# Integrating connectionist and symbolic computation for the theory of language

**Paul Smolensky, Géraldine Legendre* and Yoshiro Miyata****

Department of Computer Science and *Department of Linguistics, Institute of Cognitive Science, University of Colorado, Boulder, Colorado 80309, USA

**School of Computer and Cognitive Sciences, Chukyo University, 101 Tokodate, Kaizu-cho, Toyota, 470-03, Japan

**In this article we present some of the fundamental principles of a research program—the Sub-Symbolic Paradigm (SSP)—based on a particular approach to unifying connectionist and symbolic computation. SSP has been developed primarily for the study of higher cognitive domains, and in this article we focus on SSP research on language and grammar. The SSP principles integrating connectionist and symbolic computation are developed by establishing mathematical relationships between two levels of description of a single computational system: at the lower level, the system is formally described in terms of highly distributed patterns of activity over connectionist units, and the dynamics of these units; at the higher level, the same system is formally described in terms of symbol structures, the constraints governing them, and the processes manipulating them. Applied to natural language, these computational principles entail that a central organizing principle of grammar is *optimality*: a grammar is a means of determining which of any set of structural analyses of an input is the most well-formed. Such a *Harmonic Grammar* consists of a set of conflicting 'soft' rules or constraints, each of which is in principle violable in the appropriate context. This constitutes a novel framework for formal grammar which emerges from the connectionist computational substrate. We describe how such soft rules allow for precise treatment of a complex set of interactions of semantic and syntactic constraints in a single language, and of universal patterns of interaction among phonological constraints.**

Few would deny that over the past decade or so, neural or connectionist networks have produced an explosion of results and a great deal of interest. Yet this approach to the computational modeling of intelligent cognitive systems faces fundamental problems. The research proposed here has a major connectionist component, but it distinguishes itself from the bulk of connectionist research in the following respects:

(0) a. It is strongly guided by symbolic computation, but not a 'hybrid' in the usual sense: the connectionist and symbolic computation involved are not two components of a composite system, but *two descriptions of a single system.* We call this (integrated connectionist/symbolic computation'.

b. The emphasis is on *higher level* cognitive processes, with a main focus on *language*, which provides a particularly challenging testbed, since symbolic computation is so central to existing theory.

c. The main emphasis in the language research is on *formal grammars* for natural language, with supporting research on the grammars of formal languages.

d. The methodological emphasis is on the development and theoretical analysis of general mathematical *principles*, rather than simulation experiments on limited, specific 'models'.

We believe that (i) these properties allow our work to overcome several shortcomings of much connectionist research, and that (ii) the results already achieved in this research program show its soundness and promise. To argue (i), we begin this overview (Section 1) with a brief summary of the goals which have shaped this research, leading it to have the properties (0). In Section 2 we give a brief, high-level summary of some previous results of the research program, addressing (ii). In Sections 3–4 we present the fundamental principles of the approach, and at the end of Section 4 we compare our approach for relating connectionism and language to other approaches.

## 1 Research goals

Our approach to integrating connectionist and symbolic computation has evolved as a solution to what we take to be a profound paradox lying at the heart of cognitive science. Formal theories of logical reasoning, grammar, and other higher mental faculties compel us to think of the *mind* as a machine for rule-based manipulation of structured arrays of *symbols*. What we know of the *brain* compels us to think of human information processing in terms of manipulation of a large set of *numbers*, the activity levels of interconnected neurons. Finally, the richness of human *behavior*, both in everyday environments and in the controlled environments of the psychological laboratory, seems to defy rule-based description, displaying strong sensitivity to subtle statistical factors in experience, as well as to structural properties of information. To solve the *Central Paradox of Cognition* is to resolve these contradictions with a *unified theory of the organization* of the mind, of the brain, of behavior, and of the environment. To this end, we set the goal of a theory that is:

(1) *Theoretically integrated*

A theory:

a. that addresses the full challenges of higher-level cognition,
b. while achieving both 'horizontal' and 'vertical' integration.

'Higher-level cognition,' refers to those rather abstract domains—e.g. language, problem solving, reasoning, abstract planning—where nearly all existing cognitive theory relies heavily on some sort of symbolic computational model. A 'horizontally integrated' theory is one that provides a coherent account of the interrelation of cognitive processes across the wide range of higher- and lower-level cognitive domains. 'Vertical integration' requires a coherent theory of the interrelation of the multiple levels of organization spanning from the neural level up to the highest mental levels.

In Smolensky's paper[1] 'On the proper treatment of connectionism', an approach to the Central Paradox and the goal of Theoretical Integration called the *Sub-Symbolic Paradigm* ('SSP') was developed. In SSP, connectionism achieves vertical integration by adopting a level of description intermediate between those of neurons and of symbols, and by exploiting this intermediate ('subsymbolic') level as a bridge for bringing into contact theories of brain and theories of mind. Furthermore, it was argued, connectionism provides a computational account of unified cognitive architecture upon which can reside quite varied processes or virtual machines that serve the varying needs of diverse cognitive domains; this is a powerful means of achieving horizontal integration. Thus the same fundamental connectionist computational mechanisms can be seen to underly perceptual processes, memory, and certain higher-level processes; *different principles of organization* emerge as higher-level descriptions of different kinds of connectionist networks operating in different kinds of information-processing environments.

The other goals shaping the proposed research are methodological—the development of a theory that is:

(2) *Methodologically integrated.*

A theory:

a. that integrates the theoretical insights into the various cognitive domains contributed by the numerous and varied relevant disciplines;
b. whose development effectively exploits the diverse methodologies of these disciplines; and
c. whose content constructively feeds into these disciplines and furthers their own particular goals.

(3) *Principle-centered*

A theory:

a. centered on key analytical concepts and on the general principles governing them,
b. the principles being embedded in an overarching formal theoretical framework grounded on mathematically sound and powerful foundations.

Together, these Three Goals have led to a research program with the properties (0)—a research program which, compared to the majority of connectionist research, is more tightly integrated with symbolic research on computation and language, and based more in theoretical and mathematical analysis than in experimentation.

## 2 Previous results

Some of the previous results of the SSP research program can be briefly summarized at a high level as follows:

(4) a. SSP develops a mathematical formalism showing precisely how a mental representation can be *simultaneously* a fully distributed pattern of numerical activities at one level of analysis *and* the functional near-equivalent of a symbolic structure when analysed at a higher level[2-4].
  b. SSP shows in mathematical detail, illustrated by computer simulations, how mental processing can be simultaneously a massively parallel process of spreading activation at one level of analysis and, at a higher level, a kind of parallel holistic manipulation of symbolic structures—even those containing recursive embedding[2,4,5].
  c. SSP and related connectionist research demonstrate that the overall effects of spreading activation can often be analysed at a higher level as a process of *optimization*, in which a representation is constructed that maximizes a connectionist measure of well-formedness we call *Harmony*[6-16].

d. SSP shows how to combine the three preceding results (a–c) to define a new formalism for grammar, a formalism which has been successfully used to address long-standing problems in phonology and syntax to which solutions within purely symbolic theory have been problematic. This formalism—*Harmonic Grammar* —constitutes a novel integration of connectionist and symbolic computation, and rests on both symbolic and connectionist technical advances[14,17–21].

e. SSP combines (a–c) to shed new light on a central problem in the foundations of cognitive science, emphasized by Fodor and Pylyshyn[22] in their highly influential critique of connectionist theory: the explanation of how higher cognition can achieve, with finite and fixed resources, competence that is highly systematic, coherent, compositional, and productive[23,24].

These results constitute direct progress in the achievement of the Three Goals. The cognitive problems addressed in (4d,e) are among the most central in higher cognition, falling squarely under Theoretical Integration (1a). Results (4a–c) provide the supporting pillars of a vertically integrated theory, while involving general notions of connectionist computation that cut across many cognitive domains and simultaneously support horizontal integration: a significant step towards Theoretical Integration. Considerable Methodological Integration has been achieved by targeting problems in higher-level cognition which rely heavily on symbolic computation, which have been of central interest in related fields such as linguistics, by adopting the formulation of these problems that have been developed by the practitioners of those fields, and by giving a major role to the theoretical constructs and established methodologies of these other disciplines.

The final goal of Principle-Centering is served by all the results[6]. In each case, the results take the form of powerful and general cognitive principles, centered on novel concepts for understanding cognition that arise from viewing a lower-level connectionist computational model and a higher-level symbolic computational model as two descriptions, at different levels of analysis, of one and the same computational system. Some of these principles are presented below as (5)–(8). The connectionist technical innovations of SSP have contributed substantially to the soundness and power of the mathematical framework supporting a connectionist theory of higher cognition. Original symbolic technical contributions have also provided innovations in grammar formalism. Key have been mathematical bridges between the continuous, numerical model of computation underlying connectionism and the discrete, symbolic computation of virtual machines that emerge as natural higher-level approximate descriptions of appropriately designed connectionist systems. These mathematical techniques provide the kind of technical leverage needed to make progress on the Central Paradox of Cognition.

## 3 Emergence of symbolic from connectionist computation

We now present a high-level summary of three principles which are fundamental to the SSP approach:

(5) a. When analysed at the lower level, mental representations are distributed patterns of connectionist activity; when analysed at a higher level, these same representations constitute symbolic structures.

b. When analysed at the lower level, mental processes are massively parallel numerical activation spreading; when analysed at a higher level, these same processes constitute a form of symbol manipulation in which entire structures, possibly involving recursive embedding, are manipulated in parallel.

c. When the lower-level description of the activation spreading processes satisfies certain mathematical properties, this process can be analysed on a higher level as the construction of that symbolic structure including the given input structure which *maximizes Harmony*: the Harmony can be computed either at the lower level as a particular mathematical function of the numbers comprising the activation pattern, or at the higher level as a function of the symbolic constituents comprising the structure.

The two closely related principles (5a, b) employ tensor calculus to design the global properties of complex activity patterns, enabling the construction of recursive connectionist representations and networks with symbolic higher-level properties. The technique is called *tensor product representation*[2,4].

The lower- and higher-level descriptions of a representation (5a) are related as follows. The higher-level description is that of a symbolic structure characterized by a set of structural roles, each of which may be occupied by a filler, which is a constituent symbolic structure. The corresponding lower-level description is a vector (of activity values of connectionist units) which is the superposition or sum of vectors each representing one of the constituents; these constituent vectors consist of the tensor or generalized outer product of a vector representing the filler times a vector representing its structural role. Tensor calculus allows these representations to be defined recursively, so that fillers which are themselves complex structures are represented by vectors which in turn are recursively defined as tensor product representations.

Structure-sensitive symbolic processing (5b) of these representations is achieved by means of operations from tensor calculus which check conditions on constituents and which use linear transformations to move constituents in given structural roles to new ones, or to modify the fillers in given roles. Such operations are naturally embodied in connectionist networks[2,3,5,25,26].

Principle (5c) combines the tensor analysis techniques underlying (5a, b) with another technique from mathematical physics: Lyapunov functions. In the simplest cases, the core of the Harmony (Lyapunov) function can be written at the lower, connectionist level simply as the quadratic form $H = \mathbf{a}^T \mathbf{W}\mathbf{a}$, where $\mathbf{a}$ is the network's activation vector and $\mathbf{W}$ its connection weight matrix. The linear character of tensor product representations and the bilinear nature of $H$ imply that the Harmony can be computed at the higher, symbolic level[14]: $H = \Sigma_{c_1,c_2} H_{c_1;c_2}$; each $H_{c_1;c_2}$ is the Harmony of having the two symbolic constituents $c_1$ and $c_2$ in the same structure (the $c_i$ are actually constituents *in particular structural roles*, and may be the same).

This research has developed the concrete mathematical techniques needed to perform computations using principles (5a–c), and has realized these computations in computer simulations. One simple simulation, designed purely to demonstrate the formal capabilities of the technique, takes as input a distributed pattern of activity representing the tree structure underlying an English sentence, determines by inspecting the structure whether the form is that of an active or passive sentence, and, accordingly, produces as output a distributed representation of a tree structure encoding a predicate-calculus form of the semantic interpretation of the input sentence[2]. The network performs all the required symbol manipulation in parallel, and handles entire embedded sub-trees (e.g. complex NPs) as readily as it does simple symbols.

### *Methodology*

This research is made possible by methodological innovations intimately involving all the Three Goals. The insights and techniques required come from symbolic modeling of higher cognition, connectionist computation, discrete mathematical analysis, continuous mathematical analysis, symbolic computing, and numerical computing. The work requires identifying (a) what is central in symbolic computation for theories of higher-level cognition, and (b) how that can be formalized within discrete mathematics, (c) what is central in the higher-level characterization of connectionist computation, and (d) how that can be formalized within continuous mathematics; and then (*e*) how these elements of continuous and discrete mathematics can be unified in terms of general principles, concrete means of formal calculation, and computer simulation. Only when these disparate activities (a-e) function as a unit can they produce the results described here.

## 4 Optimization in grammar

A particularly challenging domain which provides an excellent testbed in which to develop the principles (5) underlying integrated connectionist/symbolic computation is *language*: specifically, the study of the grammars of natural and formal languages. We will view the grammar of a language in two ways: *descriptively*, as a function which identifies via an abstract specification the correct linguistic structure to output for each given input, and *algorithmically*, as a device for actually constructing this output (note 1).

In descriptive grammar, as studied particularly within theoretical linguistics, an extremely powerful methodology has developed in which a central role is played by the study of the *well-formedness* of various structures. This notion applies not only to traditional linguistic problems such as those of phonology and syntax, but also to problems of central interest to computational linguistics, such as semantic interpretation: purely syntactic structures such as parse trees are not the only structures that can be separated by a grammar into those which are well- and ill-formed — the same is true of structures that combine both syntactic and semantic information. Thus, e.g. in a number of unification-based approaches to syntax and semantics, the 'correct' semantic interpretation of an input sentence is analysed as the semantic part of the well-formed structure which contains the input, together with associated syntactic and semantic information (e.g. ref. 27).

Thus a powerful concept around which to build a connectionist-grounded theory of grammar is that of linguistic well-formedness. And the principles (5) turn out to be exactly what we need to do just that. The further fundamental principles underlying current SSP work on grammar are[14] (note 2).

(6) a. The well-formedness of a linguistic structure is measured by the Harmony of that structure.
  b. Descriptively, the grammar assigns to an input that linguistic structure which is most well-formed, i.e., has maximal Harmony. The descriptive grammar can therefore be specified by the Harmony function itself, which measures the well-formedness of all possible linguistic representations that could be assigned to an input.
  c. Algorithmically, the grammar is a Harmony-

maximizing connectionist network, the Harmony function of which specifies the descriptive grammar.

## Numerical theory

The theory now proceeds along two somewhat different paths, one numerical, the other non-numerical. The former is based on the following principle, a direct consequence of (5c) (ref. 14):

(7) a. The explicit form of the Harmony function can be computed to be a sum of terms each of which measures the well-formedness arising from the coexistence, within a single structure, of a pair of constituents in their particular structural roles.

b. Thus the descriptive grammar can be identified as a set of *soft rules* each of the form:
If a linguistic structure $S$ simultaneously contains constituent $c_1$ in structural role $r_1$ and constituent $c_2$ in structural role $r_2$, then add to $H(S)$, the harmony value of $S$, the quantity $H_{c_1,r_1;c_2,r_2}$ (which may be positive or negative; and $c_1, r_1$ may $= c_2, r_2$).
A set of such soft rules (or 'constraints,' or 'preferences') defines a *Harmonic Grammar*.

c. The constituents referred to in the soft rules include both those that are given in the input and the 'hidden' constituents that are assigned to the input by the grammar. The problem for the algorithmic grammar is to construct that structure $S$, containing both input and 'hidden' constituents, with the highest overall Harmony $H(S)$.

The distinction between well- and ill-formed inputs, according to this theory, is a numerically graded one: the higher the value of $H(S)$ for the structure $S$ assigned by the grammar to an input, the more well-informed is that input. The soft rules in a Harmonic Grammar can potentially interact very strongly; the harmony-maximizing structure, and its degree of well-formedness, can be highly sensitive to combinations of factors in the input.

Following goal (2), Methodological Integration, we take as a starting point for a Harmonic Grammar analysis of a particular linguistic phenomenon the working hypothesis that the particular kinds of constituent structures posited by the best current linguistic theories of that phenomenon are indeed valid higher level descriptions of the relevant representations. We then study patterns of well-formedness judgments to identify candidate constituent interactions; this gives us a set of candidate soft rules (7b), in which the numerical constants $H_{c1,r1;c2,r2}$ are unknown. These numerical values can then be automatically determined from the well-formedness judgments elicited from native speakers by an appropriately generalized version of the connectionist learning algorithm, back-propagation[32].

## Applications to syntax/semantics

We have applied Harmonic Grammar (henceforth, HG) to the study of a particular phenomenon in the syntax and semantics of natural language: *unaccusativity* or *split intrasitivity*[17] (note 3). This phenomenon can be introduced to those unfamiliar with it as follows. Of central importance in the interaction of syntax and semantics is the *argument structure* of verbs, which, on one view, relates the synthetic roles of a verb's arguments to the semantic roles of the interpretation of those arguments as participants in the described event (note 4). Perhaps not surprisingly, given its central role in language, argument structure has turned out to be quite a challenging problem, and a number of linguists have focused attention on the simplest case: events with one participant described by intransitive verbs.

In practically every language in which intransitive verbs have been carefully examined, it appears that they split into two classes, depending on whether their single argument behaves like the subject or direct object of a transitive verb in that language. In some languages, the distinction is reflected in the morphological marking of the verb: in Lakhota (a Siouan language), the two arguments of a transitive verb are directly encoded into the verb via a morpheme $X$ corresponding to the subject and a morpheme $Y$ corresponding to the direct object. Intransitive verbs contain a single morpheme .corresponding to their single argument, which in some cases is $X$ (typically agentive verbs, but not always), and in others is $Y$ (typically non-agentive verbs, but not always[33,34]). In languages with impoverished morphology such as English and French, there is no immediately visible distinction among intransitive verbs; yet a distinction can nearly always be observed in certain syntactic phenomena. For example, French has a construction in which the main verb *croire* 'believe' occurs with a participial complement, corresponding to English *I believe John gone*. In the French construction, only the direct object of transitive verbs or the argument of *some* intransitive verbs can appear as the object of *croire*; the subject of transitive verbs and the arguments of the remaining intransitive verbs cannot. The same distinction can be observed in a half-dozen other syntactic constructions in French[35,36].

A variety of theoretical approaches have shed light on unaccusativity phenomena. Some have emphasized the parallel between the split among intransitives in certain syntactic contexts and the corresponding split in behavior of the subject and direct object of transitives, as just exemplified for the *croire* construction. They advocate a 'deep' syntactic distinction, claiming in essence that the argument structure of some intransitive verbs dubbed 'unaccusative' calls for a deep direct object, rather than a deep subject, as called for with 'unergative' intransitives[37-39]. In such treatments, a deep direct object is often claimed to be a necessary condition for an intransitive verb to appear in a given (e.g., *croire*) construction, but such a condition is usually not sufficient, as the acceptability of the resulting sentence is sometimes sensitive to semantic and aspectual properties of both the intransitive verb and its argument[35,40]. This is one reason that different constructions do not allow exactly the same set of 'unaccusative' verbs: the problem of *unaccusativity mismatches*. The deep syntactic approach must incorporate an explicit account of the interacting semantic and aspectual factors if it is to give a complete account of the phenomena, including the complex pattern of mismatches.

Another line of work attempts to establish that such semantic and aspectual factors are themselves sufficient to provide a complete account, and that a deep syntactic distinction is unnecessary (e.g. refs. 41, 42). This controversy cannot be separated from such important and extremely controversial issues as mono- vs. multi-stratal syntax. But the level of complexity in the interaction between semantic and syntactic factors that underlie split intransitivity is often relegated to a secondary place in the disucssions: broad tendencies are emphasized, while the extent of their validity, and the factors leading to their violation, are typically neglected.

As part of the research on unaccusativity in French including our HG work, we are carrying out detailed studies based on a data base of acceptability judgements which we have assembled: 8393 sentences involving 183 intransitive and 225 transitive verbs in 11 syntactic environments. Studies of the 3608 sentences in this data base involving intransitive verbs have corroborated some of the claimed universal semantic and aspectual tendencies, not corroborated others, and identified new regularities. Our conclusion at this point in our study is that, in French, (a) semantic and aspectual factors play a major role in unaccusativity phenomena; (b) their role is more complex than has been previously proposed; (c) they are not individually or conjunctively sufficient to provide a complete account; (d) a major role is also played by a deep syntactic distinction; and (e) the syntactic and semantic factors interact strongly.

Our HG account has co-evolved with and contributed to our study of syntactic and semantic accounts; in fact, it builds on and integrates these accounts. The HG account involves representations of deep grammatical functions (DGFs) subject and direct object, and so incorporates the syntactic approach. Unlike symbolic theories, however, a given intransitive verb does not *require* its argument to have one or the other DGF—instead, it has a *preference* for one over the other; any linguistic structure in which this preference is violated has its well-formedness (Harmony) reduced by a particular amount which characterizes the *strength* of that verb's preference. This preference is encoded in one of the lexical soft rules in the HG account. This rule interacts strongly with (a) other syntactic soft rules, (b) a set of semantic soft rules, and (c) soft rules concerning syntactic/semantic correspondences (note 5).

The HG framework integrating these different types of constraint on well-formedness allows these constraints to interact strongly enough to account for the French data. Based on an earlier data base which we assembled, it correctly accounts for the acceptability judgments of all but 3 of the 885 sentences involving 143 intransitive verbs in five syntactic constructions[18]. A more comprehensive account currently under development[19] now correctly accounts for the well-formedness judgements of all but 104 of 8393 sentences involving 183 intransitive and 225 transitive verbs embedded in 11 syntactic constructions. With intransitives, the focus of the study, all but 14 of 3608 sentences are correctly accounted for.

### *Non-numerical theory*

In applying HG to phonology, Smolensky and Alan Prince (of Department of Linguistics and Center for Cognitive Science, Rutgers University) made a striking discovery: in a wide variety of phonological problems, the numerical strengths of soft rules arrange themselves so that the rules form *strict dominance hierarchies*. In these hierarchies, the preferences or constraints can be ordered from weakest to strongest in such a way that each constraint is stronger than all the weaker constraints *combined*; thus a given constraint must be satisfied (if possible), regardless of whether that entails violation of any number of weaker constraints—unless satisfying the constraint requires violating still stronger constraints that can otherwise be satisfied. In such situations, all the information carried by the numerical strengths of the soft rules can be re-expressed non-numerically as the ordering of the rules in the dominance hierarchy. Thus, in this special case, principle (7) can be reformulated in non-numerical terms. After considerable further development, which includes the development of a novel symbolic framework called *Harmony-Theoretic Phonology*, this principle assumes the following form[20,21]:

(8) a. A descriptive grammar is an axiomatically defined algebraic preference relation $\succ$ among linguistic structures; $S_1 \succ S_2$ is interpreted as '$S_1$ is more Harmonious (well-formed) than $S_2$'.

b. Given an input $I$, such a grammar assigns an output that linguistic structure $S$ which contains the input and which is maximally Harmonious; i.e. $S \succ S'$ for all other structures $S'$ containing $I$.

c. The well-formedness relation $\succ$ among linguistic structures is defined compositionally from well-formedness relations among the substructures from which the linguistic structures are built.

d. Most of the basic well-formedness relations and means of combination needed for the grammars of individual languages are universal: they appear in the grammars of all languages. What primarily distinguishes the grammars of individual languages is *the particular ways the universal well-formedness rules are combined* (e.g. the particular ordering of constraints in dominance hierarchies).

### *Applications to phonology*

This non-numerical formulation has been successfully applied to a variety of problems in phonology. A few examples of the problems which have been treated to date include: (a) the universal typology of basic syllable structure; (b) a detailed analysis of the remarkable Berber syllabification system (note 6); (c) classic interactions of various phonological processes such as those exhibited in Lardil and Yawelmani; and (d) the universal typology of stress systems. In (a), e.g., the universal typology arises simply by considering all possible dominance orderings of the following universal well-formedness relations:

(9) a. A syllable is more well-formed if it has a filled onset position;
b. A syllable is less well-formed if it has a filled coda-position;
c. A syllable is less well-formed if it contains unrealized (deleted) segments;
d. A syllable is less well-formed if it contains epenthetic (inserted) segments.

The new formulation of phonology based on principle (8) makes a number of theoretical contributions; e.g.:

(10) a. A precise formal framework is provided for powerful kinds of constraint-based reasoning, some of which are new, and some of which have previously been available, but only informally.

b. Accounts of the interaction of various phonological processes based on the ordering of rules in a sequential derivation are replaced by a declarative characterization of phonological well-formedness in terms of the relative strengths in a given language of preferences, most universal, some language-particular. In place of language-specific derivational processes, we have the universal process of Harmony maximization (a process which in fact underlies the connectionist account of many other cognitive domains, besides language, including lower-level processes such as perception and memory retrieval—note the implications for horizontal Theoretical Integration, (1b)).

c. Formal means are provided for deriving universal typologies from universal preferences, and for situating language-particular systems within a theory of universal phonology. For example, what previously appeared to be a rather singular, bizarre syllabification system in Berber can now be formally analysed as a natural, albeit extreme, special case of the universal theory of syllabification.

Results (10b,c) constitute significant progress in the development of a formal, declarative, constraint-based theory of universal phonology, a theory which has been rather elusive despite much effort directed towards it (note 7).

### *Methodology*

The two approaches comprising HG are made possible by integrating several methodologies: elicitation and analysis of well-formedness judgements by native speakers, theoretical analysis of the structure of linguistic representations, development of novel symbolic formalizations of optimization, mathematical analysis of connectionist computation, and the design of specialized connectionist processing architectures, learning algorithms, and network analysis techniques.

It is worth contrasting the methodology developed here with the two most prevalent methodologies currently practiced for relating connectionism and language; this illustrates the import of the Three Goals of Section 1. Our strategy, embodied in principles (6) through (8), is: to abstract from particular connectionist models to general connectionist principles; to use these to derive a general grammar formalism; to test the descriptive adequacy of this formalism by using it to develop specific analyses of particular linguistic data (which is formally described both at a lower level as a connectionist net and at a higher level as a set of soft rules); and, finally, to test the explanatory adequacy of the formalism by using it to characterize universal properties of the grammars of human languages. By contrast, the most typical approach to applying connectionism to language[45–54] is to identify some linguistic phenomenon of interest, construct specific data sets that exhibit it, train and test some particular connectionist network on these data, and then try to draw more general linguistic conclusions that go beyond these particular data. Thus the typical strategy attempts to connect connectionism and language by encoding particular data into particular networks, while the SSP strategy is to connect general high-level principles of connectionist computation with general linguistic principles: a direct manifestation of goal (3), Principle-Centering. Our point is not that the typical approach does not yield interesting and important experimental results about what connectionist networks can learn, represent, and compute—our claim is rather that the alternative SSP approach can offer a complementary kind of contribution which relies heavily on the integration of connectionist and symbolic computation.

Other examples of work striving to explicitly integrate connectionist and linguistic principles include refs. 55–59. This work has brought to light some valuable relationships between connectionist computation and linguistics, although we note that the connectionist computational principles involved in this other work—e.g. principles of activation and inhibition of adjacent prominence values in a linear sequence; principles governing similarity and continuity in temporally adjacent output feature vectors—are more low-level and computationally weaker than the principle of Harmony maximization. The result is that, like the typical connectionist approach to language just discussed, contact between connectionism and linguistics is attempted at a rather lower level than is the case in the SSP work we have described. For example, while the Goldsmith–Larson approach[57] yields some suggestive computer simulation experiments on Berber syllabification, extensive analyses of this problem have been developed using both the Numerical and Non-numerical SSP approaches. The latter includes a connectionist implementation which bears considerable resemblance to the Goldsmith–Larson network—but because our network is mathematically derived to maximize a specially-designed Harmony function, precise theorems concerning the correctness of its competence can be proved; furthermore, it is integrated into the broad grammatical framework of HG, and a universal theory of syllabification (10c)[21].

Another alternative strategy[60–63] could aptly be termed *implementationalist*, in that it uses connectionist mechanisms to directly implement (often quite serial) symbolic rule application. Advocates of this methodology argue that it results in major revision of symbolic theory, but we believe that the implementational relationship it enforces between connectionist and symbolic computation calls on the weaknesses, rather than the strengths, of both: the kind of connectionist network used for this sort of implementation typically fails to exploit the power of connectionist computation resulting from learning algorithms, distributed representation, mutual constraint satisfaction, and optimization; and the kind of symbolic representations and operations that get implemented in these kinds of networks are typically quite impoverished. Such an unsatisfactory computational compromise imposes severe limits on a vehicle for advancing the theory of language, we believe, although advocates of the approach might well make a virtue of such constraints.

Taken in the broad context of research relating connectionism to higher-level cognition in general and to language in particular, we believe that the integrated connectionist/symbolic computational techniques we are developing in this research will be among those that make an enduring contribution, either directly through long-term survival or by leading in turn to better methods. The Sub-Symbolic Paradigm, we argue, makes possible significant progress towards achieving the Three Goals, and towards the resolution of the Central Paradox of Cognition.

---

## Notes

1. For example, the phonological component of a grammar might receive as its input a string of phonemes constructed in the morphological component by concatenating the phonemes of a verb stem with the phonemes of a verb ending. The phonological component's job might then be to output a structure in which (a) some phonemes may have been altered to meet various phonological constraints in the language, and (b) hierarchical structure has been added which groups phonemes into syllables, syllables into metrical feet, etc., and (c) accentual structure has been added, marking varying degrees of stress on syllables, etc. Another example would be the syntactic/semantic component of a grammar, which might receive as input a string of word tokens, and might produce as output a structure in which (a) the words are tagged as to lexical class; (b) the string is parsed into phrases; and (c) semantic structures capturing various aspects of the string's meaning are included.

2. This principle is conceptually related to ideas applying Harmony Theory to linguistics which have been proposed in Goldsmith's *harmonic phonology*[28,29] and Lakoff's *cognitive phonology*[30,31] although the formal development described here has no counterpart in either Lakoff's or Goldsmith's work to date.

3. We chose unaccusativity phenomena as our first testbed because we rated it well according to eight criteria: (a) it is one facet of a central issue in syntax/semantics (argument structure); (b) it is a phenomenon well-studied by linguists..., (c) yet it is not completely understood, and offers the potential for new contributions; (d) it sometimes exhibits strong, complex interactions of multiple syntactic and semantic factors (as we have shown in our research in French), and so lends itself to one of the strengths of HG; (e) it involves some syntactic structure (as we and others have argued),... (f) but it can be approached without dealing explicitly with issues of embedding (which the technical core of HG was not yet well equipped to handle when we began this work),... (g) yet it can serve as a stepping stone for subsequent research on related phenomena that do crucially involve embedding; and finally, (h) it is a domain in which Legendre has been working for some years, using traditional symbolic

methods[35]. It is a pragmatic entailment of (a–b) that unaccusativity is a controversial phenomenon, and we assert (c–g) not because these claims are undisputed, but because we feel we have solid arguments to defend them—in the particular case of French, where our HG work to date has focussed. As suggested in (d–g), unaccusativity has been useful in the early stages of development of HG because it allows us to test whether HG can deal with the complexity of *interactions* that involve syntactic structure, without requiring that we deal explicitly with complexity in the syntactic structure itself; yet it gets a foot in the door leading to such complexity as well.

4. For example, on this view, the argument structure of the passive verb *was kissed* tells us that in the sentence *John was kissed by Mary*, the argument of *was kissed* filling the syntactic role of *subject*—*John*—plays the semantic role of *patient* in the correct semantic interpretation: the sentence describes an event in which John *receives* a kiss.

5. An example of type (a) is a rule that says the well-formedness of a *croire* construction is diminished if the target NP is not a deep direct object; these rules embody the syntactic constraints on the relevant constructions that derive from the syntactic approach. In this syntactically simplified account, the only syntactic factors appearing in these rules is the construction in which the intransitive is embedded, and a non-surface ('hidden') variable which we interpret as the DGF of the intransitive's argument. An example of a semantic rule of type (b) is one asserting that the well-formedness of *croire* constructions is reduced if the target NP is volitional; these rules capture the semantic and aspectual tendencies of the semantic approach. (The properties figuring in these rules are telicity, progressivizability, volitionality, animacy, and definiteness; our study of the French data has shown acceptability judgments to be sensitive to all these factors.) Finally, an example of a linking rule of type (c) is one stating that the well-formedness of a structure is increased if the argument of a progressivizable predicate is assigned the DGF direct object.

6. In most languages, a vowel (consonant) must be parsed as the nucleus (onset or coda) of a syllable, so the possible parses of a phoneme string into syllables is fairly constrained. However, in Berber[43], virtually every phoneme can be parsed into any syllable position, greatly increasing the complexity of the syllabification process, descriptively as well as algorithmically.

7. Indeed, prominent generative phonologists have even tried to argue that such a theory is nonexistent[44].

## References

1. Smolensky, P., On the proper treatment of connectionism, *Behav. Brain Sci.*, 1988, **11**, 1–74.
2. Legendre, G., Miyata, Y. and Smolensky, P., Distributed recursive structure processing, in *Advances in Neural Information Processing Systems* 3 (eds. Touretzky, D. S. and Lippman, R.), Morgan Kaufmann, San Mateo, CA, 1991, pp. 591–597. Slightly expanded version in *Scandinavian Conference on Artificial Intelligence-91* (ed. Mayoh, B.), IOS Press, Amsterdam, pp. 47–53.
3. Smolensky, P., On variable binding and the representation of symbolic structures in connectionist systems, Technical report, Department of Computer Science, University of Colorado at Boulder, February 1987, Technical Report CU-CS-355-87.
4. Smolensky, P., Tensor product variable binding and the representation of symbolic structures in connectionist networks, *Artif. Intell.*, 1990, **46**, 159–216.
5. Dolan, C. P. and Smolensky, P., Tensor Product Production System: A modular architecture and representation, *Connection Sci.*, 1989, **1**, 53–68.
6. Cohen, M. A. and Grossberg, S., Absolute stability of global pattern formation and parallel memory storage by competitive neural networks, *IEEE Trans. Syst. Man, Cybern.*, 1983, **13**, 815–825.
7. Golden, R. M., The 'Brain-State in a Box' neural model is a gradient descent algorithm, *Math. Psychol.*, 1986, **30–31**, 73–80.
8. Golden, R. M., A unified framework for connectionist systems, *Biol. Cybern.*, 1988, **59**, 109–120.
9. Hinton, G. E. and Sejnowski, T. J., Analysing cooperative computation, in *Proceedings of the Fifth Annual Conference of the Cognitive Science Society*, Erlbaum Associates, Rochester, NY, May 1983.
10. Hinton, G. E. and Sejnowski, T. J., Learning and relearning in Boltzmann machines, in *Parallel Distributed Processing: Explorations in Microstructure of Cognition, Volume 1: Foundations*, (eds. Rumelhart, D. E., McClelland, J. L. and the PDP Research Group), MIT Press/Bradford Books, Cambridge, MA, 1986, pp. 282–317.
11. Hopfield, J. J., Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 2554–2558.
12. Hopfield, J. J., Neurons with graded response have collective computational properties like those of two-state neurons, *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 3088–3092.
13. Hopfield, J. J., Learning algorithms and probability distributions in feed-forward and feed-back networks, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 8429–8433.
14. Legendre, G., Miyata, Y. and Smolensky, P., Harmonic Grammar—A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations, in *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum, Cambridge, MA, July 1990, pp. 388–395.
15. Smolensky, P., Schema selection and stochastic inference in modular environments, in *Proceedings of the National Conference on Artificial Intelligence*, Washington, DC, August 1983, pp. 378–382.
16. Smolensky, P., Information processing in dynamical systems: Foundations of Harmony Theory, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations* (eds. Rumelhart, D. E., McClelland, J. L. and PDP Research Group), MIT Press/Bradford Books, Cambridge, MA, 1986, pp. 194–281.
17. Legendre, G., Miyata, Y. and Smolensky, P., Harmonic Grammar—A formal multi-level connectionist theory of linguistic well-formedness: An application, in *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum, Cambridge, MA, July 1990, pp. 884–891.
18. Legendre, G., Miyata, Y. and Smolensky, P., Can connectionism contribute to syntax? Harmonic Grammar, with an application, in *Proceedings of the 26th Meeting of the Chicago Linguistic Society*, (eds. Deaton, K., Noske, M. and Ziolkowski, M.), Chicago, IL, April 1990, in press.
19. Legendre, G., Miyata, Y. and Smolensky, P., Unifying syntactic and semantic approaches to unaccusativity: A connectionist approach, in *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, (eds. Sutton, L. and Johnson, C. with Shields, R.), Berkeley, CA, February 1991, in press.
20. Prince, A. and Smolensky, P., Harmony-Theoretic Phonology: Optimality and constraint interaction in generative grammar. In preparation.
21. Prince, A. and Smolensky, P., Notes on connectionism and Harmony Theory in linguistics. Technical report, Department of Computer Science, University of Colorado at Boulder, July 1991, Technical Report CU-CS-533-91.

22. Fodor, J. A. and Pylyshyn, Z. W., Connectionism and cognitive architecture. A critical analysis, *Cognition*, 1988, **28**, 3–71.
23. Smolensky, P., The constituent structure of connectionist mental states: A reply to Fodor and Pylyshyn. *Southern J. Philos.*, 1987, **26** (Supplement), 137–163. (Reprinted in *Connectionism and the Philosophy of Mind* (eds Horgan, T. and Tienson, J), Kluwer Academic, Dordrecht, 1991, pp. 281–308 and *The Philosophy of Psychology: Debates on Psychological Explanation*, (eds. Macdonald, C. and Macdonald, G.), Basil Blackwell, Oxford, to appear.
24. Smolensky, P., Connectionism, constituency and the language of thought, in *Meaning in Mind: Fodor and his Critics* (eds. Loewer, B. and Ray, G), Basil Blackwell, Oxford, 1991, pp. 201–227.
25. Dolan, C. P., Tensor Manipulation Networks: Connectionist and Symbolic Approaches to Comprehension, Learning, and Planning. Ph D thesis, Department of Computer Science, University of California, Los Angeles, CA, June 1989.
26. Dolan, C. P. and Dyer, M. G., Symbolic schemata, role binding, and the evolution of structure in connectionist memories, in *Proceedings of the IEEE First International Conference on Neural Networks*, San Diego, CA, June 1987, vol. II, pp. 287–298.
27. Shieber, S. M., *An Introduction to Unification-Based Approaches to Grammar*, Center for the Study of Language and Information, Stanford, CA, and University of Chicago Press, Chicago, IL, 1986.
28. Goldsmith, J. A., *Autosegmental and Metrical Phonology*, Basil Blackwell, Oxford, 1990.
29. Goldsmith, J. A., Phonology as an intelligent system, in *Bridges between Psychology and Linguistics: A Swarthmore Festschrift for Lila Gleitman* (eds. Napoli, D. J. and Kegl, J. A.), Cambridge University Press, Cambridge, in press.
30. Lakoff, G., A suggestion for a linguistics with connectionist foundations, in *Proceedings of the Connectionist Models Summer School* (eds. Touretzky, D., Hinton, G. E. and Sejnowski, T. J.), San Mateo, CA, Morgan Kaufmann, 1988, pp. 301–314.
31. Lakoff, G., Cognitive phonology. Paper presented at the UC-Berkeley Workshop on Rules and Constraints, May 1989.
32. Rumelhart, D. E., Hinton, G. E. and Williams, R. J., Learning internal representations by error propagation, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations* (eds. Rumelhart, D. E., McClelland, J. L. and the PDP Research Group), MIT Press/Bradford Books, Cambridge, MA, 1986, pp. 318–362.
33. Legendre, G. and Rood, D., On the interaction of grammar components in Lakhota: Evidence from split intransitivity, in *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society*, February 1992.
34. Williamson, J. S., Patient marking in Lakhota and the Unaccusative Hypothesis, in *Proceedings of the Fifteenth Meeting of the Chicago Linguistic Society*, Chicago, IL, 1979
35. Legendre, G., Unaccusativity in French, *Lingua*, 1989, **79**, 95–164.
36. Legendre, G., Split intransitivity: A reply to Van Valin 1990. Technical Report ICS-TR-92-3, University of Colorado Institute of Cognitive Science, 1992.
37. Burzio, L., *Italian Syntax: A Government-Binding Approach*, Reidel, Dordrecht, Holland, 1986.
38. Perlmutter, D. M., Impersonal passives and the Unaccusativity Hypothesis, in *Proceedings of the Fourth Berkeley Linguistic Society Meeting*, 1978.
39. Perlmutter, D. M., Multiattachment and the Unaccusative Hypothesis: The perfect auxiliary in Italian, *Probus*, 1989, **1**, 63–119.
40. Levin, B. and Rappaport, M., An approach to unaccusative mismatches, in *Proceedings of the Nineteenth Meeting of the North Eastern Linguistic Society*, 1989.
41. Dowty, D., Thematic proto-roles and argument selection, *Language*, 1991, **67**, 547–619.
42. Van Valin, R. D., Semantic parameters of split intransitivity, *Language*, 1990, **66**, 221–260.
43. Dell, F. and Elmedlaoui, M., Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber, *J. Afr Lang. Ling.*, 1985, **7**, 105–130.
44. Bromberger, S. and Halle, M, Why phonology is different, *Linguistic Inquiry*, 1989, **20**, 51–70.
45. Berg, G., Learning recursive phrase structure: Combining the strengths of PDP and X-Bar Syntax. Technical Report 91-5, Department of Computer Science, State University of New York and Albany, 1991.
46. Jain, A. N., Parsing complex sentences with structured connectionist networks, *Neural Computation*, 1991, **3**, 110–120.
47. McClelland, J. L. and Kawamoto, A. H., Mechanisms of sentence processing: Assigning roles to constituents, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 2: Psychological and Biological Models* (eds. Rumelhart, D. E., McClelland, J. L. and the PDP Research Group), MIT Press/Bradford Books, Cambridge, MA, 1986, pp. 272–325.
48. Miikkulainen, R. and Dyer, M. G., Encoding input/output representations in connectionist cognitive systems, in *Proceedings of the Connectionist Models Summer School*, (eds. Touretzky, D., Hinton, G. E. and Sejnowski, T. J.), Morgan Kaufmann, San Mateo, CA, 1988, pp. 347–356.
49. Mozer, M. C., A focussed back-propagation algorithm for temporal sequence recognition, *Complex Syst.*, 1990, **3**, 349–381.
50. Pollack, J. B., Recursive auto associative memory: Devising compositional distributed representations, in *Proceedings of the Tenth Annual Meeting of the Cognitive Science Society*, Erlbaum Associates, Montreal, Canada, August 1988.
51. Pollack, J. B., Recursive distributed representation, *Artif. Intell.*, 1990, 46, 77–105.
52. Rumelhart, D. E. and McClelland, J. L., On learning the past tenses of English verbs, in *Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models* (eds. McClelland, J. L., Rumelhart, D. E. and the PDP Research Group), MIT Press/Bradford Books, Cambridge, MA, 1986, pp. 216–271.
53. Servan-Schreiber, D., Cleeremans, A. and McClelland, J. L., Graded state machines: The representation of temporal contingencies in simple recurrent networks, *Mach. Learning*, 1991, **7**, 161–194.
54. St. John, M. and McClelland, J. L., Learning and applying contextual constraints in sentence comprehension, *Artif. Intell.*, 1990, 46, 217–258.
55. Elman, J. L., Distributed representations, simple recurrent networks and grammatical structure, *Machine Learning*, 1991, **7**, 195–226.
56. Goldsmith, J. A., Local modeling in phonology, in *Connectionism: Theory and Practice* (ed. Davis, S), Oxford University Press, Oxford, in press.
57. Goldsmith, J A. and Larson, G., Local modeling and syllabification, in *Proceedings of the 26th Meeting of the Chicago Linguistic Society: Parasession on the Syllable in Phonetics and Phonology* (eds. Deaton, K., Noske, M. and Ziolkowski, M.), Chicago, IL, April 1990, in press.
58. Hare, M., The role of similarity in Hungarian vowel harmony: A connectionist account, *Connection Sci.*, 1990, **2**, 123–150.
59. Larson, G., Local computational networks and the distribution of segments in the Spanish syllable, in *Proceedings of the 26th Meeting of the Chicago Linguistic Society: Parasession on the Syllable in Phonetics and Phonology* (eds. Deaton, K., Noske, M. and Ziolkowski, M), Chicago, IL, April 1990, in press.
60. Rager, J. and Berg, G., A connectionist model of motion and government in Chomsky's government-binding theory, *Connection Sci.*, 1990, **2**, 35–52.
61. Touretzky, D. S., Towards a connectionist phonology: The 'many

maps' approach to sequence manipulation, in *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum, Ann Arbor, MI, August 1989, pp. 188–195.

62. Touretzky, D. S. and Wheeler, D. W., Sequence manipulation using parallel mapping networks *Neural Comput.*, 1991, 3, 98–109.
63. Wheeler, D. W. and Touretzky, D. S., A connectionist implementation of cognitive phonology, in *The Last Phonological Rule*, (ed. Goldsmith, J. A.), University of Chicago Press, Chicago, IL, in press.

ACKNOWLEDGEMENTS. One of the authors, P. S., wishes to acknowledge that this research has been supported by NSF grants IRI-8609599, ECE-8617947, IST-8609599, BNS-9016806, and DBS-9209265; by the Sloan Foundation's computational neuroscience program; by the Optical Connectionist Machine Program of the University of Colorado Center for Optoelectronic Computing Systems and by the University of Colorado at Boulder Council on Research and Creative Work. G. L. wishes to acknowledge that the research presented here was partly supported by NSF grant DBS-9209265 and by the University of Colorado at Bounder Council on Research and Creative Work. Y. M. wishes to acknowledge that the research reported here has been supported by the Optical Connectionist Machine Program of the University of Colorado Center for Optoelectronic Computing Systems (sponsored in part by NSF/ERC grant CDR-8622236 and by the State of Colorado Advanced Technology Institute).

---

# Japan's Fifth Generation Computer Project: A summary

## The FGCS project

In 1982, the Japanese Ministry of Trade and Industry (MITI) initiated a ten-year project called the Fifth Generation Computer Systems (FGCS) Project. This Project received funding of over 50 billion Yen (over 300 million US dollars) from the Japanese government and major support from the industry. The Project was aimed at developing large-scale parallel-processing machines for symbolic (non-numerical) computation. There was a definite focus on using logic programming as the machine language of these machines.

The Institute for New Generation Computer Technology (ICOT) was set up by MITI in Tokyo as part of the FGCS Project. ICOT is a consortium of Japanese companies and MITI Institutes. The Project was managed from ICOT, and most of the research was done at ICOT. Researchers from various Japanese companies spent a few years each at ICOT working on FGCS projects. There were also a number of international visitors to ICOT who participated in projects.

The Project generated enormous interest in the West, because of its technical and economic implications. Partly as a result of this, a number of other high technology computing projects were started. These included Britain's Alvey programme, the European Community's ESPRIT programme and the American Microelectronics Computer Consortium (MCC).

## Fifth generation computers

The notion of generations has been associated with the hardware technologies which have been used to develop computer systems, ranging from vacuum tubes through transistors and integrated circuits to LSI and VLSI chips. These computers can be thought to consist essentially of a central processing unit (CPU) linked to some memory. The memory units store the data that can be accessed using some notion of addresses. This memory is also used to store programs. Each program consists of a series of instructions. Program execution consists of executing each instruction in turn, using data obtained from the computer's memory. A program counter keeps track of the current instruction. In this so-called von-Neumann model, programs are procedural in that they specify to the computer (in great detail) how each problem is to be solved. These computers are usually sequential processing machines, and a majority of them are oriented towards numerical calculation. These machines are rated in terms of the number of machine language instructions executed per second, for example in millions of instructions per second (MIPS).

The term *Fifth Generation* was used to differentiate the Japanese Project from these earlier generations of computers. The fifth generation computers were aimed at symbolic computations and a high degree of parallelism. The basic thesis was that a significant number of problems today are not really numerical—they really involve symbolic computation, where it is *knowledge* that is manipulated, rather than numbers. Again, a number of problems can be solved with multiple processors working in parallel. ICOT's main goal was to develop an easier-to-use non-von Neumann computer system with parallel processing ability and an ability to use knowledge and make inferences when executing a program.

One of the major features of the FGCS Project was the decision to use a logic programming language similar to *Prolog* as the machine language of the computer. To achieve this, the Japanese developed and used a language named KL1. Logic Programming is based on the use of logic as a declarative language, in which it is possible (in an abstract sense) to specify *what* is to be solved, instead of *how* (step by step) it is to be solved. The speed of such machines is measured in LIPS (logical inferences per second).

The programming for this system and the design of the hardware was done by ICOT. The actual manufacture of the hardware was done by the Japanese companies.