P-1409-C

VOLUME 5

NATURAL LANGUAGE AUTOMATED PROCESSING AND ARTIFICIAL INTELLIGENCE: A PROGRAM PLAN

COGNOS

1984 FEBRUARY 29

DSS FILE 12ST.41414-2-03X2

NOTE

This report, A Program Plan, is the fifth volume in a series of reports addressing Natural Language Automated Processing and Artificial Intelligence. The reports in this series are:

- The State-of-the-Art
- Implications of New Technology Thrusts
- Possible Social and Economic Impacts
- Opportunities for Canada
- A Program Plan, and
- Priorities for Canada

ACKNOWLEDGEMENTS

The material in this report has been assembled on behalf of the contractor by Nordicity Group Ltd. It is based on input from various members of the study team and includes information supplied by numerous outside sources contacted through interviews.

Executive Summary

In this report a broad framework for a major national thrust in Artificial Intelligence is presented. The framework springs from a review of the state-of-the-art, coupled with a judgement as to the specific opportunities that present themselves to Canada in a global environment of rapidly increasing competition.

A national thrust is vitally necessary because AI technology lies at the heart of competitive competences in the age of informatics and automation. Also, a national thrust involving sharply accelerated R&D effort in all sectors of AI can be achieved without the creation of major new institutions, with all that this implies in terms of lengthy deliberation as to responsibilities, and revised budgetary dispositions. Rather we have opted (as is the case in the United States) for rapid program launches through existing mechanisms, coupled with advertised prioritization by goal and target settings. Only minimal institutional changes are advocated and these are designed largely to speed the flow of government demand to the private sector, and to ensure the coalesence of scarce resources.

Key report recommendations are:

- the federal government should advertise the priority of AI research and development and should assume a lead role;
- coordinated responses should be developed by groups such as the Canadian Business Opportunities Group, an Inter-departmental Committee on AI (to be established), government research laboratories (NRC, CCRS & CRC) and others;

- NSERC and DOC should sponsor a Canadian computer network modelled along the lines of the US ARPA network;

- several Centres of Excellence should form the core of the national effort in AI;

- MSED and MOSST should establish goals and targets for increased national spending in AI;

- proposed overall spending targets are:
  First Five Years $245 million
  Year Six to Ten $935 million

- private sector financing should be 30% of annual expenditures by year 5 and 67% by year 10;

- interaction with international groups should be encouraged and supported at all levels;

- key areas for research programs include:

  - machine translation;
  - natural language processing with particular reference to management information and office automation systems;
  - lensing technology (vision systems) for robotics applications;
  - expert systems for medical and educational applications; and,
  - image analysis with respect to satellite-related resource programs.

- suggested areas for development programs include:

  - semi-automated machine translation systems;
  - improved computer-aided learning systems;
  - enhanced image analysis systems for resource surveillance;
  - expert systems for medical diagnosis, natural resource management and telecommunications; and,
  - intelligent command and control systems for military applications.

The urgency of the competitive situation in AI is such that we must start now with specific funding action. Scarce scientific manpower will not stay in Canada unless a domestic magnet is created. Action can proceed in parallel with planning. It is our sense that a race of crucial economic and technological importance has already started.

IMPLEMENTATION PLAN

TABLE OF CONTENTS

IMPLEMENTATION PLAN

1.0 Introduction

The Opportunities for Canada Report submitted under this study has, with respect to Artificial Intelligence (AI), described a Canadian environment which is generally typical of all frontier technology areas:

- a fairly widespread but thinly populated research activity in universities with some pockets of significant world class capability;
- a rapidly developing interest within Government, in terms of both research laboratory activity (eg. DOC, NRC & EMR) and in terms of Research and Development (R&D) requirements (eg. Secretary of State and DOC);
- the evolution within, and outside the Federal Government, of coordinative mechanisms designed to stimulate more coherent approaches to R&D activities (eg. the Federal Government Inter-Departmental Coordinating Committee, the Canadian Institute for Advanced Research);
- a minimal interest in the Canadian private sector, both as Users and Suppliers, and as suppliers of venture capital; and,
- some indications of gathering strength on the import side, as foreign sources of technology begin to seek access to the Canadian domestic market.

In effect in Canada, there is nascent interest, but not as yet any critical mass of activity. Nor is there, as yet, in Canada, any planning framework or significant institutional response to the opportunities represented by AI.

The task underlying this Implementation Plan is to suggest levels and areas of Research and Development activity in Canada to meet the argued urgency underlying Artificial Intelligence technology. It also deals with the institutional responses necessary for the rapid development of

technological critical mass, having regard to both 'research' and to 'design and development'. In this latter connection, we have to deal with the question as to whether, in view of the potential importance of AI technology, a comprehensive national master plan along Japanese lines should be considered, and whether major new funding and planning institutions should be created, or whether we should opt for more piecemeal action through existing institutions, albeit within some general budgetary framework. We have opted for the second approach.

In our view, given the embryonic nature of Canadian R&D activities in AI, and the urgency implicit in the competitive environment, it is highly doubtful whether a detailed all embracing 'master plan' and/or a master institution would be feasible, credible or efficient, compared with more selective approaches launched through existing institutions. It is more important for a whole variety of reasons to announce targets, set priorities and develop and launch specific action programs. National efforts in AI must be put into motion.

A further aspect which deserves attention relates to funding modalities. Such funding as is generated, for AI R&D, particularly by Government, must be planned and implemented so as to deliberately build up cooperative activity, ie:

- between Universities on specific areas of research;
- between Government laboratories and industry;
- between User and Supply elements of the private sector, on projects of User interest; and

- between the Supply elements themselves, as a means of building up industrial capability.

Essentially, this point bears on the issue of 'centres of excellence', which is dealt with later in these comments.

Demand Formulation

In the initial stages of the proposed accelerated assault on AI, the creation of demand is critical, in order to attract growing interest and resources. Government has a lead role to play as a customer in triggering action in this field of advanced endeavour. The extent to which government's interest is advertised and its funding activities orchestrated will determine the private sector funding response. It is private sector funding which must eventually carry the burden of the commercialization process. Private venture capital will tend, as it does in the United States, to track national interest. It might be added in this connection that recent Federal Government initiatives in the tax field have started to generate very positive responses to high technology in the private sector, and larger pools of venture capital are beginning to form to support such activities. In this context, a sharper focus on AI technology on the part of government is likely to result in a 'capture' of these new levels and sources of funding.

2.0 Approach to Implementation

The broad approach which we propose in respect of implementation, is as follows:

- Research:

  - to suggest an advertised prioritization of AI research on the part of the Federal Government, as a means of attracting increasing national attention and resource allocation, to the field;

  - to suggest specific responses designed to improve the conduits for research funding; and

  - to suggest areas of technology for specific attention and to recommend increased levels of funding support.

- Development

  - as with Research, to advertixe priority at the national level; and

  - to initiate very specific mission oriented demands which will bring Canadian industrial capability to bear in as concerted a way as possible.

In this approach and as indicated in the Opportunities for Canada Report, we have recognized the lead role which the Federal Government needs to play, as a stimulator of advanced technology deemed to be in the national interest. It is however important once again to emphasize that success in terms of generating a broad and sustained national thrust in respect of AI technology will be achieved only if and when the technical, managerial and financial resources of the Canadian private sector are drawn into this field of endeavour.

From the 'funding' point of view alone, the level of effort which will be required over the next ten to fifteen years is such that it simply is not possible to conceive of the burden being carried solely by the Federal Government. From a competitive 'market oriented' perspective moreover, it is unreasonable to assume that the thrust can be sustained on other than acute responsiveness to market requirements.

For these and other reasons, a prime requirement of a Federal Government R&D strategy on AI must be to engage, as rapidly and massively as possible, the interest and resources of the private sector.

2.1 Priorities

In both the areas of Research and Development we have noted the need for visible prioritization of AI by the Federal Government as a field seen to be of crucial national interest and opportunity. The more that the priority can be advertised, the more rapidly will occur the needed amassing of resources in support of it.

Specifically it is suggested therefore that:

- at Cabinet envelope levels, the Ministry of State for Economic Development (MSED) help to draw attention to Artificial Intelligence as a field of crucial technological and economic importance and as a field in which Federal funding initiatives are likely to be developed;
- at other Federal science policy focal points such as the Ministry of State for Science and Technology (MOSST) and the Office of the Science Advisor to Cabinet, similar prioritization be formally and publicly announced;
- that the Science Council of Canada, which has already taken a lead role in the development of public awareness of AI technology, be encouraged to continue its efforts;
- that, at Federal research establishments, such as NRC and CRC, the publication of planned thrusts in AI technology be accelerated;

- that NSERC be encouraged to announce the addition of artificial intelligence (AI), natural language processing (NLP), and machine translation (MT) to the set of areas funded under its Strategic Grants Program;

- that, through its economic development agreements with the Provinces, the Federal Government seek opportunities to promote specific thrusts in AI technology; and

- that, as was the case with R&D as a totality in 1978, targets for national expenditures in AI be evolved and announced.

This prioritization activity is crucial. Although it may in fact be merely the external and visible expression of interests already developing in the Federal Government, the surfacing of these interests are important.

2.2 Coordinated Responses

In addition to prioritization, it is absolutely vital that coordinative efforts within the Federal Government be increased with respect to the need for speed of implementation. The following are some examples:

- Federal Mission Department Requirements

  In recent years the Department of Supply and Services has set up a small and highly effective organization known as the Canadian Business Opportunities Group (CBOG) whose responsibility has been to coordinate Federal, and Federal/Provincial requirements for the supply of new technology equipment and services, and to use such coordinated requirements to support industrial source establishment in Canada (eg. the recent water-bomber production program emerged from such an initiative).

It is recommended that CBOG be formally charged now with the task of rapid identification of mission-department requirements in Artificial Intelligence technology with a view to early establishment of Canadian supply sources.

- Inter-Departmental Committee on AI

A Federal Inter-departmental Committee to coordinate interest and policy on AI should be established. It is important that this mechanism be used to assist in encouraging launching and supporting specific initiatives, in both research and development. Consideration should be given to tasking it with accelerating the placement of funding programs with the research community and the private sector, complete with targets and time-scales. It is in our view essential that this Committee stimulates action. A recently organized inter-departmental committee examining policy issues related to artificial intelligence, very large scale integration, and computer architecture could serve as the basis for our proposed committee.

- NRC, CCRS and CRC

Some Government research laboratories are already becoming involved in AI, and others are considering doing so:

- NRC is establishing a CAD/CAM centre in Winnipeg, and its mandate includes AI activities;
- CCRS is doing work in image analysis and will be increasing its involvement in AI;
- the CRC may well evolve an 'informatics' role relating to 'office of the future' technology with an AI component.

The pressing need in respect of such activities is that they be planned to maximum collaboration, perhaps through joint ventures with University and Industry, on specific AI projects.

- Federal/Provincial Sub-Agreements

  In areas of broad mutual economic/industrial/technological interest, the Federal Government negotiates agreements with Provinces for cooperation. They are supported by funding from both regimes. Under these umbrella agreements sub-agreements can be, and are, concluded on specific topics.

  It is recommended that these sub-agreements be considered as a means of merging R&D activities in AI, where there are applications of mutual interest. Machine translation, medical expert systems and computer aided instruction for educational and job training applications, may be other areas of interest in this connection.

  We refer to medical expert systems as an example because of the existence of strong scientific expertise in Canada, and because of the presence of large pools of funding dedicated to medical research. This field may therefore be institutionally suitable for a rapid launch.

- Mission Department Coordination - Purchasing

  In addition to the CBOG role in coordination of Federal requirements noted above, there are cases, such as Machine Translation requirements, where two or more Departments can discern areas of very specific mutual development interest. In such cases, a lead agency (in this case Secretary of State) should act rapidly to set up common R&D objectives (with DND, for instance, which has strong translation requirements) and to encourage rapid R&D procurement based on common specifications. It is suggested that such a program could be supported by the Source Development Fund.

In these and other cases, purchasing activity should dictate as a pre-requisite, responses which demonstrate high levels of cooperation between the private sector and university capabilities, so as to help coalesce entrepreneurial and scientific resources.

- Computer Networking

NSERC and DOC should sponsor and support a Canadian computer network modelled along the lines of the US ARPA network.

The twin emphasis in the above recommendations are:

- the rapid coordination of demand;
- the involvement of the university and industrial sector in Federal activities.

Both of them address the major objective of the earliest possible, maximum involvement of the private sector in the exploitation of AI technology.

- Centres of Excellence

Given the scarcity of research and development resources, in manpower, and the rising intensity of competition for those resources (evidenced in the attraction of scientific manpower towards the United States), it is inevitable that 'centres of excellence' will need to be formed in Canada. These can be conceptualized as pools of talent brought to bear on specific technological objectives for sustained periods. In our view this objective can be achieved without large new 'institutions' but rather through well orchestrated programs blessed with clear delineation of responsibilities.

> The model which could be used is the "Action Concerte" in which Federal, University and Industrial capabilities are brought to bear on a specific goal, under a single funding and management umbrella. In Machine Translation, for instance, it should be possible to devise a three-year R&D program, with a government lead agency, with say the University of Montreal undertaking a broad, background research program plus providing scientific support to system development, and a selected industrial partner or partnership carrying out system and hardware development. Such a centre of excellence would enjoy both a specific technological goal, and an amalgam of talents which should help to position Canada in world markets. It is clearly possible to conceive of several such 'centres of excellence' forming the core of the national effort in AI.

2.3 Proposed Goals and Targets

With respect to Artificial Intelligence, the field of endeavour has already been publicly sensitized to some degree by virtue of the intense discussion of national programs in other countries (eg. the Japanese fifth generation computer project). It is suggested therefore that it would be appropriate for MSED and MOSST not only to publicise the priority to be accorded to AI, but also to establish goals and targets for increased national (as opposed to merely Federal) spending in AI.

In the early years of the program, the bulk of such spending would need to come from government. However, as demand begins to emerge, it is reasonable to project a larger proportion of national expenditures to be forthcoming from the private sector.

2.3.1 National Target

At this seminal stage, the establishment of specific targets is of necessity arbitrarily based, with provision for refinement in succeeding years as experience develops. As estimated in the Opportunities for Canada Report, the current level of spending on AI research in Canada is on the order of $5,000,000 per annum. This is deemed to be so

minimal an expenditure as to guarantee mediocrity if sustained. This is particularly so in light of the sharply accelerating efforts in other countries. Further, one major consequence of sustaining effort at that level is that such scientific capability as we now possess is very likely to be drawn away to the United States and other countries as major programs are announced, funded and staffed in those countries. Standing still is not therefore a feasible option.

One interesting aspect of this problem of the capturing of scientific resources is the speed at which foreign interests are funding Canadian research capabilities in AI. The University of Waterloo computer capabilities for instance, are to an increasing extent being 'contracted for' by Japanese corporations.

With regard to national 'targets' we have deliberately addressed the issue in a 'macro' way. Although some suggestions are made about specific project funding, as in MT, we think it important to reflect the increasing urgency underlying AI in the form of overall annual expenditure objectives. These expenditures would be a compound of all national research and development activities, including grants, contracts, private sector contributions, and the like.

2.3.2 Canadian R&D Framework Strategy

It is suggested therefore that a broad ten-year Canadian R&D strategy in AI should provide for:

- a first five years of 'catch up' effort in which we endeavour to produce a critical mass of research and development capabilities;

- a second five years at which, with critical mass achieved, the overall percentage annual increases decline somewhat, relative to a larger base.

2.3.3 Proposed Overall Spending Targets

First Five Years(Phase 1)

In the first three years we suggest a target of doubling expenditures each year. In years four and five we suggest increases of 75 percent and 50 percent, respectively. Assuming the accuracy of the base figure of $5 million, this would provide for the targets shown in Exhibit One.

Second Five Years (Phase 2)

Having established this critical mass and level of expenditures, it is suggested that the following five years, which will be effectively market driven in that the bulk of the expenditures would be private sector sourced, should be targeted for a twenty percent annual increase, resulting in a total 5 year R&D effort of $935 million (see Exhibit Two).

2.4 Public/Private Sector Financing

As noted, the public sector must, of necessity, finance the initial push to critical mass, with the proviso that private sector response and funding will develop rapidly until it assumes the major share of the burden.

It is suggested for purposes of initial review that:

Exhibit One

Phase One: Target R&D Expenditures on AI

Year	R&D Expenditures related to AI	Absolute growth over previous year	Percentage increase over previous year
		- ($ million) -	
1984	$ 10	$ 5	100%
1985	20	10	100%
1986	40	20	100%
1987	70	30	75%
1988	105	35	50%
5 year Total	245		

Exhibit Two

Phase Two: Target R&D Expenditures on AI

Year	R&D Expenditures related to AI	Absolute growth over previous year	Percentage increase over previous year
		- ($ million) -	
1989	$126	$21	20%
1990	151	25	20%
1991	181	30	20%
1992	217	36	20%
1993	260	43	20%
5 year Total	935		

- in the first five years, the Federal portion of funding should decline from virtually 100% to 70% of national expenditures;
- in the second five years, Federal expenditures should further decline to a steady 33% reflecting more normal patterns for technology support. This pattern is illustrated in Exhibit Three.

The numbers appear large but are not in comparison with efforts in other major industrialized countries (see Exhibit Four). AI technology represents the central thrust towards an information oriented society and a Canadian knowledge-based industrial economy. The Federal expenditures over the first five years, for instance, represents less than the amount recently committed to the establishment of a helicopter source establishment program. By contrast, the suggested 5 year commitment to AI R&D pales in comparison to only one year of federal funding for energy R&D which amounted to $338 million in 1982-83 or the Federal Government's total R&D expenditures of $2.94 billion in 1982-83.

2.5 Allocation of Funds

In the matter of the allocation of funds, government responsibility should be mainly in areas of special national interest which includes fields such as machine translation, communications and remote sensing. For each of these areas a guideline for the allocation of funds should be up to 15 percent of the suggested government expenditure. The allocation of the remainder of the funds should be largely determined by the market and the demands and interests of the research community.

Exhibit Three

Sources of AI R&D Funding
Industry vs. Government

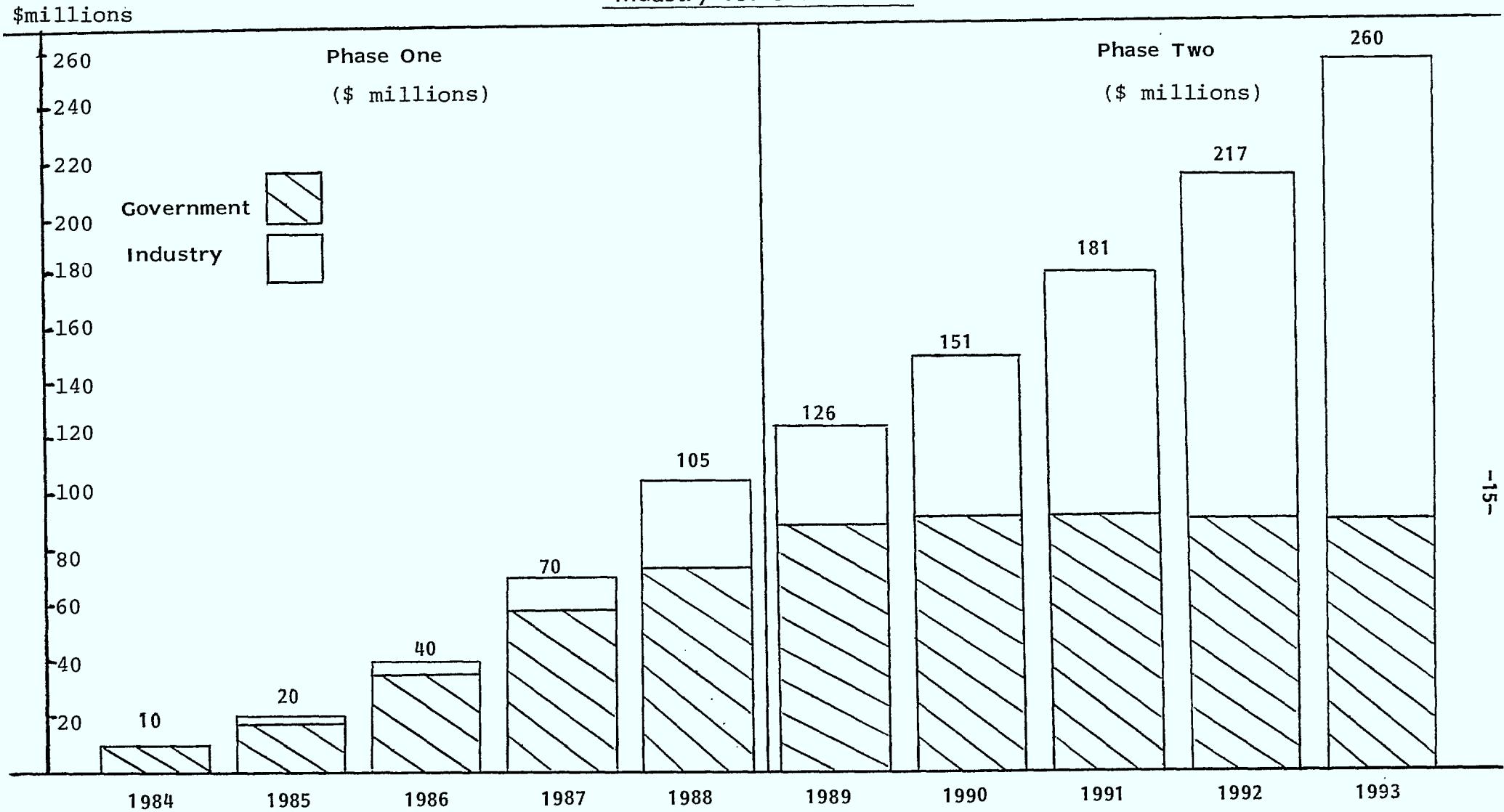


Exhibit Four

Estimated Total AI Related R&D Effort in Major Industrialized Countries

Country/Organization	Estimated Total AI Related R&D Effort	Anticipated Time Frame
	- (U.S. $ million) -	
Japan	$750 - $1,000*	0 - 10 years
United Kingdom	$630 **	0 - 10 years
United States	$2,000 - $2,500***	5 - 10 years
European Economic Community	$1,500 - $1,600	0 - 5 years

*rough order estimates based on available data

**publicly announced British government commitment to AI R&D

***estimated government and private sector funding of AI R&D

2.6 Role of Research & Development

For both general and specific reasons, it is essential that a national thrust in AI should be sustained and underpinned by a strong effort in basic research, both in government laboratories and in the university sector. The general reasons require little elaboration. Manpower is, and is going to be, a crucial resource. National success will depend on an increasing flow of trained scientific personnel.

In terms of AI there are additional reasons for mapping out a strong and sustained (ten year) research effort.

- As has been indicated in other reports submitted under this contract, the AI field is heavily interactive and interdisciplinary. For example, effective long-range system solutions in MT, for instance, can only be achieved based on intensive and wide ranging research efforts in computer technology. The nature of extent and direction of breakthroughs have still to be determined.
- To an unusual degree, the effort in AI is international. In order to acquire information and know-how, the Canadian research community will need to contribute its share to the global pool of knowledge, and will need to be heavily interactive with the international community.

With regard to development, the early years of the AI thrust program will, as noted, be characterized largely by government leadership, not only as a provider of research funding, but also as a stimulator of demand. This demand should emerge as funded development programs anchored to mission department requirements, such as MT systems, expert systems and the like. Contracts associated with such mission should specify requirements in the broadest possible terms, so that industry can be brought into the technology at the earliest phase of system definition.

Taking MT as an example one could therefore envision, in 1984 the placement of a three year MT system contract, jointly sponsored by the Secretary of State and DND, calling for successive phases of system specification, system development, prototyping, and demonstration. The magnitude of such an effort would likely be $10,000,000-$15,000,000. A range of such projects would be at the core of the AI development effort, and would represent a significant portion of the national target for AI R&D expenditures.

Arising out of our Phase 1 financial projections, a suggested breakdown of the proposed national targets for AI is shown in Exhibit Five and is expressed as expenditure as opposed to commitment.

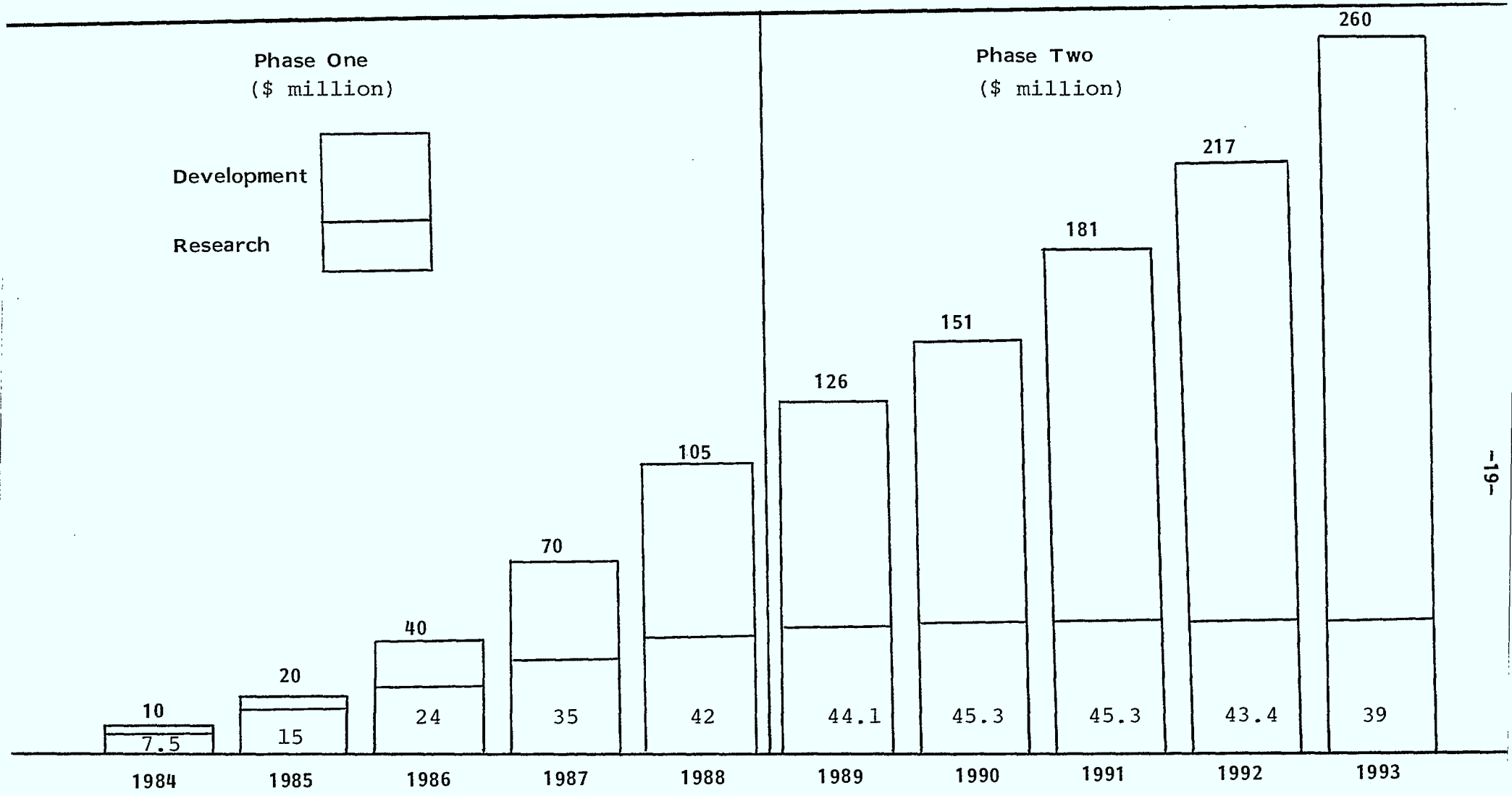
It is emphasized that from the hardware and commercialization point of view the best form of stimulation is the placement of orders. Not only does this method provide for the meeting of Departmental mission requirements, it also injects discipline as to technical goals and brings industry into the market place, with know-how, at the earliest stage. Speed 'off the ground' with regard to such developmental type contracts, is therefore absolutely crucial to the emergence of the private sector, not only as an industrial capability but as an eventual contribution of financial resources.

3.0 Marketing and Foreign Technology Considerations

The Department of External Affairs and the Department of Supply and Services will have the key roles to play in the tracking and analysis of export markets and in the establishment of Canadian sources responsive

Exhibit Five

Expenditures on AI Research
vs. Development



to such markets. Further, agencies such as NRC and CRC will be interested in the monitoring of foreign technological activity in their respective areas of thrust.

Because of the developing national responses to the Japanese fifth generation initiative, the elements of foreign technology screening and scientific and developmental international joint-ventures will be of critical importance to AI.

In our view, this international activity should be specifically recognized within a Canadian R&D strategy framework, in respect of:

- supporting university based scientific cooperation abroad;
- encouraging, at Government and Mission Department level, the development of Memoranda of Understanding with other countries on specific areas and applications of AI technology in such fields as MT, NLP, and expert systems;
- funding exchange visits of scientists, scholarships, and scholarly meetings;
- screening of foreign AI technology (for instance through our Scientific Attaches); and
- deliberately promoting corporate joint ventures between foreign sources of technology and Canadian enterprises.

Again a broad funding goal should be established, primarily in order to draw attention to the need, and to help focus and track the activities.

4.0 Research and Development Programs

4.1 Research

Within the suggested increase in the broad support for AI research, there are certain areas of endeavour to which particular attention needs to be paid. These arise both from perceived areas of current scientific strength, from areas which present possible opportunities for Canada, and from areas where some domestic demand might be susceptible to coalescence. These are:

- Machine Translation;
- Natural Language Processing with particular reference to management information and office automation systems;
- Lensing Technology (vision systems) for Robotics applications;
- Expert systems for medical and educational applications; and,
- Image analysis with respect to satellite-related resource programs.

Associated with each of these fields will be the need to undertake both basic and applied research perspectives.

4.2 Development

In terms of product development, the Opportunities for Canada Report indicates specific areas where development contracts could be launched with strong probability of success. These include:

- semi-automated machine translation systems;
- improved computer-aided learning systems;
- enhanced image analysis systems for resource surveillance;

- expert systems for medical diagnosis, natural resource management and telecommunications; and,
- intelligent command and control systems for military applications.

Each of these proposed research and development activities can be further developed in terms of specific R&D business plans, and in terms of more specific technical objectives. Already they can be readily focussed by sponsoring agencies, and can represent the core of the R&D expenditure envelopes suggested in this report.

4.3 A Case for a Research and Development Program in Machine Translation and Natural Language Processing

In the Opportunities for Canada Report, it was stressed that Canada should develop a strong R&D capability in machine translation and related areas. In addition to the fact that we have needs and expertise in these areas, there is evidence that machine translation and multilingual NLP systems can, in the mid and long term, become an habitable niche on the international scene: with the exception of Japan, foreign research on natural language processing tends to focus on unilingual (mostly English) applications.

We will attempt here to outline a research and development program in the areas of MT and NLP. Our proposals are meant to capitalize as much as possible on what emerged from the previous reports as the most promising technology thrusts and the best opportunities for Canada. While issues concerning the content of an R&D program can to some extent be kept separate from issues concerning possible institutional

frameworks for program implementation, it will be appreciated that the specific program proposed here has some straightforward organizational implications.

4.3.1 General Framework

As noted in the Opportunities for Canada Report, the Canadian R&D community in AI/NLP/MT is thinly spread across the country. If Canada is to build a strong R&D capability in the future, there is an essential need for stable institutions that will make it possible to rationalize efforts in a long term perspective and to focus energies on a limited number of reasonable goals. More specifically, such institutions would provide the means to:

- gather and make widely available technical literature on AI/NLP/MT;
- gather, develop, maintain and make available state-of-the-art software tools for R&D in AI/NLP/MT;
- make available the best hardware for NLP/MT R&D;
- gather, develop, maintain and make available products of general interest for the R&D community, as for example multilingual lexical databases, large-scale grammars for French and English, etc.;
- give adequate multidisciplinary training (computer science, linguistics, 'translatology', etc.) for R&D in NLP/MT, in view of forthcoming projects;
- facilitate on-going cooperation and exchanges with foreign R&D centres; and,
- bring together the required critical mass of specialists for specific R&D projects.

Within such an environment, a number of research and development goals can be set for the short, mid and long term, and realized as sub-projects.

4.3.2 Short Term Goals (2 - 3 years)

1. Bring together Canadian resources in office automation, telecommunications, translation and MT/NLP, and undertake the development of a marketable translator's workstation. Such a tool would integrate the following features in a user-friendly way:

- micro-computer environment with provision for multi-user configurations;
- telecommunications facilities: link between text producer, translator and translation user; access to remote terminology banks and other databases;
- state-of-the-art multi-window word processing with multi-alphabet and multi-font capabilities;
- on-line dictionary lookup and dictionary maintenance facilities;
- facilities for producing concordances (key-word-in-context) to speed up terminological research;
- some NLP capabilities, such as morphological analysis working in tandem with dictionary and concordances facilities; and,
- provision for extensibility in view of forthcoming progress in relevant fields (eg, text-critiquing systems).

It should be stressed that Canada has available all of the required expertise. A link should be established between this project and other programs, such as Department of Communications programs in office automation. Although a detailed feasibility study is called for, it seems that a good workstation could be developed in less than ten person/years.

2. Research on sublanguages and text complexity must be pushed further; more sublanguages should be examined from the standpoint of their complexity for automatic processing. This research should be tied to market research in order to permit the selection of a number of promising domains for NLP/MT applications using existing technology.

3. Explore the possibility of contractual arrangements with foreign partners to develop in Canada, and market internationally, French versions of products which currently exist only in English, as for example:

- data base query systems (eg., AI Corporation's INTELLECT)
- text-critiquing systems (eg., IBM's EPISTLE).

4. Explore the possibility of developing NLP-based computer-assisted language instruction systems; some provincial governments might be interested in some form of participation in such projects; arrangements could possibily be made with the designers of existing prototype system (eg, James Soper from Digital Research). Furthermore, this could be tied to a Canadian role in the Group of Eight program on new technologies applied to education, vocational training and culture.

4.3.3 Mid-Term Goals (3 - 5 years)

1. Develop one or more MT systems. Instead of attempting to build a large-scale second-generation system, as is currently being done in some foreign countries (eg, EEC), the aim would be to develop sublanguage-oriented small-scale systems. The technology used should be basically

similar to the one used in the TAUM systems. The areas selected should be maximally simple sublanguages, as determined by the results of sublanguage studies planned as a short-term goal. Some possible candidates can already be mentioned: extension of TAUM-METEO to handle French to English forecasts, and translation of the Federal government's job notices. The development of these systems will draw on general purpose MT/NLP software and packages such as large-scale grammars made available in the framework of the entire R&D program. Under these conditions, development of an MT system for a simple sublanguage should require between 5 and 10 person-years.

2. Develop one or more systems for automatic bilingual synthesis of natural language reports from non-linguistic data. Provided general purpose NLP software is available, practical systems for simple domains (such as stock market reports and agricultural market reports) could be developed within approximately 5 person-years.

3. Subject to agreements with foreign companies established as a short-term goal, develop French versions of data base query systems and/or text-critiquing systems. Such projects would be relatively low-risk inasmuch as the aim is simply to adapt for French products that have already proven effective in English. Moreover, in so doing, we would gain expertise in areas in which there has been little domestic activity so far. While the investment required will vary with the systems, it seems reasonable to assume that in cases like INTELLECT, 2 or 3 person-years would be sufficient to produce a practical French version.

4.3.4 Long Term Research Goals (5 to 10 years)

Our short- and mid-term goals make it clear that there is a wide body of technology already available for practical applications. It is reasonable to think that beyond the years covered in our mid-term goals, improved versions of these technologies will continue to provide the base for further development projects, tackling progressively more difficult problems.

However, our R&D program would be shortsighted if it failed to include longer-term research activities, even in its early phases. It has been pointed out in the State-of-the-Art Report that there is a worldwide (except for Japan) lack of longer-term efforts in MT. As a result, a wide body of new techniques developed for NLP and AI, though in principle applicable to MT problems, have rarely been used in MT systems. Moreover, there has been little progress made on problems that are specific to translation applications.

4.3.4.1 Some Major Problem Areas for MT/NLP

There exists a broad and well identified set of problems that stand in the way of all systems that have to manipulate the content of natural language texts, and whose solution can only come through advances in semantics, pragmatics, inferencing and knowledge representation. Some of these problems are:

- semantic role of prepositions;
- semantic structure of nominal compounds;
- tense and time reference;

- scoping of conjunctions;
- scoping of quantifiers; and,
- textual linking devices such as:
  - anaphoras
  - ellipses, and
  - focus phenomena.

In addition to these problems common to most NLP applications, MT systems face specific problems which, although they appear of extreme importance for cognitive science, have received very little attention lately. Foremost among these is the problem of the nature of translation rules:

- Is translation actually carried out through some universal interlingua, or is there rather a complex system of rules specific to pairs of languages? What is the nature of the huge 'bags of tricks' that experienced translators are said to have acquired (cf, expert systems)?
- If there are 'transfer rules', what is their formal nature? What constraints do they obey? How do they interact with other types of knowledge?

4.3.4.2 Research Program

In order to begin seriously tackling some of these problems within an MT system, we propose the following program, which in our opinion could give Canada a leading role in MT research by pioneering the development of practical 'third generation' MT systems.

1. Assess the possible contributions to MT theory of recent results in semantics, pragmatics and AI, as for example:
   - semantic theories such as Montague grammars and Situational semantics;

- recent developments in pragmatics, and developments to be expected in the near future with such projects as the Centre for the Study of Language and Information (Stanford University);

- knowledge representation techniques used in state-of-the-art AI systems (such as frames, non-monotonic logic, etc.); and

- recent developments in computer science, such as logic programming.

2. Examine in detail the knowledge used by expert translators, and try to devise more powerful schemes for translation rules.

3. Integrate the results of 1) and 2) into a coherent framework for practical third-generation MT systems.

4. Produce a prototype system for a restricted domain (such as, for example, the 'synopsis' part of weather forecasts which has been judged to difficult for the type of technology used in the TAUM-METEO system.

5. Apply a similar technology to other NLP applications, such as sophisticated data base query systems.

Of course, a long-term research program of the type proposed here can be adjusted to various funding scales. However, we believe that to attain a critical mass required for significant breakthroughs, a team of five to ten full-time specialists would be required. In some conditions, it would not be unrealistic to believe that in five years a full-scale prototype system could be produced. Furthermore, if the right domain is chosen, it may very well turn out that the prototype will not only be a good research vehicle, but will rapidly become of practical value. In ten years, this new technology could have a significant impact in the market.

5.0 Conclusions

In this report we have presented a broad framework for a major national thrust in Artificial Intelligence. The framework springs from a review of the state-of-the-art, coupled with a judgement as to the specific opportunities that present themselves to Canada in a global environment of rapidly increasing competition.

It is our view that a national thrust is vitally necessary because AI technology lies at the heart of competitive competences in the age of informatics and automation. It is also in our view that a national thrust involving sharply accelerated R&D effort in all sectors of AI can be achieved without the creation of major new institutions, with all that this implies in terms of lengthy deliberation as to responsibilities, and revised budgetary dispositions. Rather we have opted (as is the case in the United States) for rapid program launches through existing mechanisms, coupled with advertised prioritization by goal and target settings. Only minimal institutional changes are advocated and these are designed largely to speed the flow of government demand to the private sector, and to ensure the coalesence of scarce resource.

In our view the urgency of the competitive situation in AI is such that we must start now with specific funding action. Scarce scientific manpower will not stay in Canada unless a domestic magnet is created. Action can proceed in parallel with planning. It is our sense that a race of crucial economic and technological importance has already started. Canada must run.

NOTE

This report, A Program Plan, is the fifth volume in a series of reports addressing Natural Language Automated Processing and Artificial Intelligence. The reports in this series are:

- The State-of-the-Art
- Implications of New Technology Thrusts
- Possible Social and Economic Impacts
- Opportunities for Canada
- A Program Plan, and
- Priorities for Canada

ACKNOWLEDGEMENTS

The material in this report has been assembled on behalf of the contractor by Nordicity Group Ltd. It is based on input from various members of the study team and includes information supplied by numerous outside sources contacted through interviews.

Executive Summary

In this report a broad framework for a major national thrust in Artificial Intelligence is presented. The framework springs from a review of the state-of-the-art, coupled with a judgement as to the specific opportunities that present themselves to Canada in a global environment of rapidly increasing competition.

A national thrust is vitally necessary because AI technology lies at the heart of competitive competences in the age of informatics and automation. Also, a national thrust involving sharply accelerated R&D effort in all sectors of AI can be achieved without the creation of major new institutions, with all that this implies in terms of lengthy deliberation as to responsibilities, and revised budgetary dispositions. Rather we have opted (as is the case in the United States) for rapid program launches through existing mechanisms, coupled with advertised prioritization by goal and target settings. Only minimal institutional changes are advocated and these are designed largely to speed the flow of government demand to the private sector, and to ensure the coalesence of scarce resources.

Key report recommendations are:

- the federal government should advertise the priority of AI research and development and should assume a lead role;
- coordinated responses should be developed by groups such as the Canadian Business Opportunities Group, an Inter-departmental Committee on AI (to be established), government research laboratories (NRC, CCRS & CRC) and others;

- NSERC and DOC should sponsor a Canadian computer network modelled along the lines of the US ARPA network;

- several Centres of Excellence should form the core of the national effort in AI;

- MSED and MOSST should establish goals and targets for increased national spending in AI;

- proposed overall spending targets are:
  First Five Years $245 million
  Year Six to Ten $935 million

- private sector financing should be 30% of annual expenditures by year 5 and 67% by year 10;

- interaction with international groups should be encouraged and supported at all levels;

- key areas for research programs include:

  - machine translation;
  - natural language processing with particular reference to management information and office automation systems;
  - lensing technology (vision systems) for robotics applications;
  - expert systems for medical and educational applications; and,
  - image analysis with respect to satellite-related resource programs.

- suggested areas for development programs include:

  - semi-automated machine translation systems;
  - improved computer-aided learning systems;
  - enhanced image analysis systems for resource surveillance;
  - expert systems for medical diagnosis, natural resource management and telecommunications; and,
  - intelligent command and control systems for military applications.

The urgency of the competitive situation in AI is such that we must start now with specific funding action. Scarce scientific manpower will not stay in Canada unless a domestic magnet is created. Action can proceed in parallel with planning. It is our sense that a race of crucial economic and technological importance has already started.

NATURAL LANGUAGE AUTOMATED PROCESSING AND ARTIFICIAL INTELLIGENCE.