情報処理

この二つのキーコンセプトを実体化するためにリアルワールドコンピューティングプログラムの研究開発要素としての技術項目を、以下のように3つ設定する：

- 柔らかな情報処理の理論

- 超並列超分散情報処理システム

- 実世界問題への応用のための新しい要素機能

これら研究開発項目を通じて、単に高速なコンピュータを開発することだけではなく、従来の論理的情報処理機能と新たな直観的情報処理機能を統合化することによって、たとえば、以下のような機能が実現されることが期待される。

- 不完全な情報や誤りを含む複雑に関連し合った情報を総合し、実時間に適当な（近似的に妥当な）判断や問題解決を行う機能
- 必要な情報や知識を能動的に獲得したり、具体例から一般的な知識を帰納的に修得する機能
- ユーザや使用環境の変化にシステムが自らを適応させる機能

リアルワールドコンピューティングの応用分野は、雑音を伴った環境下における状況を理解するといった不完全（不良設定）問題、社会や経済現象のシミュレーションといった大規模問題、仮想現実感を用いた情報環境の実現および自律ロボットの知的制御といっ

た実時間問題などの、「実世界問題」一般にわたる、非常に幅広いものとなることが期待できる。

#### 4 研究開発の内容

すでに述べたように、リアルワールドコンピューティングプログラムにおける研究開発の主要部分は、相互に関連し合う次の3つのカテゴリーに分けられる。

- 柔らかな情報処理の基本原理の探求による革新的技術の理論基盤の確立。
- システム基盤としての超並列システムの開発、運用。
- 広範な現実の応用システムを構築するための新しい要素機能の具体化。

これらのカテゴリー分けに対応して、(a) 応用のための新機能、(b) その理論基盤としての基礎理論、(c) (a)、(b)を支えるシステム計算基盤、という3層構造が、本研究プログラムの基本的な枠組みとなる（図 1.3 参照）。

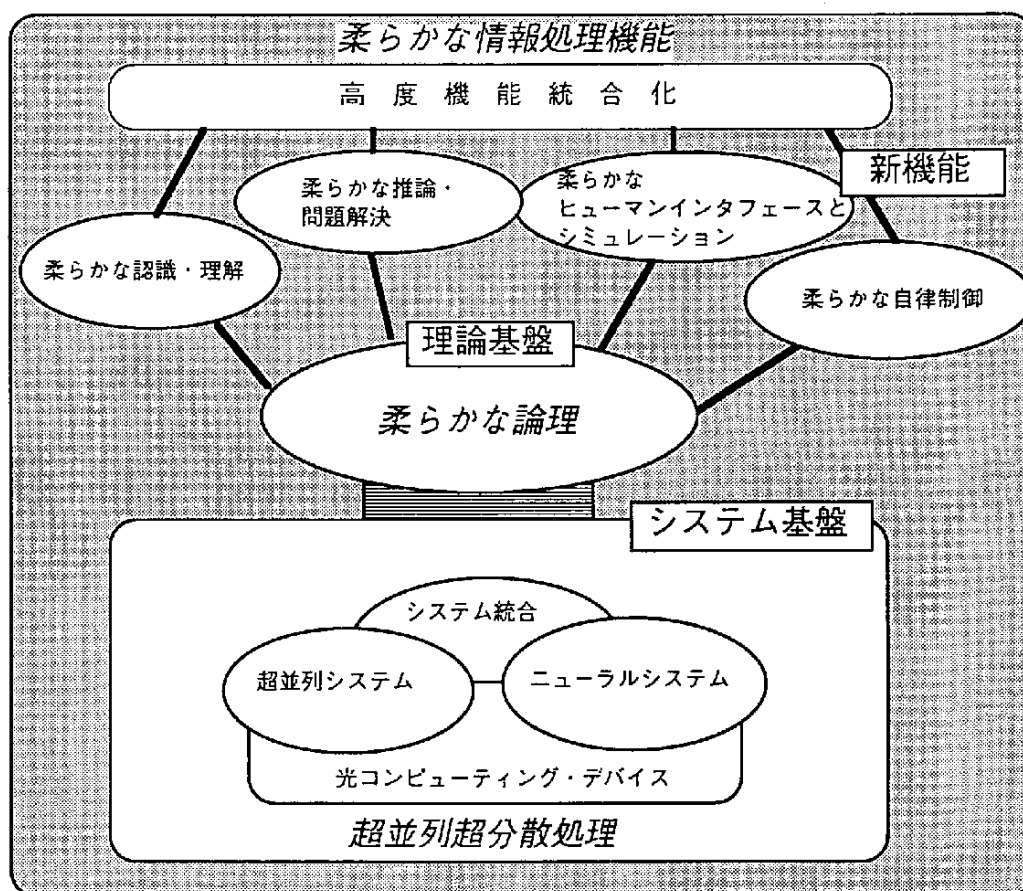


図 1.3：技術開発項目の構成

応用のための新機能に関する研究においては、広範な実世界問題を解決し実世界性を有したシステムを実現するために、共通的に重要な新しい要素機能が探求されることになる。ここでの研究課題は以下のように分類される：

- 多種多様な情報の柔らかな認識と理解
- 柔らかな情報ベースに基づいた柔らかな推論と問題解決
- 人間と機械の対話系のための柔らかな情報環境
- 柔軟かで自律的な制御

これらの要素技術を統合化する方法を探り、さらにその有効性を検証するために、いくつかの典型的な問題への応用を試みる。

理論基盤の研究開発目標は、柔らかな情報処理のための理論基盤を確立することである。このためには、まず、従来の情報処理の枠組みを次のような観点から拡張あるいは一般化することが重要である：

- 情報の表現、記憶、想起
- さまざまな情報の統合処理
- 情報の価値および要素モジュールの能力の評価
- 学習および自己組織化

特に、「多様な情報の統合」と「学習および自己組織化」は、最も基本的な項目であり、これらを超並列超分散情報処理システムの上で具体化する方法が、主なキーポイントとなるであろう。

リアルワールドコンピューティングでは、時空間的に分散した大量の情報を、それらの相互作用を配慮しながら高速に処理することが求められる。リアルワールドコンピューティングのシステム基盤として、いくつかの処理レベルで並列分散処理を利用できる計算システムの開発が不可欠である。このために、次に掲げる3つの視点から、研究開発を進めることが重要である：

- 汎用超並列システム
- 専用システムの1つとしてのニューラルシステム
- 光情報処理システム

同時に、これらの技術を統合する方法についても研究を進める。



## 5 基本方針

本研究プログラムの主目的は、単一のマシンを開発することではなく、重要でありながら未だ確立されていない要素技術の可能性を探求することである。この挑戦的かつきわめて基礎的な目的を達成するために、プログラムは、下記のような方針にのっとり推進されるべきである。

- 柔軟な実施体制：

本研究開発の推進に当たっては、統合性、象徴性、共通性の高い研究を集中研において、個別性、要素性の高い研究を分散研においてという研究課題の適切な配分を行ない、両者間の有機的で柔軟な連携を確保する。

- 競争原理の導入：

研究開発の前期では、競争原理（様々なアプローチ間の競合による研究の効果的な進展）を導入し、中間評価時にはその結果に基づいて研究開発課題を絞り込む。

- 学際性、国際性：

本研究開発の基礎的かつ挑戦的な目標を達成するため、学際的、国際的な連携を促進する。そのため、電総研や大学等の研究機関との共同研究を積極的に行ない、また、国内外の大学等の研究機関に対し再委託の公募等を行なう。

- 研究成果の公開性：

研究成果の公開性を保つため、研究開発の進捗状況および成果を国内外の会議等で報告、公開し、また、シンポジウムやワークショップ等を積極的に開催する。

- 研究インフラの整備：

以上の柔軟な実施体制と研究成果の公開性を支えるために、世界的な規模の分散研究基盤として、高速ネットワーク環境の整備を行なう。

第 2 部：  
研究開発計画

# 1 理論基盤

## 1.1 研究の枠組み

実世界の環境（人間を含む）は常に不確定性と変化に満ちている。しかしながら、われわれ人間は変化に富んだ環境の中で、不完全で曖昧さを含んだ情報を用いて、統合的に意志決定を行ない、きわめて柔軟に行動することができる。

このような人間における情報処理機能は、これまで計算機によって行なわれてきた仮想的な、あるいは限定的に設定された世界における完全な情報を用いた固い情報処理と対比的に、「柔らかな情報処理」あるいは「リアルワールドコンピューティング」として特徴づけられる。

リアルワールドコンピューティングプログラムにおける理論基盤の研究の目的は、人間の持つ柔軟な情報処理能力を情報処理の新しいパラダイムとして実現するための理論的基盤を確立することである。

これまでも柔軟で並列分散的な情報処理に関して各分野で多くの理論的研究がなされてきた。例えば、パターン認識、多変量データ解析、確率・統計的推論、ファジ理論、ニューラルネットワークによる計算・学習・自己組織化、そして正則化・最適化などに関する諸研究である。

柔軟な情報処理に理論的基盤を与えるには、このような研究をさらに深く押し進める必要があると同時に、これらの分野に潜在する共通な論理基盤として「柔らかな論理」（柔軟な確率的推論）の理論的枠組みを明らかにし、リアルワールドコンピューティングに対する統一的な理論的基礎を構築することが重要である。そこでは、学習および自己組織化機能と結びついた確率・統計的な対象や問題の定式化と非線形なダイナミクスによる情報処理の研究に力点が置かれることになる。

以上の目標のために、従来型の情報処理の枠組みを情報の表現・処理・評価などあらゆる点から拡張・一般化し、基礎理論を体系化する必要がある。

次に挙げる2点は共通的な最も重要な課題である：

- 多様な情報（および非階層的な情報処理モジュールの）の統合
- 学習と自己組織化（最適化と適応）

これらの課題を超並列超分散的な処理メカニズムの上でどのように実現するかということが、柔軟な認識、推論、制御と言った応用に向けての新機能を開発する上で一番のキー・ポイントとなると思われ、そのための理論基盤の研究が特に重要である。また、個々の項目の研究において、リアルワールドコンピューティングシステムが持つべき性質である「開放性」、「頑健性」、「実時間性」との関係が深く考慮されるべきである。

情報処理の新しいパラダイムを確立するために、自然から学ぶということも非常に重要であるだろう。このために、分野の壁を越えた研究交流のためのフォーラムにおける情報交換の促進などによって、人間の脳、生物の進化過程、生態系などに関する科学的な知見の利用を促進することが必要であろう。

理論基盤の研究課題は次のように分類されることが考えられる：

- 多種多様な情報の柔軟な表現
- 情報や処理モデルの評価
- 柔軟な記憶と想起
- 情報と処理モジュールの統合
- 学習および自己組織化
- 最適化技法

これらは、いろいろな応用に向けての新機能のための理論的基礎を提供するものと考えられる。

理論基盤の研究においては、一般的で統一的な基礎理論の構築をその主要目標とする。したがって、以下の研究課題のリストにおいては、すべての応用分野とすべての実現手法とに共通的な一般的な研究項目についてのみ述べる。それぞれの応用分野や超並列システム、ニューラルシステム、および光コンピューティングといった実現手法に個別に関わる理論・計算モデルは、それぞれの章において別に議論される。

## 1.2 研究課題

### 1.2.1 情報の柔軟な表現

視覚情報や音声信号などに代表される低次の感覚データから高次の記号的言語までの広汎な情報を統一的に扱い、散層的で柔軟なシステムを構築するためには、情報表現の柔軟なフレームワークを確立し開発することが重要な課題となる。このフレームワークではまた、情報の確からしさをも扱うことができる必要がある。さらに、いろいろな種類の情報を表現できるだけでなく、それを用いて連想記憶や学習あるいは自己組織化といった手続きを効率よく実現できるようにでなければならない。

重要な研究課題としては以下のようなものが考えられる：

- 多次元データ表に基づく情報表現
- 分散的かつ疎な情報表現

- 個別の処理ごとの適切な情報表現
- 確率的オートマトン、隠れマルコフモデル、ランダムマルコフ場、確率的決定リストなどを用いた知識表現
- 制約としての情報や知識の表現

### 1.2.2 情報および処理モジュールの評価

実世界との能動的な相互作用や学習、自己組織化などを実現するために、リアルワールドコンピューティングシステムは情報の価値を評価するための体系だった枠組みを持つ必要がある。情報統合のプロセスで用いられる正則化のための評価を行なう基準やシステムの能力の評価基準も合わせて必要である。

重要な研究課題としては以下のようなものが考えられる：

- 学習や自己組織化のための入／出力情報の評価
- 適切な情報の統合のための正則化条件の評価
- システムの信頼性と柔軟性の評価
- 処理モジュールの協調と統合のための評価
- 実環境との能動的な相互作用のための入力情報の評価

### 1.2.3 柔軟な記憶と想起

人間の記憶の高度に洗練された機能は、人間の柔軟な情報処理を実現する上での一つの鍵となっている。リアルワールドコンピューティングシステムは、柔軟な記憶機能と、連想的にさまざまな知識を取り出す機能を持つべきである。したがって、連想記憶の理論的解析と、柔軟な連想記憶のための新しい効率的なメカニズムの開発が必要になってくる。

重要な研究課題としては以下のようなものが考えられる：

- 確率的な推論による連想プロセス
- 構造的な類似性を用いた連想
- 時系列の連想記憶
- 非線形のダイナミックスを用いた動的な連想記憶

#### 1.2.4 情報の統合および処理モジュールの統合

推論、予測、計画のような情報処理はいろいろな情報や制約、知識の統合プロセスであると考えられる。このような情報統合プロセスを解析するという理論的研究を続けていくとともに、そのプロセスを制御する柔軟な方法を開発していくことも重要となるだろう。

多変量データ解析の手法は情報統合手法の一つの候補である。しかしながら、従来の手法の多くは線形変換に限定されたものであり、それらを非線形に拡張することが重要である。ニューラルネットワークモデルは非線形への拡張の一つの方向を与える。また、正則化理論も、さまざまな種類の情報を統合するための一つの理論的基礎を与えるであろう。

多数の処理モジュールによる協調的な処理も重要な問題である。これもまた統合化の問題と考えることができる。

こうした情報の統合については、従来それほど多くの研究が行なわれていないため、人間の認知過程における高次の情報の統合機能の研究も重要であろう。

重要な研究課題としては、以下のようなものが考えられる：

- 多変量データ解析やその非線形的拡張である ニューラルネットワーク・モデルを用いた情報の統合
- 正則化理論にもとづく多様な制約情報の統合
- 情報統合を制御するためのダイナミクスを含んだ制約論理システム
- 人間の認知過程における情報統合の研究

#### 1.2.5 学習と自己組織化

リアルワールドコンピューティングの理論基盤のうち、情報統合の研究とならんで最も重要なものの1つは学習と自己組織化の理論である。複雑で散層的なシステムを構築するためには、これら学習機能と自己組織化機能が重要な役割を担うはずである。記憶とデータベースの自己組織化機能も重要である。このような見地から、いろいろな概念や構造を学習する機構を探求するために、学習・自己組織化の計算理論を確立し、新しい効率的なアルゴリズムを開発することが研究目標となる。

重要な研究課題としては次のようなものが考えられる：

- 確率的な知識を学習するアルゴリズム
- 散層的なシステム構造を獲得するためのアルゴリズム

- 変化する環境の下での学習アルゴリズム
- 不確かな情報にもとづく学習アルゴリズム
- 質問による能動的な情報獲得を用いた学習アルゴリズム
- 既存知識と学習プロセスを統合する方法
- 学習のためのモデルの選択

### 1.2.6 最適化の方法

情報統合化のプロセスは本質的には最適化のプロセスとして定式化することができる。また、学習、自己組織化も処理方式の最適化プロセスと考えることができる。このような最適化問題を厳密に解くことは、通常、莫大な計算量を必要とする。この困難を乗り越えるためには、大規模並列システム上で効率的に実行可能な、近似的最適化技法を開発することが重要である。

重要な研究課題としては次のようなものがあげられる：

- 模擬焼きなまし法 (シミュレーテッド・アニーリング法)、遺伝的アルゴリズム、生態学的アルゴリズム、進化アルゴリズムなどの確率的最適化技法
- ニューラルネットワークを用いた最適化技法
- その他の非線形最適化の方法

## 1.3 研究スケジュール

理論的基礎の研究スケジュールは大きく分けて、2つのステップから成る。

研究開発の前期においては、並列性や柔軟性をもつ情報処理に関連して、理論とこれまでに提示されたモデル、例えばパターン認識、多変量データ解析、ニューラルネットワークによる情報処理などが情報表現や評価などの情報処理のフレームワークの拡張または一般化の中で再評価され、それらを根底となる「柔らかな論理」の下で統一的な観点から再構成し、さらに新たな展開を図る。同時に、情報統合や学習・自己組織化などの基礎的な機能に関する研究も行なう。

後期においては、これらの基礎理論を元に、柔軟な認識機能や推論・制御など要素的な新機能のための構成的な理論とモデルの研究を行なう。ここにおいては、新機能の研究とのインタラクションを強く保ちつつ研究開発を行なうことが重要である。また、新機能におけるシステム統合化の基盤となる、種々の要素機能モジュールを高次のレベルで統合するための基礎となる理論・モデルの研究も行なわれる。

## 2 応用のための新機能

### 2.1 研究の枠組み

リアルワールドコンピューティングシステムは、画像、音声、触覚などの多種類の情報を獲得し、理解や制御された行動の計画をたてるためにそれら情報を処理することによって、人間のさまざまな活動を支援する。ここで扱うべき情報は膨大であり多様である。さらには、実世界の情報はその本来の性質から不完全性と不確実性を含んでいる。

リアルワールドコンピューティングは様々な柔軟さを備えた新機能を求めるものである。頑健性、開放性、実時間といった側面は、柔軟さについての異なる特質を反映している。したがって、柔軟さをいかにして実現するかということが、新機能研究の基本課題である。

柔軟な情報処理は、従来の情報処理能力の限界を越えるものを目指している。期待される応用分野は、認識・理解システム、柔軟な情報(データ・知識)ベース、意志決定支援システム、柔軟な問題解決、やさしいヒューマンインタフェース、適応的大規模シミュレーション、自律協調制御などと極めて幅が広い。

研究開発の目的は、(理論基盤との接点を保って)新たな基本機能を各種の応用分野において探求することである。それらの機能はそれぞれの応用分野で働くことになるか、あるいは、新しい情報システムを実証するために統合されることになる。ここでいう新機能の優れている内容は、リアルワールドコンピューティングにふさわしい新たな概念をもつ理論とアルゴリズムによってもたらされうるものであり、それらは記号とパターンの統合、新しい頑健な機能を与える学習や自己組織によって従来考えられている情報処理の隘路を克服しているものである。単に、従来技術の組合せや、特殊目的のために安易にシステムを構築することは望まれない。

次の2点が統合システムの重要な方向と考えられる：

- リアルワールド適応自律システム、
- 情報統合対話システム。

前者の意味するものは、実世界との能動的な相互作用を通じ、環境の理解や制御が自律的にできる柔軟なシステムである。後者の意味するものは、人間とシステム間の拡大された情報チャネルを通じ、人間の能力を支援・強化する柔軟なシステムである。

システム実現のための新機能の研究開発は、以下のように分類されよう。

- 画像や音声のようなパターン情報や自然言語のようなシンボル情報といった多様な情報の柔軟な認識と理解、



- 実世界の情報の直接操作を許容し、学習能力や自己組織化能力を有する柔らかな情報ベースに基づく、柔らかな推論と問題解決、
- 人間と実世界間の相互対話を実現するための、仮想現実感を含む柔らかなヒューマンインタフェースとシミュレーション、
- 実世界環境と相互作用するための、柔らかでかつ自律的な制御と統合方法。

より具体的に、2つの統合システムと新機能の4つの分類の関係について論じよう。

第1のシステムは計算モジュール群と実世界との間で協調がなされるシステムを意味する。実世界における人間活動の部分的な代替が1つの目的である。このシステムは実世界を理解し、実世界の情報から必要な概念を学習し、新たな状況に適応する。さらに、このシステムは情報の獲得と望ましい状況の達成のため、実世界に対して制御行動を行なう。このためにシステムは不確実、不完全、そして変容する実世界の特徴を扱わなければならない。

第2のシステムは計算モジュール群と人間との間の協調を意味し、それは明らかに先のものとは別の協調システムになる。人間はさまざまな情報を扱うことができる。人間は様々な情報をつくり出すこともできるが、そこには人間の意図が、あいまいな表現ではあっても、込められてもいる。人間にとって必要となっている様々な問題解決や、新たな情報の創造を支援するために、計算モジュール群にはこれらの人間の意図を含む情報を柔軟に理解し統合することが要求される。ここでは、この計算モジュール群を含むシステムは人間の知的活動を拡大する強力なツールとみなされるべきである。

各システムは、現時点ではまだ明らかになっていないそのシステム固有の基礎的な新機能から構成されるであろう。また、各新機能は前述の4つの分類のいずれかに含まれる。新機能のいくつかについては双方のシステムにも適用可能であろう。

自律情報システムを実現するためには、以下のような機能が必要である。

- シーン画像の理解、
- 人間を含む実世界モデルと自律情報システム間の音声自然言語による質疑応答、
- 行動系列の計画と環境に適応するための最適制御、
- パターン認識、刻刻に変容する世界のモデル、制御メカニズムの統合。

情報環境システムを実現するには、以下のような機能が必要である。

- 人間の発する種々の情報からその意図を理解すること、

- 大量のデータベースから価値ある情報を検索・提示するための知的対話型支援機能の実現、
- 新たな情報を創造し、実世界の遷移状態を予測する知的シミュレーション、
- 人間の要素と実世界の計算モデルとの組合せるための統合方法。

基礎的な機能は、双方のシステムの実現にどれだけ貢献するかという視点と、各機能の属する分野でブレークスルーをなしているかという視点の両者から評価されるべきである。そのブレークスルーはリアルワールドコンピューティングの側面、すなわち、頑健性、開放性、実時間性というものと関連するであろう。研究開発プログラムは、個々の機能と統合機能の双方における新たなアルゴリズムの発見を試みる。以下、研究課題を列挙する。これらは、先に述べた2つのシステムの実現に必要なと思われる研究課題である。このことはここで列挙した全ての課題を研究プログラムが行なうことを意味しない。

これらの新機能を実現する計算基盤については、ニューラルシステム、光計算システムを含む超並列分散システムによるものが想定される。

## 2.2 研究課題

### 2.2.1 柔らかな認識と理解

#### 1. 柔らかな画像理解

情報が不十分であったり不完全であったりする場合の画像処理には、実世界における画像理解の新たなパラダイムが必要である。画像理解には定義不良、設定不良といった問題が存在する。多くの研究者が、究極的には分散されたパターン情報を凝縮された言語表現に変換する手段である認識と理解の技術の研究に努力を捧げてきた。それらは特定の対象に特性に合わせるという技術の研究であった。

画像認識・理解の初期段階(色、形、位置、動き)は、最適化問題として定式化されるモジュールとして明解に定義される。そして、制約充足パラダイムに基づいて統合されるべきである。画像情報の構造を把握するには、能動的な概念化が必要である。時には、構造化の結果を改善するために画像内に制約を導入することも必要である。この場合の画像処理は制約充足計算を含む。

画像理解の柔軟性は、動きを含む画像の分割に関する新たなアルゴリズムの発見と、可能ならば世界モデルを仮定したり他の情報を用いて画像を分割したオブジェクトを統合することによって生じる。

重要な研究課題を挙げる。

- 画像データの学習、自己組織化、分割のためのアルゴリズム、
- 画像理解のための画像特徴と概念的記号の統合、
- 画像理解に基づく、動的世界を取り扱う技術の構築、
- 非剛体の物体表現のための柔軟に変形可能なモデルに基づく、人間の意図理解のための人間についての動画像認識、
- 人間のもつ直観的情報を画像から抽出する基本機構。

## 2. 柔らかな音声理解

実現すべき柔らかな情報処理の中で、音声は重要な情報のひとつである。計算機が不特定話者の会話を理解できたならば、計算機ユーザにとって強力な情報となる。音声理解の柔軟さは、動的な音声の特徴抽出と、不完全な認識という制約下におけるマッチング手法を含む認識モジュールの統合という新しいアルゴリズムの発見から生じる。この新機能を得るためには、以下のような研究課題が考えられる。

- 初期段階の音声処理と音声の動的特徴の抽出、
- 未知単語の取扱が可能な理解機構、
- 超並列計算での実現に適した単純明解な認識アルゴリズム。

実世界における柔らかな発声・合成音声に基づく、コンピュータとユーザ間のインタフェース技術が開発されなければならない。以下のような研究課題が挙げられる。

- ノイズのある記号列（音声認識の単語出力）でも頑健な構文的、意味的处理、
- 音声駆動型会話モデルと音声会話と結合した動的世界モデル、
- 補完する特徴を持つような他の情報と音声との関係。

## 3. 柔らかな自然言語理解

実世界では、自然言語で表現される膨大な生データやコード化されたデータが生成される。しかしながら、さまざまな目的、状況下で使われるために、機械による理解処理のために変換されているものはほとんどない。この分野での新機能は、膨大なコード化された自然言語を生データとして管理し、汎用的な概念表現を抽出できる頑健な理解システムの実現に焦点を当てるべきである。

新機能の鍵となる概念は、概念的かつ構造的な分割と融合による分割単位に基づいて各種の知識表現が頑健になるような、自然言語の分割方法である。同時に、学習と自己組織化の機構も、それぞれがエージェントとみなされる更新された知識単位を得るために必要とされる。以下のような研究課題が挙げられる。

- 大規模かつ不完全な文に適用可能できる頑健な構文解析、
- 各種の電子化辞書の自然言語理解系への組み込み方法、
- 談話から得られる文系列のための自己組織化アルゴリズム、
- 談話を扱うための概念的構造的測度に基づく相互作用を通じ、知識単位を統合する計算モデル、
- 説明指向の自然言語理解、
- 自然言語と他の情報の間の連係と補完関係。

## 2.2.2 柔らかな推論と問題解決

### 1. 推論と問題解決

柔らかな推論は、実世界で生成される不整合かつ不完全なデータを扱う。大部分のデータは記号表現されているが、推論システムは確率や尤度のような数値的表現を使うこともある。この分野では、ユーザーの不完全な要求に基づいて問題の定式化が要求されるので、知識表現の問題は重要である。推論の柔らかさは、それぞれが特化した作業課題を実行する各推論エージェント間の協調メカニズムから生じる。

社会システムにおける生産システムを含む多くの大規模システムは、協調によって全体システムを安定化、準最適化、適応させるような、数多くの異なったエージェントが存在する。この大規模システムは実世界の一部であり、ここにおける問題解決も重要な新機能となる。さまざまな新機能がこの実世界に適用するためには、超並列かつ準一様なアーキテクチャの観点からみたこれらの機能を評価、解明する研究が必要となる。

以下のような研究課題が挙げられる。

- 準大域的な関係や制約の獲得と、よく定義されていない要求に基づく問題記述の枠組、
- 類推や確率的推論に基づく柔らかな問題解決アルゴリズム、
- 推論エンジンの多数のエージェントの間での協調方式、柔らかな推論を実現するための計算機と人間との協調方式、
- 通信が制約されている条件のもとでの大規模システムの問題解決。

### 2. 柔らかな情報ベースと検索

人間の日常生活は、画像、音声、文章、触覚などさまざまな情報に囲まれている。それらには生成され制御されているものもあるが、そうでないものもある。人間は、

異なる情報の統合や柔らかな推論機能を無意識に用いて、情報検索のために利用可能なデータを使っているように見える。我々の社会には、さまざまな情報がさまざまな場所に分散されているので、コンピュータネットワークによって実現するためには、この機構を見極めることが重要である。これらのデータは統合することによっていっそう価値が高まる。

以下の研究課題が挙げられる。

- 互いに独立あるいは共通な知識表現、
- 情報間の協調をとるための学習アルゴリズム、情報間のお互いの対応をとること、
- 各情報に適したデータ構造の表現と情報間の相互作用の定式化、
- 情報検索のための推論。

### 2.2.3 柔らかなヒューマンインタフェースとシミュレーション

#### 1. 柔らかな対話

リアルワールドコンピューティングシステムの優れた使用方法は、上述した個別新機能の開発と同時的に開発されるべきである。新しい情報処理環境は新しいコンピュータと幅広く知的な物理的協調活動を可能にするものとして理解される。このような環境では、ユーザとしての人間は身振り、表情、話し言葉のような自然な手段によって計算機と対話でき、また実時間の3次元画像によって情報を受け取ることができる。これらの情報環境を実現するための関連する研究課題には以下のものを含む。

- 広帯域の多様なインタフェース（身振り・表情、言語や画像のパターン認識など）、
- 仮想現実感を使った情報表示システム、
- 人間の意図を理解するための認知行動モデル。

#### 2. シミュレーション

リアルワールドコンピューティングは、新たなシミュレーション技術を実現し、それは非常に困難な問題を解くための強力なツールをユーザに提供する。ここでは、費用と時間を要するうえに、いつも正確な結果が得られるとは限らない物理的な実験を代替するものとして、大規模かつ複雑な系のシミュレーションを可能とし、時間の制約なしに系の挙動を予測できることが求められる。未知の現象や将来の予測

(例えば地球環境や天気予報など) は実時間コンピューティングの重要な応用分野である。関連する研究課題としては以下のものがあげられる。

- 学習・適応型シミュレーション、
- 複雑・カオス的な時系列の予測と制御。

#### 2.2.4 柔らかな自律制御

実世界において実時間で動作する柔らかな自律的協調システムを実現するために要求される技術の開発に焦点を当てた研究が必要である。ロボットはその典型的な例である。また、高齢者や、身体的に障害を持つ人々の支援も応用分野である。このようなシステムは認知、決定、行動のために互いに相互作用する多様な機能モジュールからなる。解くべき問題は、これらの多様な機能をいかに統合するかの方法であり、実世界の制約のもとで望まれるゴールを達成するために機能モジュール間の相互作用の制御をいかに行なうかの方法を明らかにすることである。実世界とは、不完全性、用いることのできる情報の曖昧性、物理的世界の動的変化、用いることのできる時間や空間の制限等があるもので、前記の統合機能はこれらを考慮したものではない。関連する研究課題には以下のものがある。

- 環境・作業課題・制御の柔らかなモデリング、
- 能動的分散計測とセンサ統合、
- 作業する場面での計画と分散協調探索、
- 物体を上手く実時間で操作するための複数の計測・計画・行動エージェントの構造化と調整、
- 動的な実世界と内部的な世界モデル間の無矛盾性の維持。

### 2.3 研究スケジュール

研究開発は2段階に分けられる。第一の段階では、統合システムに必要な個別新機能を実現するために、アルゴリズムの発見を目指す研究が行なわれる。柔らかな認識と理解では、指の動きやジェスチャなどを示す動画像の理解や会話音声の理解技術の確立がその研究道程となろう。柔らかな推論と問題解決では、画像を含む特定の領域の情報ベースをつくり、それに対して柔らかな推論や問題解決のための新しいアイディアの適用を行ない、それらの新しいアイディアの発展と革新を図る。新規性をもつアイディアにおいては確率推論や制約充足などの方法によるものも含まれるであろう。また、大規模デー

データベースを自己組織化する方法の発見もその道程になろう。柔らかなヒューマンインタフェースとシミュレーションでは、人間の意図とそれを表す多様な情報表現との関係を、心理学的証拠や実験と情報処理メカニズムと結びつけることで明らかにすることがなされよう。また、認識や理解の結果を用いてつくり出される時間とともに変化する状況を表現する仮想現実感による表示技術の確立もその研究道程となろう。柔らかな自律制御においては、観測、認知、計画、行動を実世界において行なうことを適応と学習の観点から統合する方法論の確立がその研究道程となろう。

第二の段階では、第一の段階で発見された個別新機能を統合して、統合システム、すなわち実世界適応自律システムと情報統合対話システムの実現に向かう研究が行なわれる。ここにおける一つの研究道標としては、実時間で動作するシステムの実現である。この研究開発で提案される種々の新機能群についてはそれらを統合するために生じる組合せ問題を知的にかつ洗練された方法で解決することがその研究道程になるであろう。そのとき、個々の新機能は統合システムの中で他のそれらと協調することを通じてシステムが頑健性と開放性をもつことに貢献するであろう。

第一の段階ではその大部分は個別新機能の発見を目指す、部分的には要素新機能の統合のための研究も行なわれる。同様に、第二の段階でもその大部分は統合システムの構築研究であるが、部分的には個別新機能の発見のための研究も行なわれる。

個別の研究活動の中では、リアルワールドコンピューティングにふさわしい新たなパラダイムの構築のための関連研究を刺激するようなアルゴリズムの発見活動（ただし、この段階では新たなアルゴリズムの理論的基盤は必ずしも要求されない）、新機能の実現方法と理論基盤に基づく実世界コンピューティング技術体系の構築活動、が行なわれる。

しかしながら、個々の研究課題はプログラムの開始5年後に見直しがなされるべきである。この見直しの機会には、本研究プログラムで開発される超並列計算機上での実装可能性やそれへの反映、あるいは、実世界コンピューティングに関連する個別の技術分野でいかなるブレークスルーをなしているか、のいずれかの観点から各新機能が評価されよう。

### 3 超並列システム

#### 3.1 研究の枠組み

リアルワールドコンピューティングには多種多様な情報を総合的かつ柔軟に処理する計算の枠組みが要求される。リアルワールドコンピューティングの応用例としては、柔らかな認識と理解、柔らかな推論と問題解決、柔らかなヒューマンインタフェースとシミュレーション、柔らかな自律制御などがある。これらの応用を実現するようなシステムは、数多くのモジュールにより構成され、モジュール間あるいはモジュール内という様々なレベルでの並列分散処理が行われる。

これらの応用に適用可能な、並列オブジェクト指向、データフロー、データ並列、ニューラルネットワーク、確率的情報処理、などといったいくつかの並列パラダイムが提案されてきた。リアルワールドコンピューティングは、このようなパラダイムの組み合わせにより実現されるであろう。リアルワールドコンピューティングに基づいたこのような柔らかな情報処理においては、さまざまなパラダイムに基づいた多くのレベルの並列性がある。これらの新しいパラダイムは超並列システムと自然に適合し、実世界の問題を実用的な時間内に解くために膨大な計算を必要とする。このことは、超並列システムがリアルワールドコンピューティングを可能とする計算パワーを実現するためには不可欠であることを示している。またこのような超並列システムは、複数パラダイムを効率良く実行するようなシステムである必要がある。

リアルワールドコンピューティングの計算基盤としての超並列システムはそれ自体が柔軟である必要がある。ハードウェアの柔軟性に関しては、耐故障性と信頼性が重要である。アーキテクチャの観点からは、超並列システムは種々のパラダイムを効率良く支援する汎用的なものでなければならない。ソフトウェアは適応性、自己組織性、最適化などの機能を持つ必要がある。超並列システムは最適な性能を得るため、それ自身が環境に適応し、ユーザには過大な要求をしない。このような観点から、本研究プログラムにおける超並列システムに関する研究では、アーキテクチャ、オペレーティングシステム、言語、そしてシステム環境の開発を対象とする。

#### 3.2 研究課題

##### 3.2.1 超並列アーキテクチャ

以下に示すものは、汎用の超並列システムを開発するために推進されるべき基礎技術である。

##### 1. モデル

複数のパラダイムをサポートする柔らかな実行モデルは、言語モデルとハードウェ



アの間隙を埋める能力が必要である。また仮想的なコンピュータを実要素プロセッサにマッピングすることを許容するような柔軟さが必要である。

## 2. アーキテクチャ

アーキテクチャは様々なパラダイムを効率良くサポートする必要がある。したがって特定の応用を指向したものではなく、汎用のアーキテクチャでなければならない。超並列システムの効率的な実現のためには、将来のデバイス技術や実装技術を考慮したハードウェアアーキテクチャを研究することも重要である。

## 3. 相互結合網

相互結合網は計算速度と同等の高速な通信を提供しなければならない。また、動的な負荷分散、全体的な同期、全体的な優先度制御もサポートする必要がある。高速な相互結合網はシリコン技術による実現だけでなく光技術についても研究されるべきである。

## 4. 頑健性・信頼性

超並列システムで予想される部分的な故障に耐えるようなハードウェア指向の頑健性が研究されなければならない。システムを構成する要素は、自己診断、自己修復機能を持つ必要がある。システム全体はメンテナンスのためのアーキテクチャや機能を持ち、システムの信頼性を維持しなければならない。

リアルワールドコンピューティングプログラムで開発されるシステムは汎用の超並列システムでなければならない。汎用性は複数のパラダイムをハードウェア実行モデルにソフトウェア的にマッピングすることで実現されよう。ハードウェア自体が種々のタイプのアプリケーションに対して適応する機能を持つことが可能ならば、そのハードウェアは汎用と考えることができるであろう。実世界コンピューティングにおいては、膨大な数のモジュールやプロセスが並列的に実行されるため、高速なプロセススイッチ機構と低コストな同期の機構をサポートするようなアーキテクチャがシステムとして統合的に実現されなければならない。

プログラムの前半において  $10^4$  要素プロセッサを持つプロトタイプシステムが開発される。いくつかのハードウェアプロトタイプを試作するが、これらのプロトタイプのうち1つはソフトウェア開発や新機能を検証するためのプラットフォームとして設計製作される。超並列モデルや超並列アーキテクチャに関する基礎研究も同時に行なわれる必要がある。プログラムの後半においては、 $10^6$  要素プロセッサを持つシステムの開発が期待される。このシステムは様々な種類のリアルワールドコンピューティングのアプリケーションが実時間で実行可能な性能を持つであろう。そのアーキテクチャはこのプロ

グラムにおいて研究される新しい超並列コンピューティングモデルに基づいたものになるであろう。

### 3.2.2 オペレーティングシステム

オペレーティングシステムの主要な目的は、計算資源を並列プロセスに対し効率良く配分すること、ハードウェアの詳細を知る必要のない、人に優しいインターフェースを構築することである。逐次計算機のオペレーティングシステムと同様、超並列オペレーティングシステムは、汎用、対話型、マルチユーザシステムでなければならない。超並列システムのオペレーティングシステムは、様々なプロセスが高いスループットで並行的に動作することを支援するよう設計されなければならない。多数のプロセッサを管理するために、人間社会にみられるような階層的な構造になるであろう。数多くの並列なプログラムの同時実行を支援するためには、オペレーティングシステムがプログラム間の資源分割の調整と実行時における資源再配分を行う必要がある。局所性、グループ化、および平衡原理が分割のガイドラインとなる。

オペレーティングシステムのひとつの目的は、プログラマが並列プログラムを書きやすくするために、汎用並列システムの抽象レベルを高めることである。この抽象化を通じて、オペレーティングシステムは上記のガイドラインに従うことで、資源とユーザープログラムの協調を管理する。

超並列システムにおけるオペレーティングシステムの研究課題を以下に示す。

#### 1. 階層構造

プロセッサの柔軟な管理を目指して機能的に分散された管理システムを実現するためには、オペレーティングシステムは階層的な構造を必要とするであろう。階層的構造は超並列システムに拡張性をもたらす。ソフトウェア、ハードウェアにより、作業や階層構造を制御する機構が必要である。また、並列性の制御やクリティカルセクションの実行におけるオーバヘッドを削減することが考慮されなければならない。

#### 2. ネットワーク管理

高度な知的ルーティング、アドレッシング、同期、デッドロック回避、流れ制御、そして故障の回避は、柔軟なネットワーク管理システムに組み込まなければならない。

#### 3. 資源管理と負荷分散

超並列システムにおいては、同期に起因するオーバヘッド、アクセスの衝突、そして通信オーバヘッドがより大きな問題となるであろう。これらの問題を解決するた

めに、オペレーティングシステムは自律的に管理情報を収集し、統計的あるいは適応的に管理することが可能でなければならない。効率的なスケジューリングや負荷分散のために、メモリ管理や資源管理の仮想化について研究する必要がある。たとえば、ページングと仮想記憶は、記憶管理のアイデアとして従来型コンピュータシステムにおいて大きな成功を収めた。これらの利点を並列システムにも期待することは妥当であろう。プロセッサ間での負荷をバランスさせるために、プロセスグループを他のプロセッサの集合に移動させることも望まれる。有効なプロセス移動を支援するようなハードウェアが、オペレーティングシステムと同じ位必要である。

#### 4. 耐故障性

超並列システムにおいて、資源管理は予想される構成要素の故障率を許容するような方式で実行されなければならない。したがって、故障回避システムを通常の処理として扱う必要がある。故障を回避するために、多重ルーティング処理が必要となる。超並列システムを効率良く管理するために、高度な知的ルーティングや故障回避機構が柔軟なネットワーク管理システム内に取り込まれていなければならない。

### 3.2.3 超並列システムにおける言語

超並列システムにおけるプログラミング言語では、数多くのプロセスの統合的なオペレーションの記述が可能でなければならない。いくつかのパラダイムに基づいたいくつかの言語が研究されるべきである。問題は、並列性を抽出する方法とシステムが提供可能な並列度を最大限に利用して実行する方法である。そのようなシステムのユーザープログラムは非常に複雑なものとなろう。大規模な並列プログラムの開発、検証、試験、デバッグ、そして保守は、逐次コンピュータに較べより複雑になるであろう。百万のプロセッサに拡張可能な種々のコンパイル技術や実行時処理の技術を研究しなければならない。実際的な応用が記述できる高水準言語を幾つか開発する必要がある。以下に示す項目について研究を行なう。

#### 1. 言語モデル

言語モデルは超並列システムに実現される柔軟なプログラミング言語のための記述モデルである。言語モデルは、根底にあるアーキテクチャに対して基本的なプログラミングの抽象化を与える。そのモデルは単純で計算パワーを損なうことなく十分にアーキテクチャに近くなければならない。同時に、ソフトウェアのプログラミングを容易にし、可搬性を高め、再利用を可能とする強力な手段を提供しなければならない。このような性質を持つひとつの有力な候補として、並列オブジェクト指向モデルが挙げられる。言語モデルの研究において、柔軟な言語、協調的なアクションを記述するモデル、継承やリフレクションに対する基礎研究を推進すべきである。

継承はオブジェクトをコンピューティング環境に適応させる手助けとなる。最近の研究では、並行的な挙動や継承に関してオブジェクト指向の並行性に対する深い理論的な基礎が与えられた。言語におけるリフレクションの考え方は、根底にあるプログラム実行の枠組みを言語内で記述することである。リフレクションは、並列処理システムにおいてスケジューリング、通信、そして負荷分散といったさまざまな側面をプログラミング言語の枠組みの内部に取り入れることを可能にする。

## 2. 超並列システムにおける高水準言語

超並列システムにおける高水準言語は百万プロセッサのスケールにおけるプログラミングの容易さと計算の記述能力を主な目的としなければならない。いくつかの並列処理パラダイムに基づいたいくつかの高水準言語が考慮されるであろう。そのような言語のなかの一つの有力な候補として、並行オブジェクト指向、関数的・宣言的な制約に基づいたアプローチを適切に組み合わせたものとなる言語があるであろう。このような高水準言語はいくつかの中間的なプログラミング言語の階層から構成されるであろう。共通の基本言語によって、いくつかの高水準言語における共通的なプログラム抽象化を提供することが可能となる。効率的な階層間のコンパイル技術は、より高水準な言語の利用を阻害することなく開発されなければならない。並列制約充足の枠組みの中では、見込み計算のサポートも推進される。並行オブジェクト指向言語に関しては、現在のオブジェクト指向モデルでは百万単位の処理を考慮していないため、以下に示すような拡張が必要となるであろう。

- 複雑な問題を階層的に分解することが可能な記述系の導入
- 種々のメッセージ伝搬システム
- オブジェクトを適応・進化させるためのセルフリフレクション機能の導入
- オブジェクト間の関係の宣言的な記述

### 3.2.4 システム開発・プログラミング環境

#### 1. プログラミング環境

複数のパラダイムをサポートすることによって、ユーザーはいくつかのプログラミング言語の中から選択できる自由を与えられる。したがって複数パラダイムのプログラミングをサポートするようなプログラミング環境の開発の必要性がある。デバッグ、負荷分散のグラフィック表示による監視や解析、通信特性などのツールが必要である。これらはハードウェアによりサポートされるため、これらの要求を考慮してアーキテクチャの設計を行わなければならない。

## 2. システム開発支援環境

超並列システムの開発をサポートする環境に対する要求は従来のものとは異なるふたつの特徴を含んでいる。ひとつは相互結合網に対する支援である。相互結合網の開発においては、耐故障性、動的負荷分散、大域同期などを含むいくつかの目的に対する要求や制限が検討されなければならない。相互結合網の全体的な機能と性能は、システムレベルのシミュレーション以前に評価される必要がある。もうひとつは要素プロセッサのアーキテクチャ開発の支援である。超並列システムにおける要素プロセッサのアーキテクチャ開発においては、相互結合網を含むさまざまなサブシステムの機能的な評価を通じ、要素プロセッサの基本的な機能が決定されていなければならない。

### 3.2.5 システムの評価

応用分野は、リアルワールドコンピューティングのための超並列計算の有効性を検証するために取り上げられなければならない。設定されたモデルから将来の出来事を予測するシミュレーションは、このような目的の一例である。シミュレーションは微視的な要素間の相互作用を記述することで巨視的な振る舞いを予測することが可能となる。このような超並列システムの応用として、ダイレクトマッピングなどのパラダイムを用いることによる非線形な問題あるいは多体問題などの処理を行なうことがあげられる。また、現象が複雑すぎたり確率的な現象を記述するための知識が不十分なため、そこで支配する方程式を記述することが困難なようなシミュレーションも扱えるようになると期待される。

## 3.3 研究スケジュール

最終的に開発されるべき超並列システムは百万オーダーの要素プロセッサを持つことが期待される。しかしながら、そのようなシステムを一気に開発するのは大きな危険を伴うと考えられる。したがって、プロトタイプ化による評価が重要となる。プロトタイプシステムはこの計画の中間で構築される。開発に要する技術としては、現在利用可能な技術と最終システムで要求される技術との中間的なものとなるであろう。半導体技術の進歩を睨む必要がある。ソフトウェア開発のためのツールや新機能のための研究プラットフォームシステムも開発される必要がある。プロトタイプシステムのひとつがプラットフォームとして設計、開発されるであろう。このプラットフォーム上でのソフトウェア開発は、本研究プログラムにおいて高水準計算するための基礎として大変に重要である。最終的なシステムは超並列システムとニューラルシステムとの統合になる。

超並列システムの研究開発は、以下に示すような2段階に分けることができる。

## 1. 前期

いくつかのハードウェアプロトタイプシステムがこの段階で開発される。プロトタイプのひとつはプラットフォームとして設計、開発されなければならない。それ以外のプロトタイプシステムは、最終的な超並列システムの間目標となるように設計、開発されなければならない。前期の終了時には、プラットフォームシステムを含むプロトタイプシステムは、最終的なシステムを設計するという立場から評価される。超並列のモデルとアーキテクチャに対する基礎研究も並行して進められる。言語やオペレーティングシステムに関する概念の形成やそのプラットフォームシステムでの実験もここで行なわれる。超並列システムのための共通的な基盤言語の設計と開発が最初に行なわれる。言語モデルに関する研究は応用指向の高水準言語の抽象化の手段を提供するために行なわれる。後期において実際の応用問題を解決するために使われるいくつかの高水準言語やオペレーティングシステムはプラットフォーム上に実装されていなければならない。いくつかのコンパイル技術やランタイム処理の技術はいくつかのプロトタイプシステムを用いて研究される。新しいユーザープログラミング環境もまた、百万オーダーの並行オブジェクトの同時実行を容易にするために開発されなければならない。基礎言語の上位には、ユーザーの要求に応じた問題指向の高水準言語も提供しなければならない。

## 2. 後期

最終システムは、前期で開発されたプロトタイプシステムの評価に基づいて、本研究プログラムの後半で設計される。最終システムを構築するための基礎となる技術は、プロトタイプシステムの中から選択される。前期で開発された重要な技術は、最終システムの設計において統合される。本研究プログラムの後半では、前期で研究された超並列モデルに基づいた超並列システムの開発に重点をおく。これは、本研究プログラムの前半で得られた理論的な基礎や新機能の研究成果を踏まえたものとなるであろう。この最終システムとニューラルシステムは、さらに1つの情報処理システムとして統合される。後期における超並列システムのソフトウェア研究は、超並列のオペレーティングシステムや言語システムの実現に焦点が置かれる。高水準言語モデルの拡張についての研究も並行して行なわれる。最適化コンパイルの技術に関する研究の成果はプラットフォームシステム上に実装されるとともに、最終システム上でのプロトタイプ化を図る。さらに、複雑さを扱うために大規模並列プログラムを階層的に分割する方法や解析の枠組みを研究しなければならない。

本研究プログラムに参加する多くの研究者は、彼らの研究所からプラットフォームシステムを使うことができない。これは、研究所とプラットフォームシステムを高速の広域ネットワークで結ぶことで可能となる。このためのネットワーク環境の

全体構成に関する研究開発もなされなければならない。さらに、システムソフトウェアは、標準ユーザーインタフェースや現在使われているシステム環境を支援する必要がある。少数の要素プロセッサで構成される小規模なプラットフォームシステムを各研究所に設置することも考えられるであろう。多くの研究者が共通の状況下で研究できるということは、実世界コンピューティングのための共通な概念と技術を与え、本研究プログラムの推進のために非常に重要である。

## 4 ニューラルシステム

### 4.1 研究の枠組み

神経細胞の処理速度はシリコン論理ゲートに較べて5桁から6桁遅いにもかかわらず、脳は最速のデジタルコンピュータよりもはるかに速く計算処理を行う場合がある。ニューラルシステムは、生物が用いている組織化原理を利用するように構築されたシステムである。脳を模したこのようなニューラルシステムは、特定の情報処理を高速に処理し、知的な振る舞いを行うことが期待される。ニューラルネットワークによる情報処理が注目されるのは、ニューラルネットワークでは、多数の単純な処理ユニットが互いに密な情報交換を行ないつつ、「協調と競合」の原理にしたがって処理を行っている点である。さらに、従来型のコンピュータで動くプログラムは結合強度で表現されているため、システムが学習や適応を行う際には、従来の数値的最適化手法を用いることが可能である。このような特徴から、ニューラルシステムは柔らかな情報処理のプラットフォームとなり得ると考えられる。

近年、様々なニューラルネットワークが提案されてきており、パターン処理、最適化問題における近似解法、制約充足処理などに適していることが示されている。しかしながら、それらのネットワークは小規模なアプリケーションに限定されたものである。それらのアプリケーションで使用されるニューラルモデルは非常に単純なものであり、学習方法はほとんどの場合、バックプロパゲーションを採用していた。柔らかな情報処理の実現は、過去のアプリケーションと比較してより難しいものであり、我々は、新たなニューラルモデル、学習方法、自己組織化などといった、ニューラルネットワークの新たな可能性について探求しなければならない。

今日では、ニューラルネットワークは従来型のコンピュータ上に構築され、シミュレーションが行なわれているが、特に大規模なネットワークにおいては、シミュレーション速度は非常に遅い。この問題はおそらく汎用超並列コンピュータにより解決されるであろうが、コストと集積規模とを考慮すると、専用ハードウェアの出現が望まれるであろう。

本プログラムにおいては、実世界領域で動作する柔らかな情報処理システムを実現するための大規模ニューラルネットワークの可能性を探求したい。本プログラムの目標は、新たなニューラルモデルにもとづいたハードウェアおよびソフトウェアからなるニューラルシステムを構築することであり、ニューラルモデルの開発は本プログラムの重要なテーマの1つである。そのモデルは、外界との相互作用により知識を学習したり自身の構造を適応的に変化させることが可能な、大規模ニューラルネットワークをサポートするものでなければならない。

ハードウェアシステムは100万ユニットのニューラルネットワークをサポートし、



10 TCUPS (Tera Connections Updates Per Second) の処理速度を実現することが望まれる。最終段階では、柔らかな情報処理を実現するため、ニューラルシステムと超並列システムが統合される予定である。そのため、ニューラルシステムと超並列システムの研究者は、早い時期から統合のメカニズムを検討しておく必要がある。

## 4.2 研究課題

### 4.2.1 ニューラルモデル

本プロジェクトでは、柔らかな情報システムを実現するための新たなニューラルモデルを開発しようとしている。柔らかな情報システムは、実世界との相互作用により適応的に自己の構造を変化させる大規模ニューラルネットワークにより実現される。ネットワークは実時間環境において、教師あり学習あるいは自己組織化により動的に成長する。しかし、成長に際して存在している知識を忘れないようにしなければならない。これらの要求を満足するため、ネットワークは以下の特性をもって構築されるべきである。すなわち、モジュール性、階層的構造化、および機能分化である。以上の特性を実現するためには、以下に示すような各領域の研究が必要であろう。

#### 1. ニューロンユニットモデル

これまでは単純なニューロンユニットモデルが利用され、応用も狭い範囲のみで成功を収めていた。より洗練されたニューロンモデルの開発が知的な機能を実現すると思われる。例えば、カオスニューロンモデル、複素数ニューロン、ニューロン論理モデル（ニューロンに論理属性が保持される）などが、より洗練されたニューロンユニットモデルの候補として提案されている。我々はこれらの可能性を評価し、新たなニューロンユニットモデルを開発しなければならない。

#### 2. モジュール化と階層化

実世界で利用される大規模ニューラルネットワークは、外部環境との相互作用を行ない、適応や訓練によってネットワーク構造を変化させる機能を備える必要がある。そのような状況では、既存の知識を破壊することなく新たな知識を獲得し、情報を効率よく取り出す能力が必要である。これらの要求を満たすためには、知識のモジュール化と階層化とがなされなければならない。モジュール化、構造化および機能分化のための分散学習法や学習制御法を開発する必要がある。このためには以下の機構に関して考慮すべきであろう。

- 外部環境への適応による機能分化および階層構造の構築
- 実世界との相互作用によるニューラルネットワークの自律的成長

- 集中あるいは分散制御を用いたモジュール化
- モジュール間相互作用
- モジュール化と階層構造化を評価するための手法

### 3. 学習と自己組織化

リアルワールドコンピューティングにおいては、実時間学習が非常に重要である。既存の知識を破壊することなく新たな知識を追加でき、かつ一貫性を保ちながら既存の知識の削除を行えるような学習方式を開発することが重要である。大規模ニューラルネットワークの多くは、再帰ループを持っている。階層式ニューラルネットワークが空間的パターン認識に有効であるのに対し、リカレントニューラルネットワークには時間的パターン認識や時間的パターン生成に対し有効と期待される。またリカレントニューラルネットワークは、制約充足の機能を有するため、最適化問題に利用できるであろう。リカレントニューラルネットワークは今後10年以上に渡って重要な役割を果たすと思われる。リカレントネットワークに対する学習や自己組織化の機構を構築する必要がある。

ネットワークのトポロジーやサイズは、汎化能力を決定するうえで最もクリティカルなパラメータに属するものである。最適な自己組織ネットワークを構築するための学習手続きが開発されなければならない。そのようなネットワークは問題学習のための十分な大きさを持つと同時に、汎化を行なうために十分小さくしなければならない。

### 4. 連想記憶

連想はニューラルネットワークにより創出される基礎的な機能の一つである。空間パターンや時間パターン等、色々な情報が分散記憶され、ベストマッチの原理で想起が行なわれる。これら連想機能の原理を理論的に解明すると共に、その機能を実現するための工学的機構が開発される必要がある。

### 5. アナログコンピューティングの新たな原理

ニューラルシステムの情報処理は、システムがもつアナログ的な非線形力学を基盤としている。この視点から、カオス力学も含む、ニューラルシステムのアナログコンピューティングの原理について、明確化されなければならない。

### 6. 異なるパラダイムの統合

異なるパラダイムを統合したモデルに関する研究が行なわれなければならない。例えば、ニューラルネットワークと論理処理との統合、パターン処理と記号処理との組み合わせなどが、ニューラルシステムを実現する上で求められよう。統合モデル

は並列処理環境上で実行されなければならない。入出力表現は、様々なパラダイムを統合する際のインタフェースとなる。特に、ニューラルネットワークの入出力表現の選択は、ネットワークの処理能力に大きく影響する。したがって、ニューラルシステムの開発に際しては、入出力表現に関する理論的立場および実験的立場からの研究が重要となる。

#### 4.2.2 ニューラルシステムのハードウェア

ニューラルシステムのハードウェアは、外部環境との相互作用により構造を適応的に変化させるようなモジュール化された大規模ニューラルネットワークをサポートしなければならない。ニューラルシステムは 100 万ユニットのニューラルネットワークをサポートすべきであろう。このように大規模なニューラルネットワークは、1000 個のニューロンが完全結合されたサブニューラルネットワークから構成されるようなものになる。目標とする処理速度は 10TCUPS (Tera Connections Updates Per Second) である。専用ハードウェアの設計においては、現時点でニューラルネットワークモデルが明確にされていないこと、将来において他のさまざまな新しいモデルが発表されるであろうということに注意しなければならない。ニューラルシステムのアーキテクチャは、汎用かつ拡張性のある機構を盛り込まなければならない。本研究プログラムの初期段階ではさまざまなアーキテクチャに関する研究が実施される必要がある。

ニューラルシステムのハードウェアは以下の 3 種に分類される。

- ニューロアクセラレータ
- VLSI ニューロチップ
- ニューラルネットの工学的実現

ニューロアクセラレータは、ニューラルネットワーク処理のために開発される専用の並列プロセッサである。多くのアーキテクチャが提案されているが、その典型的なものは、数百の要素プロセッサからなり、1GCPUS を実現している。VLSI ニューロチップはニューロンユニットモデルをシミュレートするニューロンユニットのハードウェアから成る。ニューロチップアーキテクチャの領域は広く、デジタル処理チップからアナログ処理チップの範囲にまで及んでいる。3 番目の工学的実現のアプローチは、ニューラルネットワークの機能をニューロンユニットモデルを使わずにハードウェアロジックで工学的手法で実現するものである。

以上の各アプローチは学習能力、スケーラビリティ等についてそれぞれの特徴を持っており、比較することは不可能である。本ニューラルシステムにおいては、2 番目および 3 番目のアプローチに焦点を当てることとする。この場合以下の様な選択が考えられる。

#### 1. デジタル回路ニューロチップ

デジタル回路は耐ノイズ性と処理精度の高さという点において多くの利点を備えており、大規模システムにおける安定的な処理に適している。デジタルニューロ回路にはコンピュータの製造技術を直接応用することが可能である。従来のアプローチに加え、パルス密度モデル等を利用する多くのバリエーションが提案されている。これらのアプローチによりニューロチップの新たな可能性が生じることも期待されよう。

#### 2. アナログ回路ニューロチップ

アナログシステムは、演算回路が少ないため、ハードウェア体積の削減が可能である。これは大規模ネットワークの開発においては非常に魅力的である。さらにアナログ回路には、動的ネットワークや、その他カオスニューロンのような複雑なニューラルネットワークを実現する可能性もある。

#### 3. ハイブリッドニューロチップ

デジタルおよびアナログの各々の利点を利用したハイブリッドチップを考えることも可能である。

#### 4. 工学的実現

上述したニューロチップはニューロンユニットモデルを直接実現するものである。これに対し、ニューラルネットワークの機能を実現するための別のアプローチも存在する。それは、ニューロンユニットモデルをシミュレートしない論理ハードウェアによる連想のようなアプローチである。これらのシステムは単純な論理回路の繰り返しにより構築される。

ニューラルシステムでは、基本的に全てのニューロンは基本的に活性値を互いに交換している。相互結合ネットワーク・アーキテクチャは設計上の重要なポイントである。時間多重方式と周波数多重方式とが、この問題に対する答えの候補となるであろう。結合ネットワークの設計は実装技術と深い関係がある。結合ネットのバンド幅は、チップやボードのピン数の制限で決まる。これらを、解決するための技術としてウェハスケールでの集積、三次元アーキテクチャ、光技術等が考えられる。また、設計のための道具としてCADやシリコンコンパイラ等も重要と思われる。

#### 4.2.3 ニューラルシステムのソフトウェア

ニューラルシステムの研究および開発においては、様々なニューラルソフトウェアシステムが必要とされる。

### 1. シミュレーションシステム

ニューラルネットワークの研究および開発においては、シミュレーションシステムが重要なツールである。ニューラルネットワーク研究者のほとんどは独自のシミュレータを持っている。大規模ニューラルネットワークのための柔軟性のある汎用ニューラルシミュレータが強力なツールとなるであろう。このようなシミュレータに対して求められる要件には、マシン独立性、拡張性、便利なユーザインタフェース、高速性、豊富なユーティリティルーチン等がある。又、シミュレータには、ネットワークの収束性等を制御するための解析ツールが付加されていることが望ましい。

### 2. ニューラルネットワークの言語

ニューラルネットワーク処理は、高水準言語によって記述されるべきである。このような言語の設計のため、以下に示す研究テーマ、開発テーマが必要となろう。

- 曖昧情報の表現
- ベストマッチ操作の記述
- 論理プログラムとの統合
- シミュレータとの統合

### 3. オペレーティングシステム

一般的に、ハードウェアのニューロン数は、ニューラルネットワークにおけるニューロン数を下回る。したがって、このギャップを埋めるための仮想機構がニューラルシステムにおいては重要である。

- ニューラルネットワークをハードウェアにマッピングするための機構
- 資源のスケジューリング

#### 4.2.4 超並列システムとの統合

ニューラルシステムは、多種多様なニューラルサブシステムから構築されるであろう。しかしながら、ニューラルシステムは、パターン認識、連想記憶、最適組合せの探求、および他の特殊な機能といった専用目的のためのプロセッサのひとつとなるであろう。このことは即ち、ニューラルシステムが他のシステムと結合あるいは統合されなければならないことを意味する。強結合から疎結合まで、様々な統合機構が考えられる。すでに提案されている統合機構には以下のようなものがある。密結合の例：ニューラルネットワークが連想メモリとして超並列システムの要素プロセッサに結合されたもの。疎結合の例：超並列プロセッサが例えば最適化問題のような問題をニューラルシステムに対して投げ掛けるようなもの。

#### 4.2.5 システムの評価

ニューラルネットワークの理論およびモデルは、4.2.1 節に列挙された機能を満足しているか否かにより評価される。ニューラルハードウェアは2つの方法により評価される。第一の評価方法は処理速度 (TCUPS) によるものであり、第二の評価方法は本計画で開発される実際のアプリケーションによるものである。ベンチマーク用アプリケーションでは、柔軟かつ高速な情報処理が実装されるべきであろう。柔軟な動的イメージの処理等がベンチマークプログラムの候補である。

### 4.3 研究スケジュール

ニューラルシステムの開発は2段階に分かれる。各段階はそれぞれ5年間継続する。

第1の開発段階では、新規の理論、モデル、およびハードウェアアーキテクチャが研究される。ニューラルモデルは、幅広く様々な方式を試みる必要がある。それらは評価され、後期において使用可能なモデルが選ばれる。モデルはやがてハードウェアに反映される。ハードウェアのプロトタイプが次期の開発段階のために構築される。プロトタイプシステムは、数万ユニットからなり、後期のニューラルモデルや新機能の実験に用いられる。この段階で利用される技術は、現時点で入手可能なものと最終システムで必要とされるものとの中間と位置づけられる。

第2段階は最終システムの構築に費やされる。第1段階の成果に基づき、理論およびモデルが精製される。ニューラルモデルは、最終システム上で効率良く実行されるよう、モデルのハードウェアへの実装法も研究されるべきである。最終システムは第1段階のプロトタイプの経験を生かし、百万ユニット規模のシステムが構築される予定である。最終マシンの開発には、その時点の最新研究結果を生かし、最新の技術が使われる。最終システムは超並列システムと統合される。統合の機構は、最終システムのアーキテクチャにおいて最も重要な点となるであろう。特に第2段階の初期において、理論およびモデルの研究成果に基づき、新たな統合アーキテクチャについて議論すべきであろう。ニューラルシステムと超並列システムの両グループは互いに関連して、新たな統合アーキテクチャを設計、構築すべきであろう。システム上にはいくつかのアプリケーションシステムが実装され、システム全体が評価されることになるであろう。

## 5 光システム

### 5.1 研究の枠組み

光はその大容量性と超並列処理性により、新しい情報媒体として期待されている。光の情報処理への利用は、柔らかな情報処理および超並列超分散情報処理の実現をめざすリアルワールドコンピューティングプログラムに対して、新しいアーキテクチャ、アルゴリズムおよび新しいデバイス技術を提供する。

本プログラムで開発すべき技術は次の3つのカテゴリーに分けられる。

- 光インタコネクション技術
- 光ニューラルシステム
- 光デジタルシステム

光インタコネクション技術は、現状の電子システムが直面しているいわゆる配線限界を克服することが目的である。また、光インタコネクションは、光ニューラルシステムや光デジタルシステムのための技術基盤として位置づけられる。

光ニューラルシステムは、画像情報などの広がりを持つ情報に対して、光によるニューロン間の接続を実現し、学習や連想を実時間に処理可能とする技術の開発を目指す。

光デジタルシステムでは、光をベースとして正確な計算を行う超並列処理の実現を目標とする。

光コンピューティング技術では、光デバイスの開発がまずその前提となる。

開発された光電子素子のモジュール化もまた、本計画の重要な研究目標の一つである。こうしたモジュールによって、高度に並列で分散的なシステムに対する光の寄与が検証される。

### 5.2 研究課題

#### 5.2.1 光インタコネクション技術

光インタコネクション技術は、VLSIに代表される先端的なエレクトロニクス技術と光通信技術を融合したものであり、伝送遅延、クロストーク、配線のための空間、電力消費、といったような電子システムにおける情報伝達に関わる問題を克服するものである。

光インタコネクションは上記のような問題を克服し、高速、大容量で柔軟な情報伝送の実現をめざす。

重要と思われる技術的課題は次の通りである：

## 1. 光インタコネクションデバイス

光インタコネクションデバイスには、光電子的な能動および受動素子が含まれる。これらのデバイスの組合せによって、時間・空間・波長の3つの領域における高密度多重技術を用いた高速、大容量、構造変更可能な相互接続ネットワークが実現できる。

ピコ秒以下の超高速な光インタコネクションデバイスの研究では、時分割多重を利用した大容量インタコネクションの実現をめざす。

空間並列および／または機能的な光インタコネクションデバイスの研究では、空間分割多重技術と多様な電子-光的機能とを利用した構造変更可能な高速インタコネクションをめざす。

波長並列および／または機能的な光インタコネクションデバイスの研究では、波長分割多重技術を用いた大容量のインタコネクションをめざす。とくに、波長分割多重技術は、大容量バスの実現に有効であろう。

マイクロオプティクスおよび開折素子のような受動光デバイスの研究は安定性が高く、高密度な光インタコネクションのためのモジュールの実現をめざす。

さらに、上記のような異なる性質／機能を持つデバイスを統合した高度な光統合デバイスの研究は、次の世代の光デバイスの基礎となる。

## 2. 光インタコネクションネットワーク

アーキテクチャーおよびデザイン技術の研究は、構造変更可能で、自己ルーティングができるようなコンパクトで高速、多チャンネルの光総合結合ネットワークの実現をめざす。

チップ内およびチップ間結合の研究は、プロセッサ間、プロセッサ-メモリ間、およびメモリ間の高速で柔軟な光インタコネクションネットワークの実現をめざす。

光電子デバイスと受動光デバイスとのモジュール化の研究は、光インタコネクション要素の統合と微小化をめざす。これによって、位置合わせの問題が回避される。

### 5.2.2 光ニューラルシステム

光ニューラルシステムでは、超並列かつ柔軟な光接続によって、学習や連想機能を実現し、画像情報などの空間的に分散された情報や波長情報を実時間処理する技術の実現を目標とする。

重要と考えられる研究開発課題は次の通りである：



### 1. 光ニューラルモデル

画像情報を直接処理するモデルでは、2次元、3次元情報を直接ニューラルネットワークに入力し、処理を行なうようなモデルの開発を目指す。

光の物理現象を導入するモデルでは、双安定現象、カオス、位相共役といった光物性に基づく現象を利用した新しいニューラルネットワークの理論的なモデルの開発を目指す。

拡張可能なモジュラーモデルでは、多数の単位モジュールから構成される拡張可能かつ柔軟なネットワークモデルを開発対象とする。

アナログデバイスに適したモデルでは、低精度ながら大規模、高速性に優れた光アナログデバイスを利用するニューラルモデルが開発の対象となる。

### 2. 光ニューラルデバイス

大規模・高速学習光ニューラルデバイスでは、電気または光の学習信号でシナプス強度を変更する光ニューラルチップの大規模化、高速化技術の開発を目指す。

画像処理用光ニューラルデバイスでは、画像を直接認識したり処理する光ニューロデバイス、画像の特徴抽出を行うニューラルデバイスの開発を目指す。

光ニューラルデバイスのモジュール化では、光ニューラルチップのモジュール化および標準化の実現を目的とする。

### 3. 光ニューラルシステム

設計技術の確立では、機能分割、階層化、分散化などの技術を確立を目指す。また、デジタルコンピュータとの統合化による機能の高精度化技術も開発対象とする。

学習手法の確立では、学習結果を構造化された知識として整理し、他のシステムに移植する技術、学習の高速化技術を対象とし、これらの手法の確立を目指す。

人間に優しい入出力技術では、マルチメディア情報を直接処理する技術の開発や、キーとなる画像で類似像を直接検索できる画像情報ベースの開発を目的とする。

#### 5.2.3 光デジタルシステム

論理演算原理を基本として、光の超並列性と柔軟接続性を生かしながら、画像情報に代表される空間的広がりを持つ情報や波長情報を処理可能とするのが、光デジタルシステムである。

本システムの開発における重要とおもわれる研究開発課題は次の通りである：

## 1. 光論理素子

超高速光論理演算素子では、2値あるいは多値論理の超高速処理を行う光デバイスを開発するとともに、低消費電力で2次元集積化する技術の開発を目指す。

空間並列光論理演算素子では、空間符号パターンとして2値符号化し、デジタル演算を行う光論理素子を開発する。

波長並列光論理演算素子では、異なる波長の組合せとして符号化し、デジタル演算を行う光論理素子を開発する。

受動光デバイスでは、マイクロ光学素子、平面光学素子、回折光学素子および高精度光学素子を開発する。

## 2. 光論理回路

光論理素子間光接続技術としては、光接続回路構成法と、その部品化を開発目標とする。

光機能モジュールでは、専用演算ユニット、並列光レジスタ、並列光メモリ、並列クロスバー接続装置、並列入出力装置などのモジュール技術の開発を目指す。

## 3. 光デジタルシステム

アーキテクチャおよび設計技術の確立では、可変並列光コンピュータの設計および3次元光回路設計用CAD技術の確立を目指す。

入出力インタフェースでは、光論理回路と電子システムとの高速情報交換を可能にする物理的、論理的インタフェースに関する技術の開発を目指す。

光機能モジュールの実現技術では、光機能モジュールのアレイ化、複数の異なる光機能モジュールの一体化のための実装技術の開発を目指す。

プログラミング言語、コンパイラ開発では、電子的コンピュータ・プログラミング言語やコンパイラと整合性の高い、光デジタルシステム用高級言語とそのコンパイラの開発を目指す。

### 5.2.4 システム開発の環境

光電子集積回路(OEIC)は、光ニューラルシステムや光デジタルシステムと同じように光インタコネクションにとってもキーとなるデバイスである。OEICの開発はシステム内における互換性を保つために、処理およびモジュール技術に関する共通のプラットフォームの上に行なわれるべきである。

ここでは次に挙げる技術の研究開発が課題となる：

- 先進的OEICプロセス技術
- OEIC用CAD技術
- 標準化およびモジュール化技術

先進的OEICプロセス技術では、高性能化、信頼性向上、低コスト化にOEICを実現するための先進的な材料プロセス技術の開発を目指す。

OEICモジュール化技術では、光インタコネクション／ニューラル／デジタルデバイス、光受動素子、電子回路、入出力回路、等を一つのパッケージにモジュール化する技術の開発を目指す。

OEIC用CAD技術では、OEICの構造設計、および回路、パッケージモジュールのシミュレーションを支援するためのCAD技術の開発を目的とする。

標準化およびモジュール化技術では、光電子入出力インタフェースの標準化と光デバイスのモジュール化を目指す。

### 5.3 研究スケジュール

光システムの開発は、2つの段階に分けて進められる。それぞれの段階の研究期間は5年間である。

前期には、光インタコネクション、光ニューラルシステムおよび光デジタルシステムに必要となる基本的なデバイス技術が開発されることになる。同時に、光部品の標準化およびモジュール化も研究課題とする。小規模な光システムのアーキテクチャに関する研究も実施する。

後期には、デバイス技術についてある程度の進歩が見えてくるので、光コンピューティングシステムのプロトタイプの開発に着手する。光デバイスの大規模な集積化を開始し、本研究プログラムの終了段階では、実世界コンピューティングのために複数の光モジュールが利用可能になるであろう。単位面積あたりのスループットとしては テラビット／平方センチ秒オーダの性能が期待される。



## 第 2 部 資料編

### A. 基本計画書概要

### B. 基本計画書（英語版）

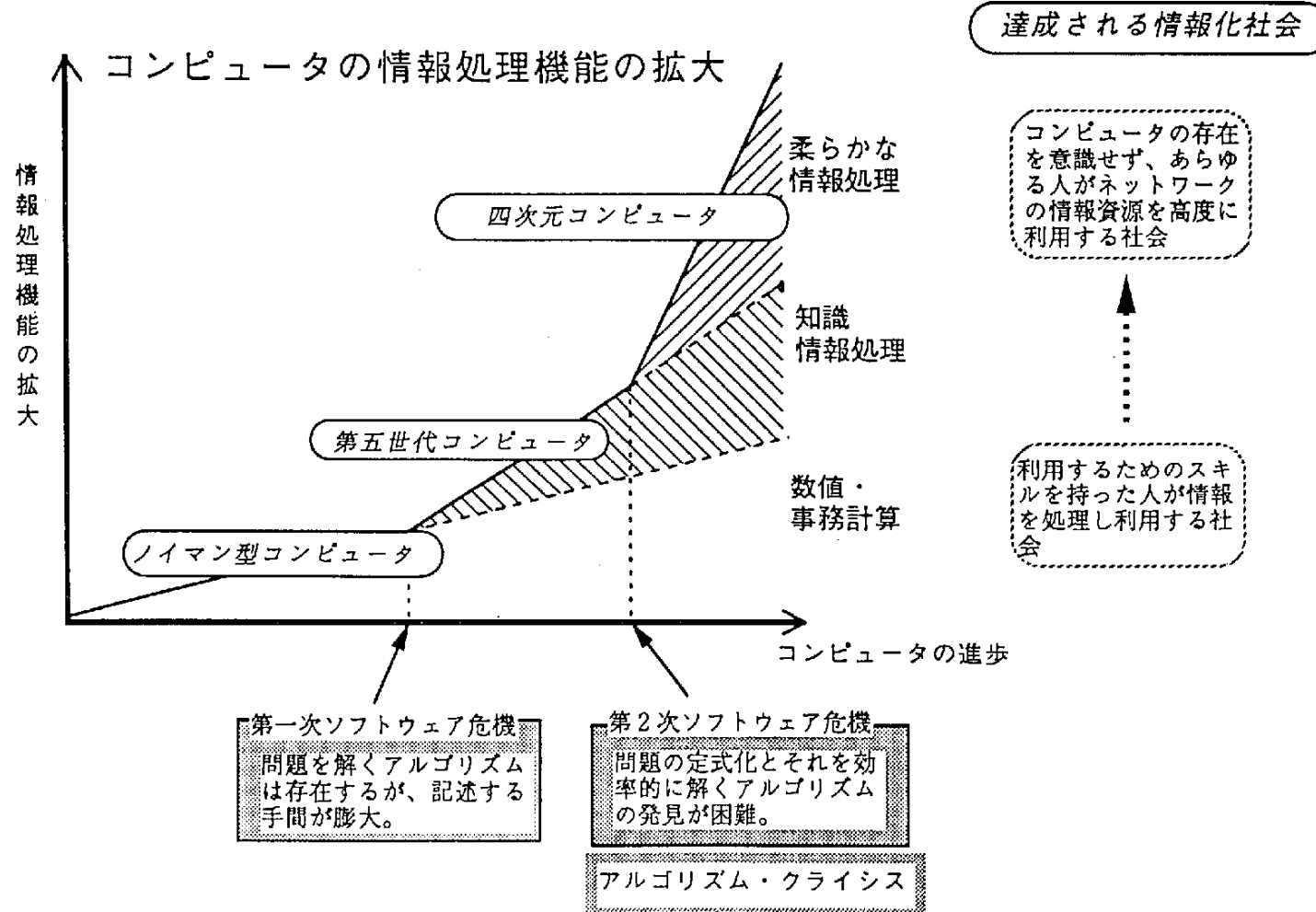
(The Master Plan for the Real-World Computing Program)



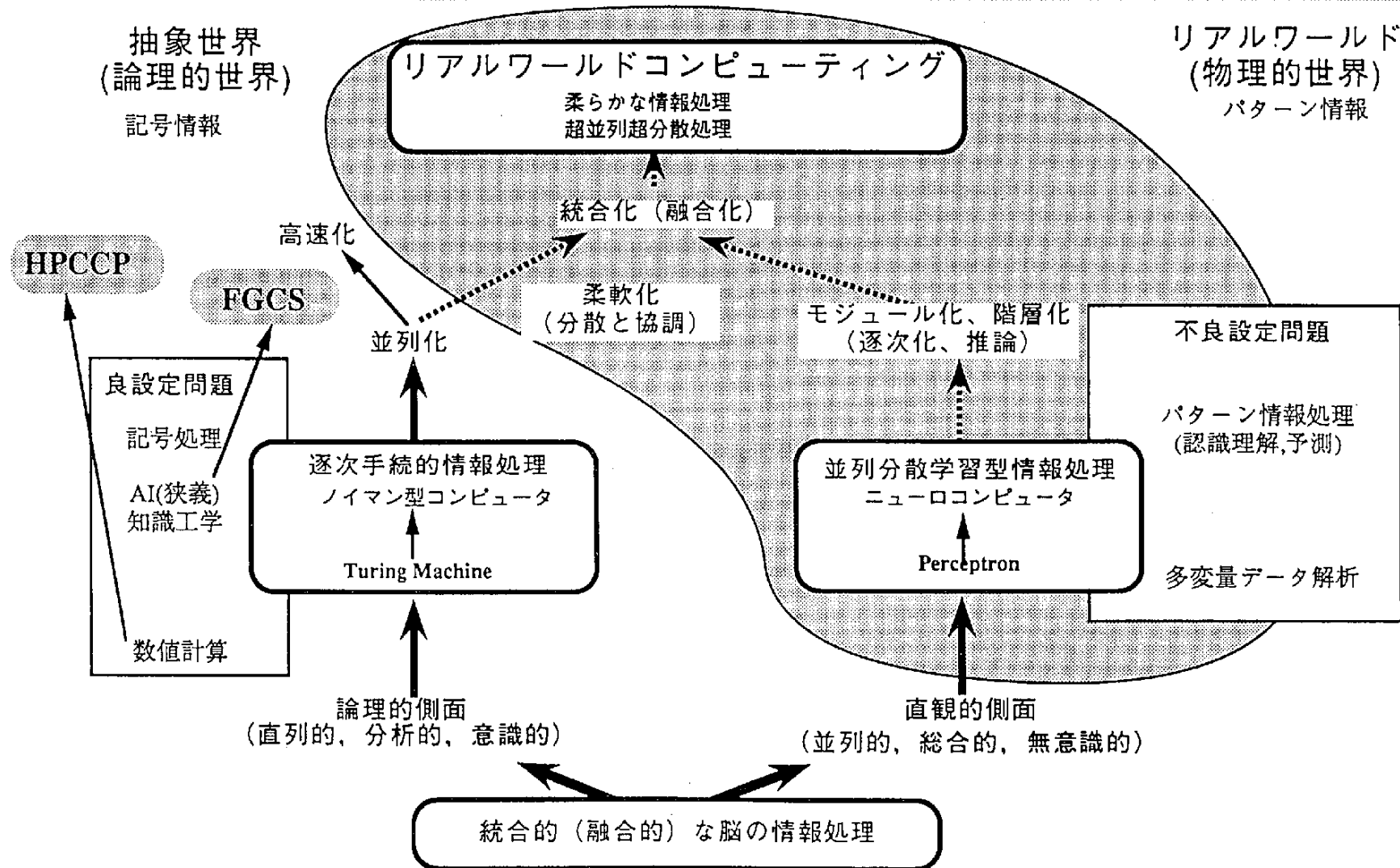
## A. 基本計画書概要













リアルワールドには多種多様な情報が存在する。そうした情報の曖昧さ、不完全さ、変容性を許容しつつ、妥当な判断や問題解決などを行うことを可能にする21世紀の情報処理技術体系の基本理念を「リアルワールドコンピューティングパラダイム」と呼ぶ。

リアルワールドコンピューティングを可能にするためには、次の二つの技術基盤が不可欠である。

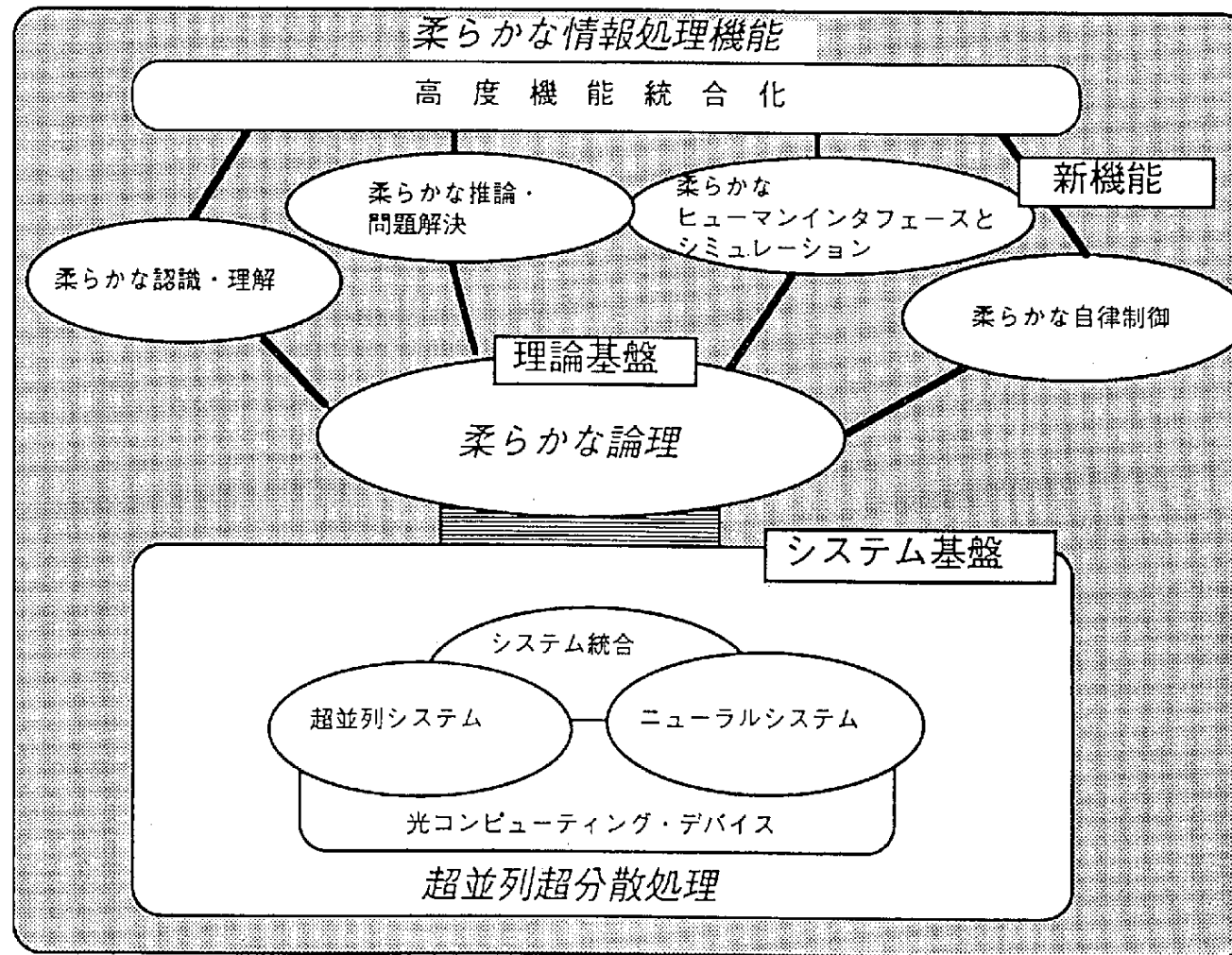
- 1) 人間に近い柔軟性と学習性、適応性をもった「柔らかな情報処理」という情報処理の枠組みのための理論基盤
- 2) 限られた時間内に、限られた物量を用いて、柔らかな情報処理を行う「超並列超分散情報処理」という計算の枠組みを反映したシステム基盤



- 1) 人類の持続的な発展の実現に向けて、基礎的かつ革新的な基盤技術の創造へ挑戦
- 2) 国際貢献の観点から、国際的な交流と協力の推進を重すると共に、研究成果を国際公共財として広く内外に公開









1. 柔軟な実施体制： 本研究開発の推進に当たっては、統合性、象徴性、共通性の高い研究を集中研において、個別性、要素性の高い研究を分散研においてという研究課題の適切な配分を行い、両者間の有機的で柔軟な連携を確保する。
2. 競争原理の導入： 研究開発の前期では競争原理（様々なアプローチ間の競争による研究の効果的な進展）を導入し、中間評価時にはその結果に基づいて研究開発課題を絞り込む。
3. 学際性、国際性： 本研究開発の基礎的かつ挑戦的な目標を達成するため、学際的・国際的な連携を促進する。そのため、電総研や大学等の研究機関との共同研究を積極的に行い、また国内外の大学等の研究機関に対し再委託の公募等を行う。
4. 研究成果の公開性： 研究成果の公開性を保つため、研究開発の進捗状況および成果を国内外の会議等で報告、公開し、またシンポジウムやワークショップ等を積極的に開催する。
5. 研究インフラの整備： 以上の柔軟な実施体制と研究成果の公開性を支えるために、世界的な規模の分散研究基盤として、高速ネットワーク環境の整備を行なう。



## 理論基盤

### 研究の枠組み

目的 : リアルワールドコンピューティングに必要とされる「柔らかな情報処理」、すなわち人間のもつ柔軟な情報処理能力を、情報処理の新しいパラダイムとして実現するための理論基盤を確立する。

方針 : ○情報処理の枠組み自体を、情報の表現・評価のそれぞれにおいて拡張・一般化することを目指す。

○柔らかな情報処理の根底にある論理基盤を「柔らかな論理」として明確化、体系化し、その立場からの統合的な理論を構築する。

### 研究課題

- (1) 情報の柔軟な表現
- (2) 情報および処理モジュールの評価
- (3) 柔軟な記憶と想起
- (4) 情報および処理モジュールの統合
- (5) 学習と自己組織化
- (6) 最適化技法



## 新機能

### 研究の枠組み

目的： リアルワールドコンピューティングの新機能として、頑健性、開放性、実時間性に基づく柔軟さを実現する技術を確立する。

方針： ○新機能の統合されたシステムの例として、リアルワールド適応自律システムと情報統合対話システムの研究を行う。

○想定されるシステムを実現するため、柔軟な認識・理解、推論、制御などの新機能と、それら要素新機能を統合するメカニズムの研究を行う。

### 研究課題

#### ○要素機能研究

- (1) 柔軟な認識と理解
- (2) 柔軟な推論と問題解決
- (3) 柔軟なヒューマンインタフェースとシミュレーション
- (4) 柔軟な自律制御

#### ○機能統合研究

- (1) リアルワールド適応自律システム
- (2) 情報統合対話システム

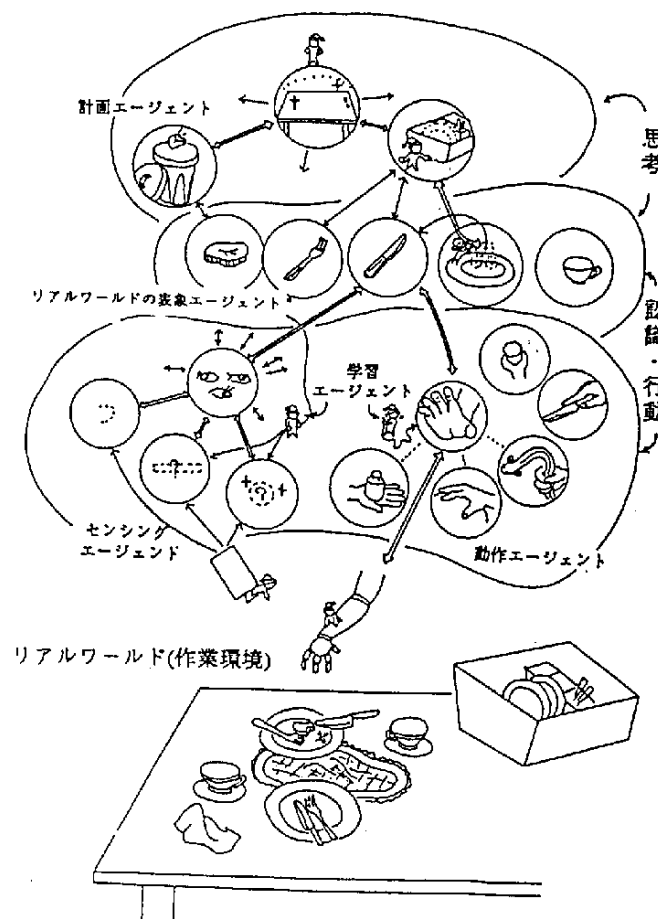




リアルワールド適応自律システム

不完全、不確実性、変容性をもつリアルワールドの情報を能動的かつ適応的に理解学習し、働きかけを行う自律的な情報処理システムの実現を目指した研究開発を行う。

例えば、多様な自律エージェントの協調によるリアルワールド作業システムを開発する。





情報統合対話システム

多様な情報と人間の間にあって、人間の意図を理解し多様な情報を柔軟かつ統合的に処理することにより、問題解決や情報創造など、人間の知的活動を支援拡張する情報処理システムの実現を目指した研究開発を行う。

例えば、画像理解、音声対話理解、自己組織化情報ベース等の統合による創造活動支援システムを開発する。





## 超並列システム

### 研究の枠組み

目的： リアルワールドコンピューティングを支える計算基盤を実現するために、汎用的な超並列システムに関する技術を確立する。

方針： ○超並列計算モデルの確立とこれに基づいた超並列アーキテクチャの構築に関する研究を進める。

○半導体技術を始めとするハードウェア技術の進展に則って、研究開発を進める。

○超並列システムのハードウェア、アーキテクチャ、ソフトウェアの研究開発を総合的に行う。

### 研究課題

- (1) 超並列アーキテクチャ
- (2) 超並列オペレーティングシステム
- (3) 超並列言語
- (4) プログラミング環境とシステム開発環境
- (5) システム評価



## ニューラルシステム

### 研究の枠組み

目的： リアルワールドコンピューティングを支える柔軟な計算基盤を実現するため、大規模ニューラルシステムに関する技術を確立する。

方針： ○リアルワールドとの相互作用を通して学習や自己組織化が行える新しいニューラルモデルの研究を行う。

○新しいニューラルモデルに基づく汎用性のある大規模のニューラルハードウェアを開発する。

### 研究課題

- (1) ニューラルモデル
- (2) ニューラルハードウェア
- (3) ニューラルソフトウェア
- (4) 超並列システムとの統合





## 超並列システム

超並列システム

複数の思考モデルを支援することのできるマルチパラダイム表現を超並列計算モデルによって実行するために、その計算基盤となる汎用超並列システムの研究開発を行う。

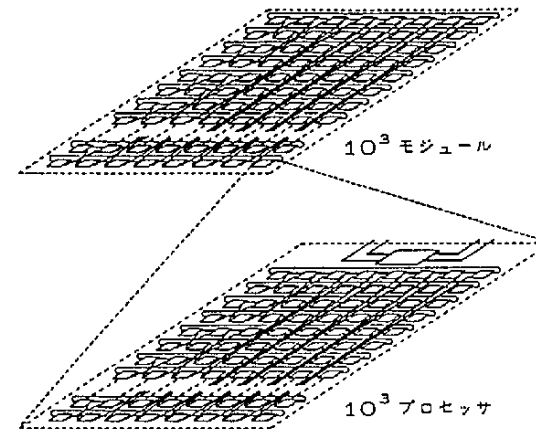
## システム統合

柔らかな情報処理の計算基盤を実現するために超並列システムと大規模なニューラルシステムの統合化に関する研究開発を行う。

## ニューラルシステム

任意のニューラルネットワークが構成できるように汎用でスケラブルなアーキテクチャを構成し、ウェハ・スケール・インテグレーションによる大規模ニューロンシステムの研究開発を行う。

システム統合



<システム>

$10^6$  要素プロセッサ

<テクノロジー>

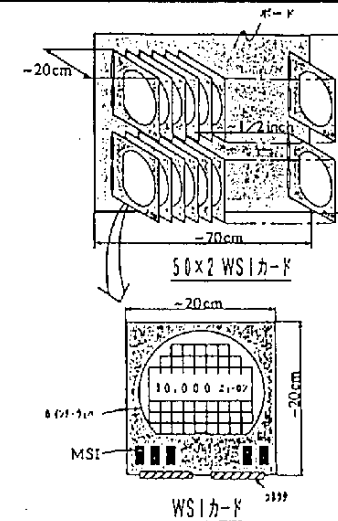
40要素プロセッサ

/チップ

25チップ/モジュール

1,000モジュール

/システム



<システム>

$10^6$  ニューロン

<テクノロジー>

$2 \times 10^7$  トランジスタ

/チップ

60チップ/ウェハ

$10^4$  ニューロン/ウェハ

100ウェハ/システム



## 光コンピューティング・デバイス

### 研究の枠組み

目的： リアルワールドコンピューティングが目指す柔らかな情報処理を実現する手段として、電子技術が不得意とする大容量情報伝送技術や画像情報などの直接処理を実現する光技術を確立する。

方針： ○光の大容量性や超並列性を利用して大容量情報伝送技術や、柔軟な光接続技術の研究開発を通して、現在の計算機がかかえる配線ネックの問題を解決する。  
○画像情報に代表される広がりを持つ空間情報や波長情報を従来の論理演算原理を用いて処理する光デジタル技術を開発する。  
○ニューロン間の膨大な結線を光で実現し、学習や連想処理などの柔らかな処理を行う光ニューロ演算技術を開発する。

### 研究課題

- (1) 光インタコネクション
- (2) 光ニューラルシステム
- (3) 光デジタルシステム
- (4) 光共通要素技術の研究



## 光デジタルシステム

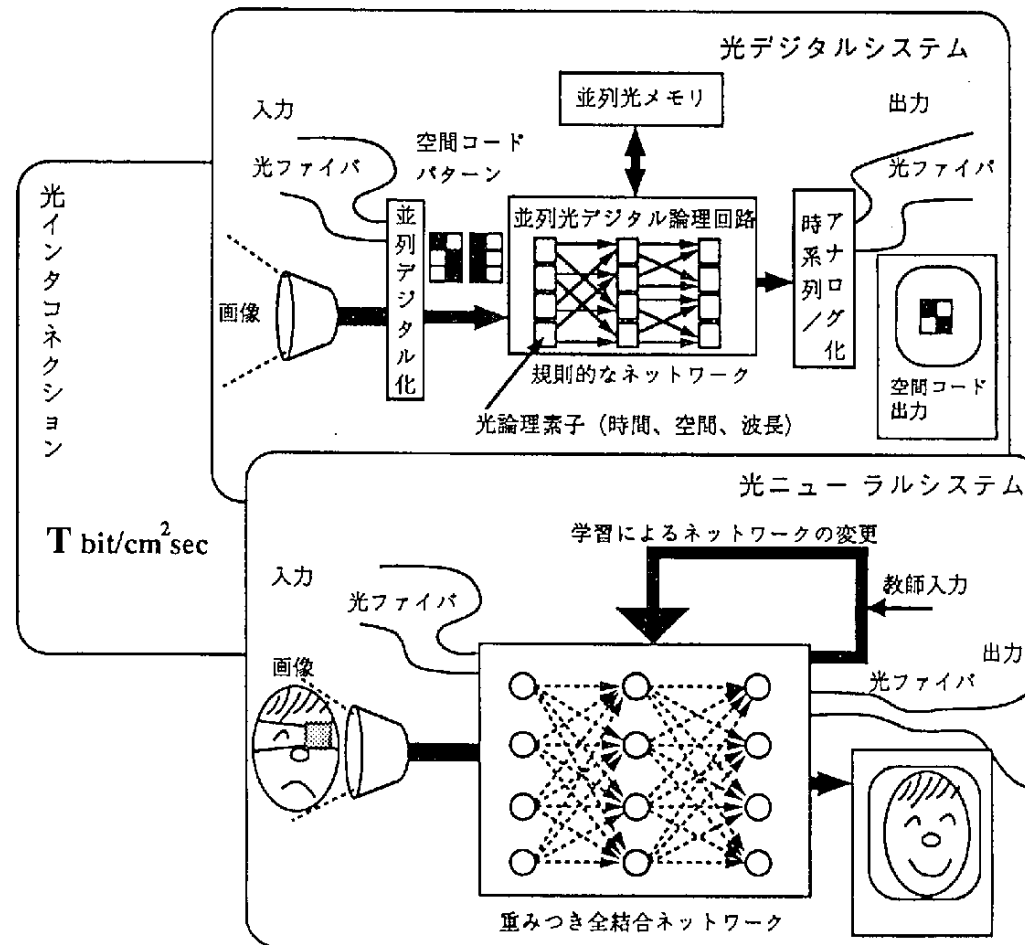
従来の論理演算原理を基本として、光の並列性を生かしつつ、画像情報に代表される一定の広がりを持つ空間情報や波長情報を処理する技術の研究開発を行う。

## 光インタコネクション

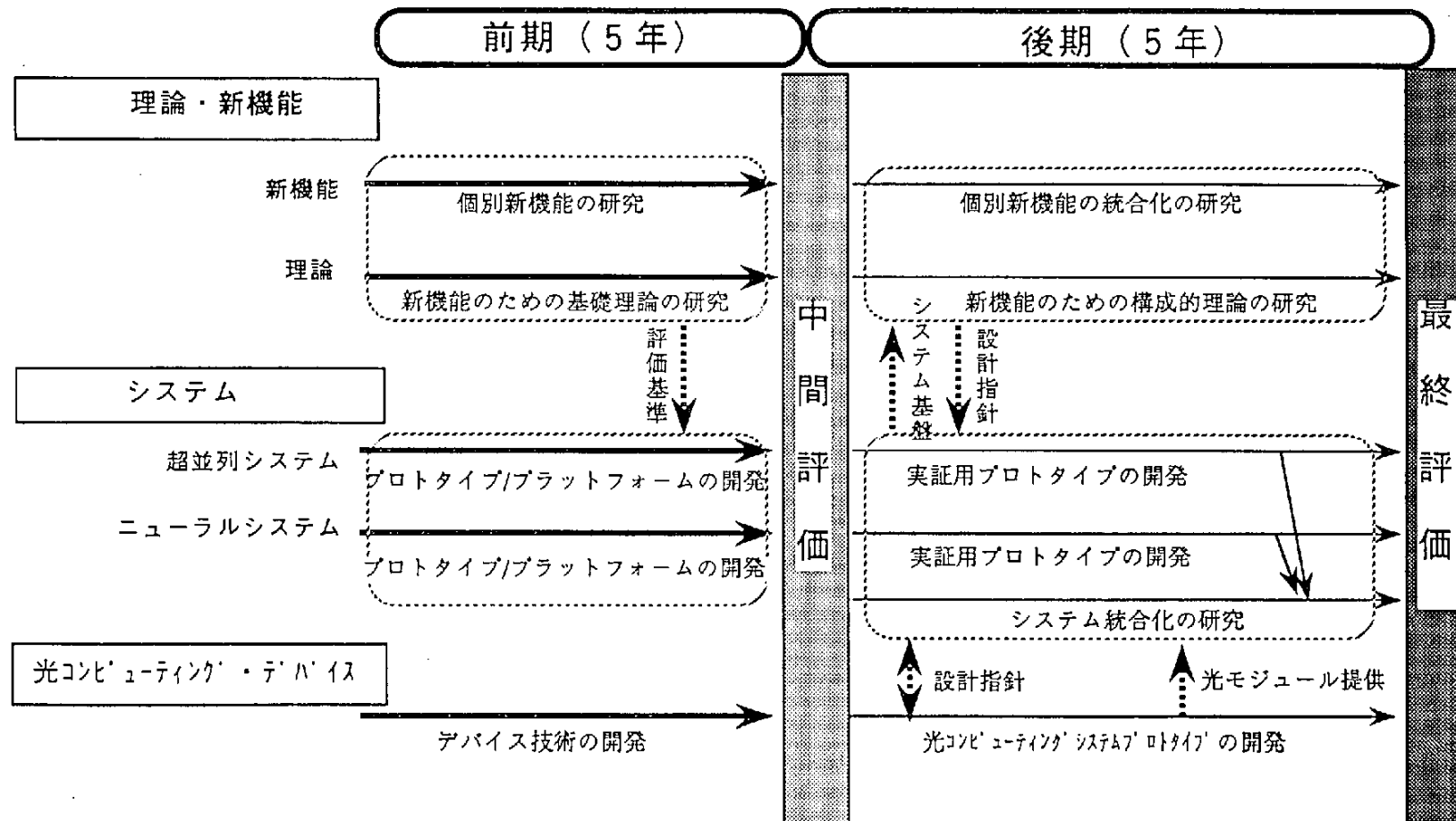
光の大容量性や超高速性の利用により情報伝送の増大を目指す技術や、光伝搬の並列性と制御性を利用した柔軟な光接続技術の研究開発を通して、現在のコンピュータの持つ配線ネックの問題を解決する。

## 光ニューラルシステム

ニューロ間の膨大な結線を光で実現し、光独自の物理特性を利用することにより学習や連想処理などの柔軟な情報処理を行うシステムの研究開発を行う。











集中研

- ・リアルワールドコンピューティングパラダイムの象徴的実現課題
- ・共通性のある基盤技術・支援技術
- ・統合化技術

(1) 理論・新機能

リアルワールドコンピューティングパラダイムの象徴的実現課題

- ・リアルワールド適応自律システムの研究
- ・情報統合対話システムの研究

(2) システム

新情報処理技術開発全体にとって共通性のあるシステム基盤

- ・超並列システムの研究
- ・システム統合化の研究

(3) 光コンピューティング・デバイス

光コンピューティングシステム技術の実証と共通支援技術

- ・光システムの研究
- ・光共通要素技術の研究
- ・先進光デバイス技術の研究

分散研

- ・分散研究拠点の個別的特性を活かすことにより効果的に推進される研究課題
- ・要素性の高い基盤技術
- ・理論研究

(1) 理論・新機能

- ・柔軟な記憶と想起の研究
- ・情報および処理モジュールの統合の研究
- ・学習と自己組織化の研究
- ・最適化の方法の研究

- ・柔らかな認識と理解の研究
- ・柔らかな推論と問題解決の研究
- ・柔らかなヒューマンインタフェースとシミュレーションの研究
- ・柔らかな自律制御の研究

(2) システム

- ・超並列アーキテクチャの研究
- ・超並列ソフトウェアの研究
- ・ニューラルシステムの研究

(3) 光コンピューティング・デバイス

- ・光インタコネクションの研究
- ・光ニューラルシステムの研究
- ・光デジタルシステムの研究



## B. 基本計画書（英語版）

(The Master Plan for the Real-World Computing Program)



# Contents

## Part 1. Fundamental Principles of the Master Plan

1	Purpose of the Research and Development	2
2	Background	3
3	Real-World Computing Paradigm	5
4	Contents of the Research and Developments	8
5	Fundamental Policy	10

## Part 2. R & D Plan of the Real-World Computing Program

1	Theoretical Foundation	12
1.1	Research Framework	12
1.2	Research Topics	13
1.2.1	Flexible Representation of Information	13
1.2.2	Evaluation of Information and Processing Models	14
1.2.3	Flexible Storage and Recall of Information	14
1.2.4	Integration of Information and of Processing Modules	15
1.2.5	Learning and Self-Organization	15
1.2.6	Optimization Methods	16
1.3	Research Schedule	16
2	Novel Functions for Application	18
2.1	Research Framework	18
2.2	Research Topics	20
2.2.1	Flexible Recognition and Understanding	20
2.2.2	Flexible Inference and Problem Solving	23
2.2.3	Flexible Human Interface and Simulation	24
2.2.4	Flexible Autonomous Control	25
2.3	Research Schedule	25
3	Massively Parallel Systems	27
3.1	Research Framework	27
3.2	Research Topics	27

3.2.1	Massively Parallel Architectures . . . . .	27
3.2.2	Operating Systems for Massively Parallel Systems . . . . .	29
3.2.3	Languages for Massively Parallel Systems . . . . .	30
3.2.4	Environments for System Development and Programming . . . . .	32
3.2.5	System Evaluation . . . . .	32
3.3	Research Schedule . . . . .	32
<b>4</b>	<b>Neural Systems</b>	<b>35</b>
4.1	Research Framework . . . . .	35
4.2	Research Topics . . . . .	36
4.2.1	Neural Models . . . . .	36
4.2.2	Hardware for Neural Systems . . . . .	38
4.2.3	Software for Neural Systems . . . . .	39
4.2.4	Integration with Massively Parallel Systems . . . . .	40
4.2.5	System Evaluation . . . . .	41
4.3	Research Schedule . . . . .	41
<b>5</b>	<b>Optical Computing Systems</b>	<b>42</b>
5.1	Research Framework . . . . .	42
5.2	Research Topics . . . . .	42
5.2.1	Optical Interconnection . . . . .	42
5.2.2	Optical Neural Systems . . . . .	43
5.2.3	Optical Digital Systems . . . . .	45
5.2.4	Environments for System Development . . . . .	46
5.3	Research Schedule . . . . .	46

**Part 1.**  
**Fundamental Principles of the Master Plan**

# 1 Purpose of the Research and Development

The main purpose of the Real-World Computing Program is to lay the technological foundation of the highly advanced information society of the 21st century.

The development of computers has been very rapid and computers have come to provide a huge amount of computing power. Nevertheless, in many respects, including for examples, pattern recognition, problem solving based on incomplete information, and learning capability, computers still lack the flexibility of processing and are far behind the human being.

In this program, we aim at establishing the bases for flexible and advanced information technologies which are closely allied to humans and able to process a variety of diversified information in the real world. They are essential for the highly advanced information society in the 21st century in creating the cooperative relationship among human and computers and producing the innovative and generic technologies.

Information systems in the 21st century will be based on not a single but various key technologies such as massively parallel computing, optical computing, neural computing, and logic-based computing. These technologies should be flexibly integrated into information systems in order to cope with real-world problems. The goal of this program is, first of all, a challenge for these significant but not yet established technologies. We try to establish the theoretical foundation for these emerging technologies, explore their potentials and give some prospects of the style of integration. Some important real-world problems will be attacked for confirming the possibility and usefulness of the technologies. Through opening these experiences to the public, we aim at contributing to the development of the common knowledge and wealth of human-kind.

In order to accomplish our very fundamental and ambitious goals, promoting international and interdisciplinary cooperations in research is very important. Supporting the collaboration among industries, national institutes, and universities is also imperative.



## 2 Background

Supported by the development of computer and communication technologies, the information society will produce an innovative change, not only in the industrial activities such as new wealth and services, rationalized production and distribution system, but also in the qualitative improvement of national living standard, advancement of localities and the educational and cultural aspects.

In particular, as the information network will be developed from intra-organization network inside an enterprise to that of inter-enterprises and further to those that uniting homes and individuals, the subject of information processing will change from enterprises to homes and individuals. As a result, information to treat will greatly increase not only in quantity but also in variety, and a new technological basis which enables everyone to easily take advantage of the various information resources of the network to a higher degree will be required as a basis for such information network society.

Under the social background as well as technological requirements in various application fields of information processing, it is needed that the computers become more humanized and possess the capability to assist and collaborate with humans in the real-world environment.

Historically, computers have achieved a remarkable development in support of the technological progress. In the first stage, computers developed along the line of conventional von Neumann architectures in such fields as numerical computation, document processing, managing data-base in storage and retrieval, since those application fields have clear algorithms and are suitable to process by conventional computers.

The second stage of development was the direction of manipulating logic and symbol, casting light to the intelligent thinking process of human such as logical (deductive) inference. The research fields of AI and knowledge engineering contributed to provide computers with capability of handling symbolic representation of knowledge and inference rules. In 1981 Japan proposed the concept of the Fifth Generation Computers as an approach to new generation computing. It intended to provide computers with powerful logical inference capability and open the door to the world of large-scale knowledge information processing.

Now computers are far superior to the human beings in the ability of solving the well-defined problems where the algorithms for solution exist and can be clearly stated in programming languages. Nevertheless, computers are still not so flexible compared with human's flexible information processing in the real world where many problems to solve are ill-defined and hard to describe in algorithm.

Hence, in order to cope with the real-world problems and open a new horizon of information processing technology, it is important and necessary to pursue the funda-

mental principles of human-like flexible information processing by casting light to the intuitive or subsymbolic level of human information processing and to embody those as new information technologies on the basis of the developing hardware technologies.

In recent years a large amount of fundamental knowledge on information processing of human beings has been accumulating, and hardware technologies for massively parallel systems are expected to come to our hands at the beginning of the 21st century.

Taking these situations into account, we propose the concept of "Real-World Computing" as a new paradigm of information processing, which aims at giving the "real-worldness (or flexibility)" of human information processing ability to information processing systems (see also Figure 1.1).

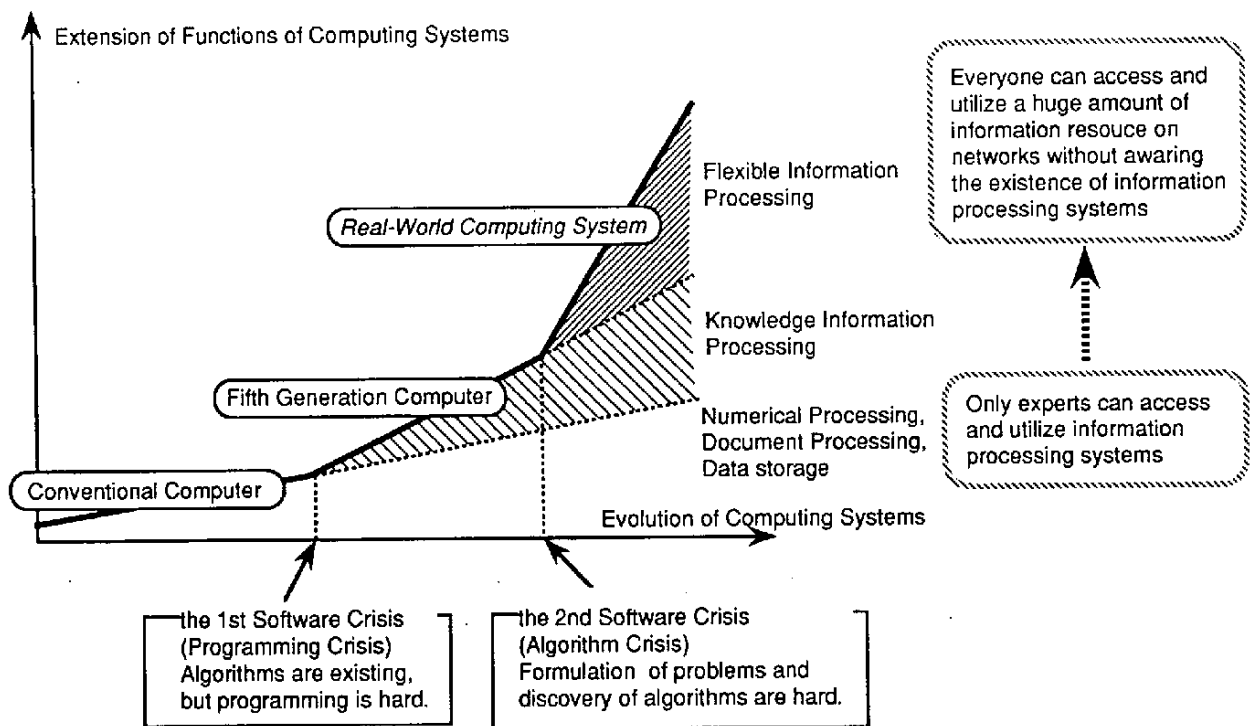


Figure 1.1: Extension of the functions of computing systems.

In order to make the 21st century a globally prosperous information network society where human-kinds and information processing systems collaborate more closely and flexibly, the Real-World Computing Program, a challenge for novel fundamental generic technologies in the fields of information processing, should be carried out immediately.

### 3 Real-World Computing Paradigm

Information processing in general is a function that has been acquired by creatures in the course of evolution in surviving and adapting to the changing real-world environment. It is a many-sided function that enables us to recognize the surrounding environment, make predictions, and then plan actions based on the predictions. It may be classified into the following two categories:

**A:** Logical information processing

**B:** Intuitive information processing

Logical information processing is characterized by the words: conscious, analytic, serially intensive, and digital. To the contrary, intuitive information processing is characterized by the words: unconscious, synthetic, parallel distributed, and analog.

Basically, information processing technology is intended to complementing or substituting both aspects of our information processing function as a total. Historically, however, machines which are suitable for functions of aspect A were implemented as conventional computers and progressed very rapidly mainly supported by the tremendous growth of hardware technologies. On the other hand, in the so-called "intuitive information processing ability" aspect, conventional computers are far behind human beings. The lack of flexibility of the computers is due to the unbalanced development of them.

In our daily life, we human beings evaluate various kinds of information including ambiguity or uncertainty and acquiring necessary information for making predictions, plans, or decisions in quite a flexible way in a changing environment. Development of systems which have these flexible information processing functions is now one of the most important requests common to various fields of today's advanced information processing aiming at the further-developed information processing functions such as pattern information processing, knowledge information processing, and intelligent robots.

In order to develop such systems, it is most important to explore the intuitive aspects of information processing functions of human beings and to integrate the above two aspects. As the result of integration, we get the new paradigm of information processing which will be the foundation of a more human-like, flexible information systems. We call such a flexible information processing as the real-world computing (see Figure 1.2).

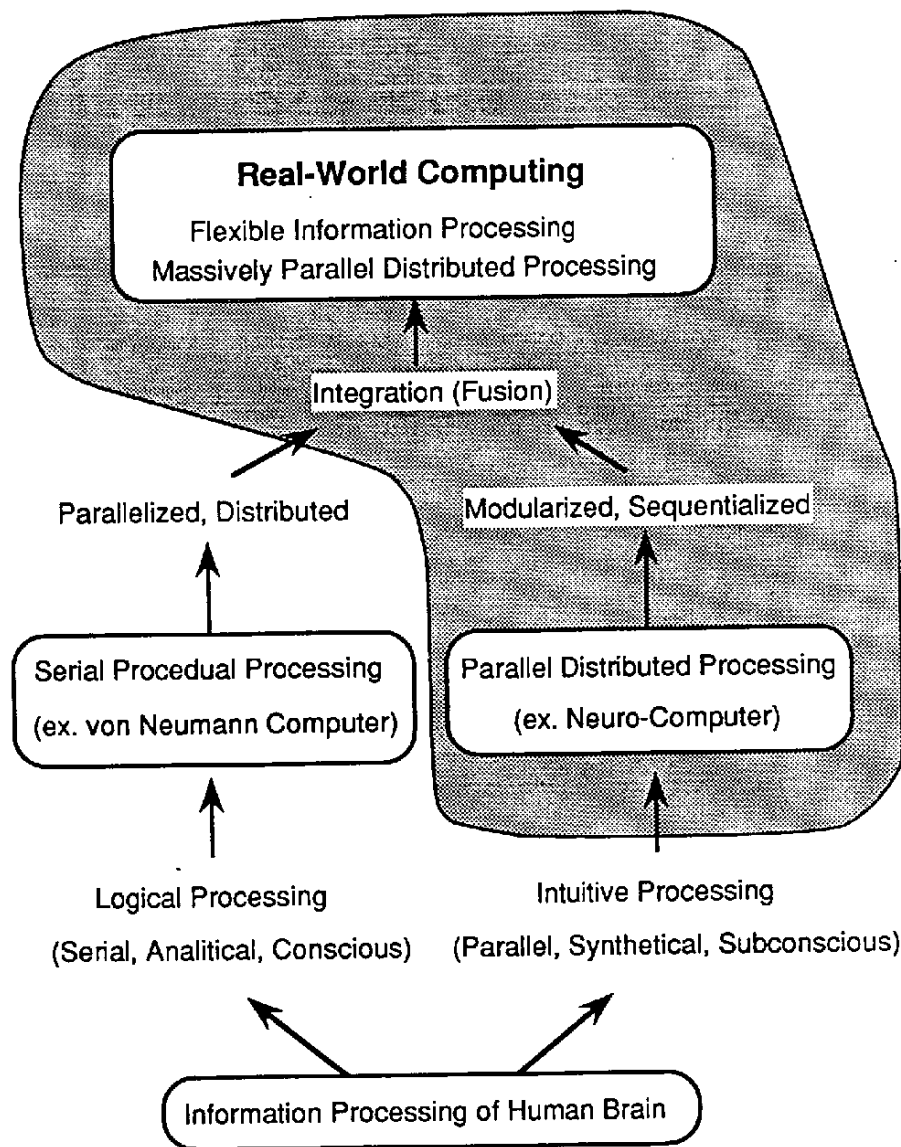


Figure 1.2: Bifurcation and unification of two aspects of information processing.

Real-world computing systems are characterized by some proper properties as shown in the followings:

- Openness
- Robustness
- Real-time

“Openness” means that systems can adaptively and autonomously change or extend themselves to cope with unexpected situations they encounter in the real world. “Robustness” means toughness and the stability of the system behavior in distort of input

data or fluctuating environment. "Real-time" means the systems can respond within a reasonably short time.

The reason why human beings can maintain their flexible information processing ability characterized by the above real-worldness is that our brains incorporate distributed representation of information, massively parallel processing, learning and self-organizing ability, and information integration ability. Hence, the following two seem to be key concepts for realizing the above characteristics of real-world computing systems.

- Flexible information processing

Information processing which ables to process ambiguous or uncertain information and has the capability of adaptation and learning.

- Massively parallel and distributed processing

Information processing which processes multi-modal, multi-variate, and strongly correlated information in a massively parallel and distributed manner.

For implementing the above two key concepts, we define the following three technological issues as essential elements of the Real-World Computing Program:

- Theory for flexible information processing,
- Massively parallel and distributed information processing systems,
- Novel elemental functions for real-world applications.

Through research and development on these issues, the following functions will be achieved not only by developing high-speed computers, but also by integrating conventional logical information processing functions with intuitive information processing functions.

- The function to integrate a variety of complicatedly related information including ambiguity or errors and to make an appropriate (approximately acceptable) decision or solution within a reasonable time.
- The function to actively acquire necessary information and knowledge, and to learn general knowledge inductively from examples.
- The function to adapt the system itself to the users and the changing environment.

Applications of real-world computing systems are expected to include wide range of real-world problems, for examples, incompletely specified (ill-defined) problems such as understanding of situations in a noisy environment, large-scaled problems such as simulation of social and economical phenomena, real-time problems such as man-machine interface with virtual reality and autonomous control of intelligent robots.

## 4 Contents of the Research and Developments

The main contents of the technological developments in the Real-World Computing Program are classified into the following three mutually related categories:

- To investigate the principle of flexible information processing and to establish the theoretical bases of the innovative technology,
- To develop the massively parallel systems as computational bases,
- To materialize the novel functions as the fundamental elements for wide-range actual applications.

Corresponding to these categories, the three-storied structure composed of (a) novel functions for applications at the top, (b) theoretical bases in the middle, and (c) computational bases at the bottom, is the fundamental framework for the research and development of this program (see Figure 1.3).

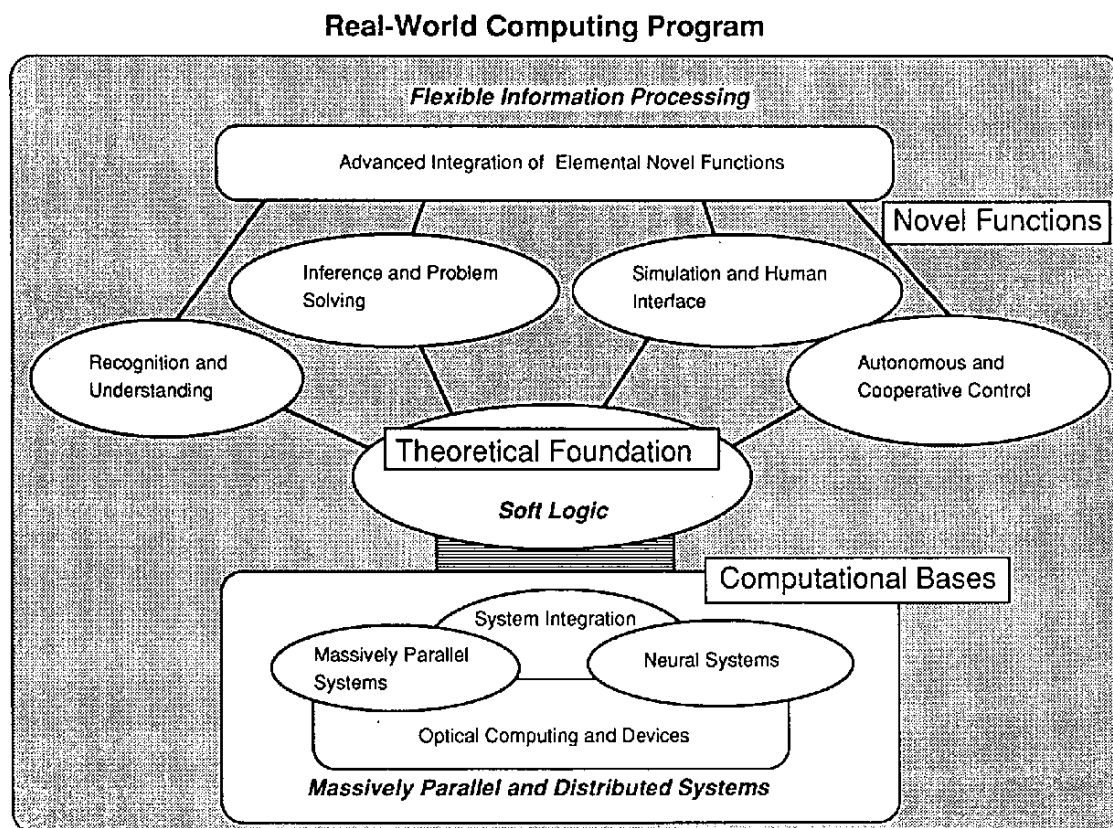


Figure 1.3: Organization of the contents of research and developments.

In the research and development on novel functions for applications, elementary novel functions which are commonly important to realize the systems to solve a wide-range of real-world problems should be investigated. These novel functions are aimed to be implemented on the computational bases. Research topics are classified into the following categories:

- Flexible recognition and understanding of multi-modal information,
- Flexible inference and problem solving based on flexible information base,
- Flexible interactive environment for man-machine communication,
- Flexible and autonomous control.

Some typical application problems will be attacked for exploring the way of integration of these elemental technologies and confirming the possibility of them.

The objective of research on the theoretical bases is to establish a new theoretical foundation for flexible information processing. For this purpose, it is necessary to expand and generalize the conventional framework of information processing in all respects of:

- Representation, storage, and recall of information,
- Integration of various kinds of information,
- Evaluation of information,
- Learning and self-organization.

In particular, "integration of multi-modal information" and "learning and self-organization" are the most fundamental issues and how to implement these in a framework of massively parallel and distributed information processing will be a major key point.

Real-world computing requires processing large volumes of spatio-temporally distributed information at a high speed while taking care of their mutual interactions. As the computational bases of real-world computing, computing systems which can exploit parallel and distributed processing at several processing levels should be developed.

For this purpose, research and development from the following three perspectives is important:

- General-purpose massively parallel systems,
- Neural systems as a kind of special-purpose systems,
- Optical computing systems.

The way of integration of these technologies will also be investigated.

## 5 Fundamental Policy

The primary goal of this program is not to develop a single computer, but to explore the possibilities of elemental technologies which are significant but not yet established. In order to accomplish this challenging and very fundamental goal, the program should be managed under the following fundamental policy.

1. Flexible formation of research organization:

To insure an effective research program, research themes are appropriately distributed so that researches which are highly integrated, symbolic, or common are performed in the central laboratory, and researches which are highly individual or elemental are performed in the distributed laboratories, and to secure organic and flexible cooperation between both parties.

2. Introduction of competitive principle:

To introduce competitive principles (effective progress of research due to competition amongst various approaches) in the first half of the research and development, and to select the research themes to be concentrated on in the second half, based on the evaluation of these results at the middle point.

3. Interdisciplinary and international cooperation:

To promote interdisciplinary and international cooperation to fulfill basic and challenging aims of this research and development. As a result, to actively perform joint researches with organizations like ETL, universities, etc., and invite sub-contract applications from domestic and overseas research organizations such as universities, etc.

4. Publication of research achievements:

To ensure publication of research achievement by reporting and publicizing the progress and results of research and development in domestic and foreign conferences, etc., and actively holding symposiums and workshops as well.

5. Establishment of infrastructure for research activities:

To establish the high-speed computer networking as the infrastructure for international distributed researches, to support the flexible research formation and distribution of research results.



## **Part 2.**

# **R & D Plan of the Real-World Computing Program**

# 1 Theoretical Foundation

## 1.1 Research Framework

The real-world environment (including humans within it) is full of uncertainty and changeability. Human beings, however, can recognize the environment, make decisions and act in quite a flexible way by integratedly processing a variety of information which is usually subjected to incompleteness and ambiguity.

Such information processing of human beings can be characterized by the term "flexible information processing" or "real-world computing" in contrast to the conventional hard information processing by computers which assumes completely given information in a pre-assumed world or problem domain.

The objective of research and development in theoretical part of this program is to lay the theoretical foundation and to pursue the technological realization of human-like flexible information processing as a new paradigm of information processing.

So far, much theoretical research has been done in respective research fields related to flexible and parallel distributed information processing. These are, for examples, pattern recognition, multivariate data analysis, probabilistic and statistical inference, fuzzy logic, and neuro-computing, machine learning, regularization and various optimization methods, and so on.

In order to provide a theoretical foundation for flexible information processing, it is important not only to continue in-depth study in these research areas but also to intend to clarify the theoretical framework of "soft logic" (flexible logical inference) commonly underlying these fields and to aim at constructing a new unified theoretical base for dynamic real-world computing. There, probabilistic and statistical formulation of problems and nonlinear dynamics in conjunction with learning and self-organization will be a key approach.

For this purpose, it is necessary to expand and generalize the conventional framework of information processing in all respects of information representation, processing and evaluation, and to systematize basic theories and elementary novel functions for application on the flexible framework.

The following two are conceived as the most fundamental issues of common concern:

- Integration of multi-modal information (and of heterarchical processing modules)
- Learning and self-organization (optimization and adaptation)

Implementing these in a generalized and flexible framework of massively parallel and distributed information processing will be a major key point to develop novel functions for application such as flexible recognition, inference and control. The characteristics of

the real-world computing systems: "Openness", "Robustness" and "Real-time" should be considered deeply in the researches on the respective topics.

In order to establish the new paradigm of information processing, it is also very important to learn and get inspiration from nature. Taking into account the findings in the scientific researches of human brains, evolution process of creatures, and ecological system, etc. should be promoted by making the chance of exchange of information at interdisciplinary forum.

The theoretical research will be categorized into the following:

- Flexible representation of information,
- Evaluation of information and processing models,
- Flexible storage and recall of information,
- Integration of information and of processing modules,
- Learning and self-organization,
- Optimization methods.

These will serve as the theoretical foundations for the novel functions in various applications and possibly for new computing architectures.

Because researches of this theoretical foundation part mainly aim at constructing general unified theory, in what follow we will show only general research topics important for all of the application domains and all of the implementation methods. Theory and models specifically related to each application area and each implementation technology such as massively parallel computing systems, neural computing systems and optical computing will be discussed in the other proper sections.

## **1.2 Research Topics**

### **1.2.1 Flexible Representation of Information**

In order to treat a wide range of information from raw sensory data such as visual images, auditory signals to high-level symbolic languages in a unified manner and construct heterarchical flexible systems, developing of a flexible framework for representing various kinds of information efficiently and effectively is necessary. This framework should be able to represent certainty of information, and not only flexible enough to represent various kinds of information, but also suitable for implementing associative memory and learning or self-organization procedures.

The followings are considered as important research topics:

- Information representation based on multi-way data table,
- Distributed and sparse representation of information,
- Information representation suitable for implementing specific processing procedure,
- Representation of probabilistic knowledge stochastic automata, hidden markov models, random markov fields, probabilistic decision lists, etc.,
- Representation of information and knowledge as constraints.

### **1.2.2 Evaluation of Information and Processing Models**

For interacting actively with the real world and learning or self-organizing from experiences, real-world computing systems should have the systematic framework for evaluating the importance of information and the capability of processing models. Evaluation criteria of regularization conditions used in the processes of information integration are also necessary.

The followings are considered as important research topics:

- Evaluation of input and output information for learning or self-organization,
- Evaluation of regularization conditions for information integration.
- Evaluation of flexibility and reliability of systems,
- Evaluation of processing models and agents for cooperative integration of them,
- Evaluation of input information for active interaction with real-world environments.

### **1.2.3 Flexible Storage and Recall of Information**

The highly sophisticated functions of the human brain's memory is a key to the flexible information processing of human beings. Real-world computing systems should have flexible memory functions to storage and associatively recall various kinds of information. Thus the research aims at establishing the theoretical analysis of associative memory and developing new efficient mechanisms for realizing flexible associative memory should be done.

The followings are considered as important research topics:

- Association using probabilistic reasoning,
- Association using structural similarity,

- Associative memory for storing time series,
- Associative memory using non-linear dynamical systems.

#### **1.2.4 Integration of Information and of Processing Modules**

Information processing such as inference, prediction, and planning can be considered as an integration process of various kinds of information and knowledge. It will be important to carry out the theoretical research to analyze these information integration processes, and develop a new flexible way for controlling the processes.

Methods of multi-variate data analysis will be provide a candidate for the methods of information integration. However, most of the conventional methods are limited to the linear transformations. Non-linear extension of these methods will be important. Neural network models are considered as giving a kind of non-linear extension. Regularization theory will provide a theoretical foundation for incorporating various kinds of constraints.

Cooperative processing by a huge number of processing modules is also an important issue. Research on the integration of processing modules is important.

On these topics, not so much research works has been done so far. Investigating on the information integration process of human cognitive systems will also be important as a source of inspiration.

The followings are considered the important research topics:

- Integration of information using multivariate data analysis methods and neural network models,
- Integration of various kinds of constraints under the regularization theory,
- Constraint logic systems including dynamics for controlling information integration,
- Explorations of integration of information in human cognitive processes,

#### **1.2.5 Learning and Self-Organization**

One of the most important parts of the theoretical foundation of real-world computing is the theory of learning and self-organization. In order to construct complex heterarchical systems, learning and self-organization should play an important role. It is also important to realize self-organization of memory and data base. In this viewpoint, conducting the research aiming at constructing computational theory of learning/self-organization for exploring the learnability of various concepts or structures and developing novel efficient algorithms will be necessary.

The followings are considered as important research topics:

- Algorithms for learning probabilistic knowledge,
- Algorithms for learning heterarchically structured knowledge,
- Algorithms for learning in a changeable environment,
- Algorithms for learning from uncertain information,
- Learning algorithms using active information acquisition,
- Methods for incorporating existing knowledge with learning process,
- Selection of models for learning.

### **1.2.6 Optimization Methods**

Information integration processes can be formalized as optimization processes. Learning and self-organization can also be considered as optimization procedure. Solving these optimization problems always demand a huge amount of computation. In order to surmount this difficulty, the research aims to develop approximately correct optimization methods which can be executed effectively on massively parallel systems will be important.

The followings are considered as important research topics:

- Probabilistic optimization methods such as simulated annealing, genetic algorithms, ecological algorithms, evolutionary algorithms,
- Optimization using non-linear dynamical systems such as neural network models,
- Other non-linear optimization techniques.

## **1.3 Research Schedule**

The research schedule of theoretical foundation will roughly be divided into two stages.

In the first stage, basic researches are intensively pursued. Theories and models so far presented in relation to parallel and learning type flexible information processing, such as pattern recognition, multivariate data analysis and neuro computing, will be reconsidered and reorganized from unified theoretical viewpoint of "soft logic" through expanding and generalizing the framework of information processing, such as information representation and evaluation. In particular, researches on elementary functions, such as information integration and learning/self-organization, will intensively be pursued.

On the theoretical bases, constructive theory and models for basic novel functions such as flexible recognition, inference and control will be developed in the second stage. At this stage, interaction with applications will be mostly promoted with respect to novel functions. In accordance with system integration of novel functions, theory and models to treat and coordinate elementary functional modules in higher levels will also be pursued.

## 2 Novel Functions for Application

### 2.1 Research Framework

Real-world computing systems support various human activities by acquiring many kinds of information such as images, speech sounds, tactile so on, and processing them for understanding or planning for controlled actions. These information to be treated are massive and modal. Moreover the real-world information has incompleteness and uncertainty because of its nature.

Thus the real-world computing requires the novel functions with flexibilities of various kinds. The aspects such as robustness, openness and real-time, reflect different attributes of flexibility. Therefore the key issue for novel functions is how to realize such flexibility.

Since flexible information processing intends to expand the abilities of information processing beyond the limitation of conventional one, expected application fields are quite wide, including for examples various recognition/understanding systems, flexible information (data/knowledge) bases, decision supporting systems, flexible problem solving, friendly man-machine interface, adaptive and large-scale simulation, autonomous and cooperative control, and so forth.

The objective of research and development is to investigate elemental novel functions (which serve as contact points with theoretical foundations) in the various application fields. Those functions are embodied in respective application fields or integrated to demonstrate new flexible information systems. The novelty of the functions should be emerged from new concepts of theory or algorithm suitable to the real-world computing, conquering the bottle-necks such as integration of symbol and pattern, learning, self-organization, giving new kinds of robust function so on. It is not desired to merely combine conventional technologies or to make ad-hoc systems for specified tasks.

There, the following two categories are conceived as the important directions of integrated systems:

- Real-world adaptable autonomous systems,
- Information-integrating interactive systems.

The former means flexible systems which can autonomously understand and control the environment through active interaction with the real world. The latter means flexible systems which support and enhance human capabilities through enlarged information channels between human and systems. The researches and developments on novel functions for realizing the systems will fall into the following categories:



- Flexible recognition and understanding of multi-modal information such as pattern information like images and speech sounds and symbolic information like natural languages,
- Flexible inference and problem solving based on flexible information bases which admit direct treatment of information and have capabilities of learning and self-organization,
- Flexible human interface and simulation including virtual reality for realizing mutual interaction between human beings and the real world,
- Flexible and autonomous control and integration methods to interact with real world environment.

Let us explain more concretely the relationship between the systems and the four categories of novel functions.

The first system means corporated systems between computing modules and the real world. The systems work for the purpose of partial replacement of activities of human beings in the real world. The systems are able to understand the real world and learn necessary concepts from the information of the real world for adapting new situations. Moreover they make actions to control the real world for both obtaining information and creating desired situations. These functions are required to cope with the uncertain, incomplete and variable characteristics of the real world.

The second system means another corporated systems between computing modules and human beings. Human beings are able to handle various information. They also used to create various information in which many kinds of their intentions are embedded yet in those kinds of ambiguous representation. Computing modules are required to both understand and integrate those informations flexibly to assist human beings for solving problems and creating new information. Then the systems should be regarded as a powerful tool to extend human intelligent activities.

Each system consists of its own fundamental and novel functions which are not revealed at the present time. Each novel function also is dropped into one of the four categories described. Some of new functions may be applicable to both systems.

In order to realize real-world adaptable autonomous systems, the following functions will be required:

- Understanding of scene images,
- Answer and question by spoken natural language between the autonomous information system and the real world model including human beings,
- Planning for action sequence and optimal control for adapting the environment,

- Integration of pattern recognition, incremental world model and control mechanism.

In order to realize information-integrating interactive systems, the following functions will be required:

- Understanding of intentions from various information produced by human beings,
- Realizing intelligent and interactive assistances to retrieve and present valuable information from a large amount of data in databases,
- Intelligent simulation to create new information findings and forecast transient state of the real world,
- Integration methods for combining human factors and the computational model of the real world.

The fundamental functions should be evaluated from viewpoints of both how they contribute for realizing these two systems and how they make breakthroughs in their own technological fields. The breakthroughs will be related with the aspects of real-world computing, namely, robustness, openness and real-time. The research and development program tries to discover new algorithms for both individual functions and integrated functions. In the following, we will list research themes which seem necessary to realize both systems. This does not necessarily mean that this program pursues all of these research themes listed.

The computational bases for realizing these novel functions will be massively parallel and distributed systems including neural systems or optical computing systems.

## **2.2 Research Topics**

### **2.2.1 Flexible Recognition and Understanding**

#### **1. Flexible understanding of images**

New paradigms for image understanding in the real world are needed for image processing for the cases of insufficient and inconsistent information. There are ill-definedness and ill-posedness on image understanding. Many researchers have dedicated their efforts to the study of recognition, understanding technology which is an ultimate means of converting distributed pattern information into compact representation expressed in language. They have studied various techniques to meet characteristics specific to subjects.

Early stages of image recognition/understanding (colours, shapes, position and movements) are defined clearly in the form of modules which may be formulated

as optimization problems, and should be integrated, for example, according to the constraint satisfaction paradigm. Systems need active conception for the acquisition of the structure of image data. Sometimes it is important to introduce constraints in images in order to improve the structurization. Image processing in this case includes the computation of constraint satisfaction.

Flexibility of image understanding emerges from the discovery of novel algorithms for segmentation of image including motion and integration of segmented objects in images by assuming the world model and using other information if possible.

Some important research themes are:

- Algorithms for learning, self-organizing and segmentation of image data,
- Integration of image features and conceptual symbols for understanding image,
- Construction technology for handling a dynamically changing world based on image understanding,
- Recognition of motion images of human beings for understanding human intentions based on flexible and deformable model for non-rigid object representation,
- Basic mechanisms for extracting intuitive informations from images.

## 2. Flexible understanding of speech

Speech is one of the important information in which flexibility of information processing needs to be realized. This information becomes a powerful one for computer users if the computer understands conversational speech which is speaker-independent. Flexibility of speech understanding emerges from the discovery of novel algorithms for extraction of dynamical speech features and integration of recognition modules including matching methods under the constraints of incomplete recognition. In order to obtain such novel function, the following research themes need to be considered:

- Early auditory processing and extraction of dynamic features of speech,
- Understanding mechanisms which can manage unknown words for the system,
- Simple and clear recognition algorithms suitable to be implemented by massively parallel computation.

Technologies should be developed to make interfaces between computers and users based on flexible spoken/synthesized speech in the uncontrolled real world. The following research themes are listed:

- Robust processing of syntax and semantics for symbol sequences (word output of speech recognition) with noise,
- Conversation models driven by speech and dynamic models of the world linked to conversational speech,
- Relationships between speech and other information such as vision which have complementary characteristics.

### 3. Flexible understanding of natural language

A very large amount of raw and coded data in natural language is produced in the real world. However, very little of it is transformed by means of machine understanding processing to be used for various purposes or in various situations. Novel functions in this field should be focused on the realization of robust understanding systems which can manage a large amount of coded natural language as raw data to extract a set of concept expressions useful for general purposes.

The key concept of novel function in this case is the segmentation of natural language so that many kinds of knowledge representations become robust based on the segmented units incorporating in conceptual and structural splitting and merging. At the same time, both learning and self-organization mechanisms are required to obtain up-dated knowledge units each of which may be regarded as an agent. The following research themes are listed:

- Robust parser applicable to a large amount of incomplete sentences,
- Methods for embedding electrical dictionaries of different kinds into natural language understanding systems,
- Self-organization algorithms for a sequence of sentences obtained from dialogue,
- Computational model to integrate knowledge units through interaction based on conceptual and structural distance to treat dialogue,
- Explanation oriented understanding of natural language,
- Interaction and complementary relationships between natural language and other information.

### 2.2.2 Flexible Inference and Problem Solving

#### 1. Inference and problem solving

Flexible inference deals with incoherent or incomplete data created in the real world. Most of data are expressed as symbols. However inference systems may use numerical expression like probability, likelihood value and so on. In this field, the knowledge representation problem becomes important because it is required to formulate problem statements for users based on incomplete requirements. Flexibility of inference will emerge from cooperation mechanism among inference agents, where each agent is able to carry out a specified task.

Many large systems in our social system, including production systems, have many different agents where cooperation makes global system stable, semi-optimal and adaptive. Problem solving is also an important function in the large system. Research to reveal these mechanisms from viewpoints of massively parallel and semi-uniform architecture is required in order to apply the various novel functions to the real world. The following research themes are listed:

- Acquisition of semi-global relationships or constraints and problem description frameworks based on incomplete requirements,
- Flexible algorithm of problem solving based on analogous or stochastic inference,
- Cooperation method between a plurality of inference engine agents or cooperation with human for realizing flexible inference,
- Problem solving of large systems under the constraint of limited communication channels.

#### 2. Flexible information bases and retrieval

Human beings are surrounded in daily life by different information, such as image, speech, text, tactile and so on; some of them are produced and controlled while others are not. Human beings seem to use available input data for information retrieval by integrating different information and using flexible inference function unconsciously. It is important to determine this mechanism in order to implement solutions by means of computer networks, since different data on different information are distributed at different locations in our society. These data become more valuable if they are integrated. The flexibility of information bases is strongly connected with self-organizing mechanism of them. The following research themes are listed:

- Knowledge representation which is common to or independent from each other,
- Self-organizing mechanism of large scale information bases,
- Learning algorithms to obtain cooperation between many kinds of information, and acquisition of correspondence to each other,
- Formalization of interaction between information and representation of data structure suitable to each information,
- Inference for information retrieval.

### 2.2.3 Flexible Human Interface and Simulation

#### 1. Flexible interface

Novel way of using the real world computing system must be developed in parallel with the development of individual novel functions mentioned above. The novel information processing environment is conceived to provide a human user with a wide range of intelligent and/or physical collaborative activities with the computer system. In such environment the human user is able to interact with the computer by such natural means as gestures, facial expressions, spoken language and to receive information by means of the real time 3-D images etc. Related research themes to realize such information processing environment includes the followings:

- Broad-band multi-modal interface (human body action, facial expression, recognition of linguistic and visual patterns, etc.),
- Information display system with virtual reality,
- Cognitive and behavior models to understand human intention by his gesture with his facial expression.

#### 2. Simulation

Real-world computing realizes novel simulation technologies which will provide a human user with a powerful tool to solve very difficult problems. Instead of doing expensive and time consuming physical experiments which may not be always precise, the computing system should be able to simulate a very large scale complex system and to predict its behavior in real-time. Prediction of untapped phenomena and future events, such as the earth environment and weather forecast, is important application of the real-world computing. Related research themes are:

- Learning/adaptation type simulation,

- Prediction and control of complex/chaotic time series.

#### **2.2.4 Flexible Autonomous Control**

Research is needed to focus on the development of technologies necessary to realize a flexible autonomous coordinated system operating in the real world in real time. A robot is a typical example. Application also includes the aids for the elderly or physically handicapped people. Such a system is composed of various function modules interacting each other for perception, decision and action. Problems to be solved are how to integrate these various functions and how to control interactions among function modules in order to achieve desired goal state under the constraints of real world: incompleteness and ambiguity of available information, dynamic change of physical environment, limitation of available time and space and so on. Followings are related theme of research.

- Flexible modeling of environment/task/control,
- Active, distributed sensing and sensory integration,
- On-site planning and distributed cooperative search,
- Structurization and coordination of multi sensing/planning/action agents for real time skillful manipulation of objects,
- Maintenance of consistency between internal world model with the dynamic real world.

### **2.3 Research Schedule**

The research schedule will be divided into two stages.

In the first stage, researches for discovering algorithms will be pursued for realizing elemental novel functions necessary to the integrated systems. The research milestones of flexible recognition and understanding will be the establishment of recognition of motion images including finger motion and gesture, and the establishment of understanding of conversational speech. One of the research milestones of flexible inference and problem solving will be to develop information bases of specified knowledge domain including image data to which new ideas of flexible inference and problem solving applied for both developing and innovating where the new ideas will include new types of algorithm for solving statistical inference problem and constraint satisfaction problem. Another milestone will be the discovery of method for self-organizing a large amount of information bases. One of the milestones of flexible human interface and simulation will be to reveal relationship between human intention and broad-band information

of human beings by combining information processing mechanisms and psychological evidences and experiments. Another milestone will be the establishment of display methodology including virtual reality for representing time-variant situations created using recognition/understanding results. The milestone of flexible autonomous control is the discovery of the principle methodology to integrate sensing, perception, planning and action in the real-world from the viewpoint of adaptation/learning.

In the second stage, researches on integration of elemental novel functions discovered in the first stage will be pursued for the purpose of realizing integrated systems called real-world adaptable autonomous system and information-integrating interactive system. One of the research milestones of the second stage will be the realization of real-world computation systems which work in real time. The intelligent and sophisticated resolution of combinatorial problem will be pursued for integrating individual novel functions developed in this program. Then each novel function in an integrated system will cooperate with others in order to create new novel functions in connection with robustness and openness of the system.

Researches for integrating elemental novel functions may be pursued partially in the first stage, while the first stage dominates researches for elemental novel functions and vice versa in the second stage.

Each research includes such activities as trying to discover new algorithms which will stimulate the related researches for creating new paradigm suitable for the real-world computing where theoretical foundations of new algorithm are not always required at the stage of discovery. The research then establishes new technologies of the real-world computing based on both implementation schemes and theoretical foundations of novel functions.

However each research theme should be reviewed after five years. During the review, each function will be evaluated on whether how it will be implemented or reflected in the parallel machine developed in the program or what breakthrough it will make in its own technological field related to real-world computing.



## **3 Massively Parallel Systems**

### **3.1 Research Framework**

Real-world computing requires a computation framework which can process various kinds of information flexibly and with integrity. Some novel functions for real-world computing applications are flexible recognition and understanding, flexible inference and problem solving, flexible human interface and simulation, and flexible autonomous control. A system which implements one of these applications is likely to consist of many modules, which can exploit parallel and distributed processing at several levels both within and between modules.

Several parallel paradigms applicable to the above applications have been proposed, including , concurrent object oriented, dataflow, and data parallel, neural network, probability based information processing. Real-world computing will probably be realized by some combination of these paradigms. In such a flexible information processing system, there are many levels of parallelism to be exploited. These new paradigms are naturally adopted to a massively parallel system and they require a huge amount of computation to solve practical problems within a reasonable amount of time. These observations show that a massively parallel system is necessary to support efficient computational power of real-world computing and it must also be a system which can efficiently execute the multi-paradigm.

The massively parallel system which is the computational base of real-world computing should be flexible in itself. The system flexibility comes in part from hardware, in part from architecture and in part from software. Hardware robustness and reliability are essential. It should be general-purpose from an architectural point, because a massively parallel system should support multi-paradigms efficiently. Software will be needed to realize the functions of adaptation, self-organization and optimization. The system is one that can adapt itself to application environments for optimal performance. This flexible system will deliver maximum performance while minimizing the burden on the users. The goals of general-purpose massively parallel systems research in this program concern the development of architecture, operating systems, languages, and system environments of massively parallel systems.

### **3.2 Research Topics**

#### **3.2.1 Massively Parallel Architectures**

The following are fundamental technologies which should be pursued in the development of general-purpose massively parallel systems.

#### 1. Model

Flexible execution models which can be bases of general-purpose architectures should have the ability to fill the gaps between the language models and hardware. Flexibility which allows a mapping of a virtual computer onto actual processing elements should also be pursued.

#### 2. Architecture

The architecture must support various paradigms efficiently. The massively parallel system should be based on a general-purpose architecture which does not focus on any special applications. It is also important to study hardware architecture in consideration of future device technology and packaging technology for an efficient implementation of massively parallel system.

#### 3. Interconnection network

The interconnection network should provide high speed communication which is comparable to computation speed. It should also provide support for dynamic load distribution, global synchronization, global priority control. In implementing the high speed interconnection network system not only silicon technologies but also optical technologies should be considered.

#### 4. Robustness/Reliability

Hardware oriented robustness which can tolerate expected component failures in massively parallel systems should be examined. System components should have self checking and self repairing facilities. The total system should have a maintenance architecture or facilities to maintain system reliability.

The system to be developed in the Real-World Computing Program should be a general-purpose massively parallel system. The general-purpose nature may be achieved by mapping the multi-paradigm model to the hardware execution model in software. If the hardware can be modified and adapted to a variety of applications, it may be considered as general-purpose. Since a huge number of modules or processes must be processed concurrently in real-world computing, architectural mechanisms which can support rapid process switching and cheap synchronization should be implemented as an integral component of the system.

In the first half of this program, a prototype system which consists of  $10^4$  processing elements should be developed. Several hardware prototype systems are developed in this stage. One of these prototypes should be designed and developed as a platform. A platform system is used for tools of software development and a research platform on novel functions. Fundamental researches on massively parallel model and architecture are also concurrently studied. In the second half, a massively parallel system to be

developed is expected to have the order of  $10^6$  processing elements. It will have the ability to execute various kinds of real-world computing applications at real time speed. The architecture will be based on the new massively parallel computing model to be studied in this program.

### **3.2.2 Operating Systems for Massively Parallel Systems**

The primary purpose of the operating systems are to assign computer resources to processes for efficiency and to build a user-friendly software environment behind which hardware details are hidden. Like operating systems for a serial computer, operating systems for massively parallel systems must be general-purpose, interactive, multi-user parallel systems.

The operating system for a massively parallel system should be designed to support the execution of various processes concurrently with high throughput. To administer massively parallel processors, the operating system may have a hierarchical structure much like human society. Support for the simultaneous execution of a number of parallel programs requires the operating systems to coordinate the partitioning of resources among the programs, as well as any dynamic repartitioning at run-time. The locality, the concept of grouping activities and the principle of balancing are guidelines for partitioning.

One goal of an operating system is to provide the programmers a higher level of abstraction of a general-purpose parallel systems in order to make parallel programming easier. Through the abstraction, the operating system manages the resources and coordinates several user programs according to the above guidelines.

The research issues of an operating system for the massively parallel system are as follows:

1. Hierarchical structure

To realize a functionally distributed management system for flexible processor management, the operating system may require a hierarchical structure. Hierarchical structure makes the system scalable for a massively parallel system. It is required that some efficient mechanisms to control activities and its hierarchical structure in terms of hardware and software. The reduction of overhead for controlling parallelism or executing critical sections should be considered.

2. Network management

Advanced intelligent routing, addressing, synchronization, deadlock prevention, flow control, and failure avoidance should be incorporated into a flexible network management system.

### 3. Resource management and load distribution

In the massively parallel system, the elimination of synchronization overhead, access contention and communication overhead will become more serious issues. To overcome these problems, the operating system should be able to collect management information autonomously and undertake statistical management or adaptive management. Memory management and virtual systems for several resources should also be pursued for efficient scheduling and load distribution. For example, paging and virtual memory are the most successful ideas in memory management for a conventional computer system. It seems reasonable to expect these benefits in parallel systems. In order to balance load among processors, it is sometime desirable to migrate a group of processes to other set of processors. Efficient hardware support is needed to support process migration efficiently as well as the operating system facility.

### 4. Fault tolerance

In the massively parallel system, resource management should be carried out in a manner which allows for expected component failure rates. Therefore, it is necessary to handle the failure avoidance system as a normal process. Multiple-route processing will also be required in order to deal with failures. In order to manage the massively parallel system efficiently, advanced intelligent routing and failure avoidance should be incorporated into a flexible network management system.

## 3.2.3 Languages for Massively Parallel Systems

The language for the massively parallel system must be able to describe the coordinated operations of a number of processes. Several languages based on several paradigms should be considered. The problem is how to extract the available parallelism in the problem domain, and be able to execute it with as much parallelism as the underlying system can provide. The user programs for such a system will be very complex. Development, verification, testing, debugging, and maintenance of large-scale parallel programs will be far more complex compared to the sequential counterparts. Various compilation techniques and run-time implementation techniques scalable to nearly one million processors should be studied. Design and prototyping of the language model and creation of the high level language models should be done. Some high level languages which can be used for practical applications should be implemented. The following items should be considered.

#### 1. Language model

Language model is a description model for flexible programming languages of

massively parallel systems. The language model can provide a basic programming abstraction for the underlying architecture. The model must be simple and be sufficiently close to the underlying architectures so as not to restrict their computing power, and at the same time provide powerful means of abstraction to promote software programmability, portability, and reusability. Here, one viable candidate is the object-oriented concurrency model. In the research of language model, the fundamental research on supporting flexible language, model of describing coordinating and cooperating actions, inheritance and reflections should be pursued. Inheritance serves to adapt the objects to a certain computing environment. Recent studies have provided deep theoretical foundations to object-oriented concurrency with regards to the concurrent behavior as well as their inheritance behavior. The idea of reflection in a language is to make the representation of the underlying program execution scheme manipulable within the language. Reflection allows the parallel system to encompass scheduling, communication, and load-balancing within the programming language framework.

## 2. High level languages for massively parallel systems

The primary goals of high level languages for massively parallel systems should be ease of programming and also the ability to describe computation on the scale of a million processors. Several languages which will be based on multiple parallel paradigms will be considered. One viable candidates of such languages will be an appropriate amalgamation of concurrent object-oriented, functional, and declarative constraint-based approaches. This amalgamation will create layers of programming languages. The common base language make it possible to provide the basic programming abstraction of the several high level languages. Efficient inter-layer compilation techniques must be developed so as not to preclude the use of high level languages. Support for speculative computation within the parallel constraint satisfaction framework will be also exploited. As for the concurrent object-oriented languages, currently available object oriented models are not intended to process more than million processes, so the following extensions will be needed.

- Introduction of a description system permitting hierarchical decomposition of complexity,
- Diversification of message propagation systems,
- Introduction of self-reflection functions for adapting and evolving objects,
- Declarative description of object relationships.

### **3.2.4 Environments for System Development and Programming**

#### **1. Programming environment**

Multi-paradigm can give users the freedom to choose among various programming languages. The need to develop a programming environment that can support multi-paradigm programming is expected. Tools for debugging, graphically monitoring and analyzing load balance, communication characteristics, etc. will also be required. Since these functions may need hardware support, the architectural design should take these requirements into consideration.

#### **2. System development support environment**

The requirements of an environment supporting the development of the massively parallel system include two features different from conventional ones. One is the support for the interconnection network development. In the development of the interconnection network, the requirements and limitations of several goals should be examined, including robustness, dynamic load distribution, and global synchronization mechanisms. The overall functions and performance of the interconnection network should be evaluable in advance by system-level simulation. The other is the support for the architecture development of processing element. In the development of an architecture for the processing elements of the massively parallel system, a set of basic functions for processing elements should be determined through a functional assessment of the various subsystems, including the interconnection network.

### **3.2.5 System Evaluation**

The application fields should be picked to verify the effectiveness of massively parallel computational bases for real-world computing. The simulations that are dedicated to the prediction of future events from established models are examples for this purpose. These simulations can predict macro-behavior by describing interactions between micro-elements. Massively parallel systems are expected to handle situations such as non-linear or many-body problems by direct mapping paradigm. They will also be expected to handle simulations where the governing equations are difficult to formulate because of the complexity of the phenomena or of the incompleteness of our knowledge to describe the phenomena stochastically.

## **3.3 Research Schedule**

The final massively parallel system to be developed is expected to have the order of a million processing elements. However, it would be risky to develop such a large

system in one step; prototyping is therefore essential. The prototype systems will be constructed in the middle of the program. The development technologies used will be intermediaries between the currently available ones and those required for the final systems. The progress of semiconductor technologies must be watched closely to set up the development plan. A platform system which is used for tools of software development and a research platform of novel functions should be constructed. One of the prototypes will be designed and developed as a platform. The software development on this platform system is very important for using the platform system as computational bases in the program. The final systems is the integration of the massively parallel systems and the neural systems.

The research and development of massively parallel systems is divided into two stages as follows.

1. The first stage

Several hardware prototype systems are developed in this stage. One of these prototypes is called platform system which is used as a platform for software implementation and evaluation. Prototype systems should be designed and developed as intermediate goals of the final massively parallel systems. At the end of the first half of the program, these prototype systems including a platform system are evaluated from the standpoint of designing final systems. Fundamental research on massively parallel model and architecture are also concurrently studied. The establishment of the concept for languages and operating system, and prototyping on the platform system is pursued in this stage. Design and implementation of the common base language is done first. The research of the language model is also pursued to provide means of abstraction of high level specialized languages. Some high level languages and operating systems which can be used for practical applications in the second stage should be implemented on the platform system. Various compilation techniques and run-time implementation techniques will also be studied on the prototype systems. Also, a new user programming environment must be created to facilitate millions of concurrent objects executing simultaneously. We must provide special purpose high level languages which will accommodate the users' needs.

2. The second stage

The final system (or systems) are designed in the second half of the program based on the evaluation of the prototype systems which is developed in the first stage. The bases of the final system are selected from the prototype systems. The important technologies which are developed in the first stage are integrated into the design of the final system. The second half of the program is dedicated to

developing the massively parallel system based on the massively parallel model which will be studied in the first stage. It will also be based on the results of research on theoretical foundations and novel functions in the first half of the program. The final system and the neural system will be integrated into a integrated information processing systems. As for the software research, the second half of the program will focus on the implementation of the massively parallel operating systems and language systems. The extension of the high level language models is also pursued. Optimized compilation techniques is implemented on the platform and its prototyping is studied on the final system. Furthermore, in order to harness the complexity, hierarchical decomposition methodology of large-scale parallel programs and analysis framework should be studied.

Many researchers in this program should be able to use the platform system from their respective research institutes. This will be accomplished by connecting the research institutes and the platform system through high speed global network. The research and development of the total structure on such network environments should be pursued. In addition, the system software should support a standard user interface and a currently used system environment. The installation of small platform systems with a smaller number of processing elements at each research institutes will be considered. The ability for many researchers to work in common circumstances, giving them common concepts and technologies for real-world computing, is very important in the promotion of this program.



## 4 Neural Systems

### 4.1 Research Framework

The brain performs certain sorts of computations many times faster than the fastest digital computer, despite the fact that the processing speed of neurons are five or six orders of magnitude slower than that of silicon logic gates. Neural systems are systems that are deliberately constructed to use some of the organizational principles that are used by living creatures. We expect that a brain-like neural system would perform special types of information processing at high speed and generate intelligent behavior.

Characteristic of information processing using neural networks, is that many simple processing units function according to the principles of "cooperation and competition" while maintaining close information exchange among them. In addition, since the programs that are used in conventional computers are expressed by the connection strength, the system can perform learning and adaptation by making use of conventional numerical optimization techniques. All these features make neural systems a promising platform for flexible information processing.

Recently various neural network models have been proposed and demonstrated to be suitable for pattern processing, finding approximate solutions to optimization problems, for constraint satisfaction processing, etc. However, these networks have been restricted to small-scale applications. The neural models which were used in such applications were very simple and the learning method used was almost exclusively the back-propagation technique. The implementation of flexible information processing is a difficult problem compared with past applications. We have to explore new abilities of neural networks : new neural models, learning methods, self-organization, etc.

Today, neural networks are usually built and simulated on conventional computers. However, simulation speed is very slow, particularly with large networks. The problem may be overcome by using general purpose massively parallel computers, but considering the cost and integration scale, special purpose hardware will be desired.

In this program, we would like to exploit the abilities of large scale neural networks in order to implement a flexible information processing system which can work in real-world domains. The goal of this program is to construct a neural system which consists of hardware and software systems based on novel neural models. The development of neural models is one of the important themes of this program. The models must support large scale neural networks which can acquire knowledge or adaptively change their structures by interacting with external environments. The hardware system is expected to support a one million unit neural network and attain 10 TCUPS (Tera Connection Updates Per Second). The final neural system will consist of special neural hardware, its software, and cooperate with the massively parallel system for flexible

information processing. The integration mechanisms with the massively parallel system should be developed by cooperating with the research groups of the massively parallel systems.

## **4.2 Research Topics**

### **4.2.1 Neural Models**

We are going to develop new neural models which realize flexible information systems. The flexible information system will be implemented by large scale neural networks which adaptively change their structures by interacting with the real world. The network grows dynamically by supervised learning or self-organization in a real-time environment. However, the network should not forget existing knowledge when it grows. In order to satisfy these requirements, the network should be structured to have the following characteristics: modularity, hierarchical structuring, and differentiation. Realizing these characteristics will require research in the following fields:

1. Neuron unit models

To date, neural network research has mainly used simple neuron unit models, and has succeeded only in limited areas of application. The development of more sophisticated neuron models will lead to generate functional intelligence. For example, chaotic neuron models, complex-number neurons, and neuron logic models (in which neurons possess logical properties) are proposed as candidates for sophisticated neuron unit models. We have to evaluate their capabilities and develop new neuron unit models.

2. Modularity and hierarchical structure

Large scale neural networks of the real world must interact with the external environment, and change their network structure by adaptation or training. In such a situation, new knowledge should be acquired without the destruction of existing knowledge, and information should be efficiently retrievable. In order to meet these requirements, modularization and hierarchical structuring of knowledge should be developed. We have to develop a distributed learning method and learning control method for modularization, structuring, and differentiation. The following mechanisms should be considered:

- Construction of differentiation and hierarchical structure by adapting to the external environment,
- Autonomous growth of neural networks by interaction with the real world,
- Modularization using centered or distributed-control,

- Interaction between modules,
- Evaluation methods in modularization and hierarchical structuring.

### 3. Learning and self-organization

In real-world computation, real-time learning is very important. The development of learning methods which can add new knowledge without destroying old knowledge, and which can delete existing knowledge without loss of consistency is important. Many large scale neural networks contain recurrent loops. Recurrent neural networks are promising for temporal pattern recognition and temporal pattern generation, while layered neural networks are effective in spatial pattern recognition tasks. Also, recurrent neural networks can be used to solve optimization problems because of their constraint satisfaction capabilities. Recurrent neural networks will play an important role over the next ten years. Learning methods and self-organization mechanisms for recurrent networks must be established.

Network topology and/or size are two of the most critical parameters in determining generalization performance. We have to develop learning procedures that will construct optimal self-organized networks. Such networks should be sufficiently large to learn the problem, and small enough to generalize well.

### 4. Associative memories

Memory which can be implemented by neural networks has interesting characteristics. Information such as spatial or temporal patterns can be stored in a memory using a distributed representation. This stored information can be retrieved, based on the closest match mechanism. Theories of memory mechanisms such as memory capacity, topological structure of memory etc. are important research themes. We have to clarify the principles of association, and develop engineering mechanisms which implement these functions.

### 5. Analog computation models

Information processing of neural systems can be supported by their analog non-linear dynamics. From this point of view, we have to clarify new principles of analog computing in neural systems including chaotic dynamics.

### 6. Integration of models

Models which integrate different paradigms should be investigated. For example, the integration of neural networks and logical processing (e.g. using a soft logic inference model), and the combination of pattern processing and symbolic processing may be required to implement the neural system. The integrated model

should be run on a parallel processing environment. Input/output representations are interface with other models or external environments.

The choice of representation of input/output information to/from a neural network strongly influences the processing capabilities. Therefore theoretical and experimental research on input/output representations to/from neural networks is critical for the development of the neural system.

#### 4.2.2 Hardware for Neural Systems

The neural system should support a one million unit neural network which interacts with the external environment and changes its structure adaptively. Such a large neural network may be modularized and consist of sub-neural networks, each of which has a thousand neurons that are fully interconnected.

In order to satisfy such requirements, the final hardware system must achieves 10 TCUPS processing speed. In designing the special hardware, it should be noted that the model of neural networks has not been clarified yet and that various other new models would be most likely to be announced in the future. The hardware will also be required to process a wide variety of neural network structures such as layered type and recurrent type. General purpose and scalable mechanisms should be embedded in the architecture. Therefore various architectures should be examined in the early phase of this program.

The neural hardware can be classified into the following three types:

- Neuro-accelerator,
- VLSI neuro-chip,
- Engineering implementation of neural networks.

The Neuro-accelerator is a special purpose parallel processor which is tailored for neural network processing. Though many architectures have been proposed, the typical structure consists of hundreds of processing elements and achieves 1 GCUPS. The VLSI neuro-chips consist of hardware neuron unit(s) which simulate neuron unit models. The spectrum of neuro-chip architectures is wide, ranging from the digital processing chip to the analog processing chip. The third approach is an engineering implementation of neural networks. It uses hardware logic to simulate the functions which are realized by neural networks. It is impossible to compare them, because each chip has its own features such as learning capability, scalability, etc. In the neural system, we will focus on the second and third approach.

The following options could be considered.

1. Digital circuits neuro-chip

Digital circuits offer substantial advantages in noise tolerance and high accuracy processing, and are suitable for stable operation of large scale systems. The manufacturing technologies for computers can be directly applied to neural digital circuits. In addition to the conventional approach, variations are proposed such as the pulse-density model. These approaches may give new capabilities to the neuro-chip.

2. Analog circuits neuro-chip

Analog systems save hardware volume because operational circuits are fewer. It is attractive for the development of large-scale networks. Also, analog circuits can process dynamic and other complex neural networks such as chaotic neurons.

3. Hybrid neuro-chip

We can consider the digital-analog hybrid chips which make use of each of these advantages.

4. Engineering implementation of neural networks.

The neuro-chips mentioned above implement the neuron unit model directly. There is however another approach to realizing neural network functions (e.g. association) by using logical hardware, which does not simulate the neuron unit models. Such systems could be constructed by the repetition of simple logical circuits.

In a neural system, all the units basically exchange their activation values. Therefore the interconnection network between processing units is an important design point. Time-multiplexing and frequency-multiplexing methods will be candidates to eliminate the problem. The design of the interconnection network is closely related to the implementation technologies. The interconnection band-width is restricted by the pin number of the neuro-chip and the board. Wafer scale integration, three dimensional VLSI architecture, and optical interconnection will be important technologies. The development of these technologies is valuable if the technologies are commonly used in this program. Tools such as libraries of standard neuron cells will be required for neural ASICs. Silicon compilation of networks is important for embedded applications.

#### 4.2.3 Software for Neural Systems

In the research and development of neural systems, various neural software systems will be necessary.

1. Simulation systems

Simulation systems for neural networks are important tools for neural network

research and development. Most neural network researchers have their own simulators. A flexible general purpose neural simulator for large scale neural networks will be a powerful tool. The requirements for such a simulator would be machine independence, expandability, convenient user interface, high speed processing, and variety of utility routines. Analytical tools should be integrated into the simulator, to describe the algebra and geometry of neural learning and to permit transformations of the network dynamics in order to achieve better convergence and/or cognitive performance.

## 2. Neural network languages

Neural network processing should be programmed by high level languages. The following research and development themes will be required to design such languages.

- Representation of ambiguous information,
- Description of best match operations,
- Integration with logical programming,
- Integration with simulators.

## 3. Operating systems

In general, the number of hardware neurons will be smaller than the number of units in a neural network. Therefore, a virtual mechanism to fill the gap is important.

- Mechanisms for mapping neural network into hardware,
- Resource scheduling.

### 4.2.4 Integration with Massively Parallel Systems

The neural system may consist of various neural subsystems. However, the neural system will be a kind of special purpose processor for pattern recognition, associative memory, optimal combinatorial search and other special functions. This means that the neural system must be connected or integrated with other systems. Various integration mechanisms can be considered, from the closely coupled to the loosely coupled cooperative architecture. The following integrated mechanisms can be candidates:

- Closely coupled integration: neural networks are connected to each processing element of a massively parallel system as an associative memory.
- Loosely coupled integration: a massively parallel system sends a special job such as an optimization problem, to the neural system as a dedicated system.

#### 4.2.5 System Evaluation

Neural theories and models are evaluated according to whether the functions listed in section 4.2.1 are satisfied or not. The neural hardware is evaluated using two methods. The first one is the evaluation of processing speed (TCUPS). And the second one is the evaluation of the execution of real applications which are developed in this program. In the benchmark application, flexible and high speed information processing should be realized. Flexible motion image processing and information integration by multi-neural networks are candidates of the benchmark program.

#### 4.3 Research Schedule

The neural system development has two stages. Each stage is five years long. In the first developmental stage, models and hardware architectures are researched. Various kind of models should be proposed and examined in the first stage. Some models which are selected by the evaluation will be candidates for the second stage. Their characteristics will affect the design of the final system architectures.

A hardware prototype system will be constructed for the second developmental stage. The hardware system is expected to have ten thousands processing units, and will be a platform system which will be used for researching neural models and novel functions. The technologies used in the first stage will be intermediary between those currently available, and those required for the final system.

The second stage is the construction of the final system. Neural models and architectures will be refined, based on the results of the first stage. The models have to work efficiently on the final system. The implementation methods of models on the neural hardware should be researched in this stage.

The first hardware architecture prototype will be used and evaluated for further research and development. A final neural system which has the order of one million neuron processing units will be designed and constructed by using new hardware technologies and other results of the program. The final system will be integrated with the massively parallel system. The integration mechanism should be a most important aspect of the final system architecture. In particular, new integration mechanisms should be discussed to construct the final integrated system in the early years of the latter stage. Some application systems will be implemented on the system, and the total system should be evaluated.

## 5 Optical Computing Systems

### 5.1 Research Framework

Light is expected as a new information medium because of its extended transmission capacity and massively parallel processing capability. Optics provides new device technology as well as new architectures and algorithms in the Real-World Computing Program which aims at flexible information processing by massively parallel and massively distributed processing.

Optical technology to be developed in the program will be classified into three categories.

- Optical interconnection,
- Optical neural systems,
- Optical digital systems.

Optical interconnection aims to overcome so called wiring limit which electronic systems are now facing. Optical interconnection is also the key technology for realizing optical neural systems and optical digital systems.

Optical neural systems aim to realize real-time learning and associative processing of images and other distributed data by connecting neurons with light.

Optical digital systems aim to realize massively parallel processing with computational accuracy using light.

Optical computing technologies presuppose the development of optical devices.

Modularization of optoelectronic devices is also an important goal of the Real-World Computing Program. Optical enhancement into highly parallel and massively distributed systems will be verified with the modules.

### 5.2 Research Topics

#### 5.2.1 Optical Interconnection

Optical interconnection merges advanced electronics technology represented by VLSI with optical communication technology to eliminate information transmission problems in electronic systems, such as propagation delay, line to line crosstalk, space factors of wiring and mounting, and large power consumption.

Optical interconnection aims to break through the above limits and offer a high-speed, large capacity and flexible information transmission.

In order to develop optical interconnection, the following issues are important.



## 1. Optical interconnection devices

Optical interconnection devices include optoelectronic active and passive devices. The combination of these devices realizes high-speed, large capacity and reconfigurable interconnection networks, using high-density multiplexing technologies in the area of time, space and wavelength.

Ultrafast optical interconnection devices aim for sub-picosecond high-speed devices to achieve large-capacity interconnection by using time division multiplexing(TDM).

Space-parallel and/or functional optical interconnection devices aim to implement reconfigurable high throughput optical interconnect, using space division multiplexing(SDM) technology combined with their various electro-phonic functions.

Wavelength-parallel and/or functional optical interconnection devices aim to achieve large capacity optical interconnect by using wavelength division multiplexing(WDM) technology. A high capacity signal bus is realized by WDM technique.

Passive optical interconnection elements including micro-optics and diffractive elements aim to develop optical components having advantages of both stability and high-density optical interconnections.

Advanced optoelectronic integrated devices and circuits combining different material systems and/or functions aim to achieve compact devices for the next-generation following the aforementioned devices.

## 2. Optical interconnection network

Establishment of architecture and design technology aims to realize system design of compact, high-speed, and multi-channel optical interconnection networks having such flexible functions as reconfigurability and selfrouting.

Interchip and intrachip optical interconnection aim to develop high-speed and flexible optical interconnection network between processing elements, between processors and memories, and between memories.

Modularization of optoelectronic devices and passive optical elements aims to realize integration and miniaturization of optical interconnection components, which eliminates the problem of high-precision alignment.

### 5.2.2 Optical Neural Systems

Optical neural systems realize the real-time processing of images and other spatially distributed information or spectral information through learning and associative pro-

cessing using massive and flexible interconnectivity of light.

In order to develop such systems, the following issues are important.

### 1. Optical neural models

Direct image processing model aims to develop a model which permits direct input of 2-D and/or 3-D image information into the neural network and process it.

Model based on the physical phenomena of light aims to develop a novel type of theoretical model using physical phenomena of light such as bistability, chaos, phase-conjugation, etc.

Expandable modular model aims to develop an expandable, flexible model consisting of a number of unit modules.

Model suitable for analog devices aims to develop a model intended for implementing optical analog devices which are low in accuracy but excellent in large scale configuration and high speed.

### 2. Optical neural devices

Optical neural devices with large-scale, high-speed learning functions aim to develop large-scale optical array devices which can vary synaptic connection weights according to electric and/or optical learning signals.

Optical neural devices for image processing aim to develop optical devices for direct image recognition and processing, and devices for extracting features of input images.

Modularization of optical devices aims to realize modularization and standardization of optical neuro-chips.

### 3. Optical neural systems

Establishment of design technology aims to establish design methodology for the distribution of functions and hierarchization of the system, and to realize accurate processing through system integration with digital computer.

Establishment of learning method aims to develop technologies for acquiring the knowledge from the training signals, storing them as structured knowledge, and technologies for increasing learning speed.

Realization of human friendly I/O interfaces aims to develop technologies for direct processing of multimedia information, and also to develop image database allowing direct search of images by key image.

### 5.2.3 Optical Digital Systems

Optical digital systems realize the processing of images and other spatially distributed information or spectral information with logical computation principle using massive and flexible connectivity of light.

In order to develop such systems, the following issues are required.

#### 1. Optical logic devices

Ultrafast optical logic devices aim to develop high-speed binary and/or multi-valued optical devices, and their two dimensional integration with low power consumption.

Space-parallel optical logic devices aim to develop space-parallel optical devices with encoding signals in the form of spatially coded pattern.

Wavelength-parallel optical logic devices aim to develop wavelength-parallel optical devices with encoding signals in the form of a combination of light with different wavelength.

Passive optical devices aim to develop micro-optics, planar optics, diffractive optics and high-precision optics.

#### 2. Optical logic circuit

Optical interconnection between optical logic devices aims to develop the method of reconfigurable optical interconnection network and its instrumentation.

Optical functional modules aim to modularize optical logic devices, parallel optical registers, parallel optical memories, optical crossbar switches, and optical I/O units.

#### 3. Optical digital systems

Establishment of architecture and design technology aims to design reconfigurable optical parallel computers and to establish CAD technology for 3-D optical circuit design.

Input/output interface aims to develop physical and logical interface technologies which realize high-speed exchange of data between optical logic circuits and electronic systems.

Implementation technology for optical modules aims to develop implementation technology of optical functional module array, and integration technology of different functional modules.

Programming language and compiler aim to develop higher level programming languages and compilers for optical digital systems compatible with those for electronic computers, and compilers.

#### **5.2.4 Environments for System Development**

OEICs are key devices for optical interconnection as well as optical neural system and optical digital system. OEICs developments should be based on the common platform of processing and module technologies to yield the standardization for interchangeability in system.

The following subjects are required:

- Advanced OEIC processing technology,
- OEIC module technology,
- Standardization and CAD technology

Advanced OEIC processing technology aims to develop fundamental technologies for wafer growth and device process which realize OEICs with high-performance, reliability, and productivity.

OEIC module technology aims to constitute a module in a package including optical interconnection/neural/digital devices, optical passive elements, electrical circuits, and the input/output components.

CAD technology aims to develop computer-aided tools for designing OEICs structure, and for simulating characteristics of OEICs and their packaging module.

Standardization aims to standardize opto-electronic and mechanical I/O interface of OEIC module.

### **5.3 Research Schedule**

The development of optical computing systems has two stages. Each stage continues 5 years.

In the first stage, the basic device technologies for optical interconnection, optical neural systems, and optical digital systems will be developed. The standardization and modularization of optical elements are also the matter of research. The architecture study of small optical computing systems will be conducted.

In the second stage, we see some degree of progress in device technology, and will start the development of prototype optical computing systems. The large scale integration of optical devices will start, and optical modules for real-world computing

will be available at the end of the program  $Tb/cm^2sec$  order of space throughput will be expected.



———— 禁 無 断 転 載 ————

平成4年3月発行

発行所 財団法人 日本情報処理開発協会

東京都港区芝公園3丁目5番8号

機 械 振 興 会 館 内

TEL (3432)5405

印刷所 株式会社 昌 文 社

東京都港区芝5丁目2番30号

TEL (3452)4931







