FUNCTIONALLY ACCURATE DISTRIBUTED PROBLEM SOLVING SYSTEMS

Victor R. Lesser and Daniel D. Corkill

Computer and Information Science Department
University of Massachusetts at Amherst
Amherst, Massachusetts, 01003

1. INTRODUCTION AND OUR APPROACH

There is a large class of applications that seem naturally suited to the use of distributed processing hardware for which adequate distributed problem solving methodologies have not yet been developed. Many of these applications occur in situations that have a natural spatial distribution (e.g., where sensors for collecting raw data are widely distributed and/or mobile) where the development and execution of a distributed control strategy based on a non-local view of sensory data is required. In these applications, a distributed architecture which locates processing capability at the sensor site and which requires only limited communication among processors is especially advantageous and perhaps is the only practical way to perform the processing. Sensor networks (e.g., low flying aircraft, hydrophones, etc.), distributed command and control systems, air traffic control, automotive traffic light control, inventory control (e.g., car rentals), power network grids, and tasks involving mobile robots are all examples of this type of application.

The central theme of our research is that the exact emulation of a centralized approach to these applications in distributed environments is, in general, inappropriate due to the high communication and synchronization costs required to maintain sufficient and consistent local views of the overall problem solving database. However, if one is led to emulating the centralized approach only approximately, inter-node communication can be reduced, but at the risk of inconsistency and incompleteness of local views and the possibility of unnecessary, redundant, or incorrect processing. The key issue is whether an acceptable design exists which has sufficiently low communication costs and which still allows the overall system performance to be accurate and efficient enough to satisfy the task requirements.

The long range goal of our research is to develop a general theory for "functionally accurate" problem solving structures in which all intermediate aspects of the computation are required to be correct and consistent.

Such a general theory is also important to the implementation of complex applications in centralized environments. These applications are often organized in the form of a collection of independent modules. In such a structure it is often conceptually difficult to develop and expensive to maintain a complete and consistent centralized problem solving data base with which the modules interact. A theory which permits relaxation of completeness and consistency requirements would be a significant aid in the development and maintenance of these systems.

One example of algorithmic structures which exhibit functionally accurate behavior are those that involve the incremental aggregation of partial results into an overall solution. "Knowledge-based" Artificial Intelligence (AI) interpretation algorithms developed recently for speech, image, and signal interpretation applications [Erman and Lesser 1975, Barrow and Tenenbaum 1976, Rosenthal, Rumel, and Zucker 1976, Rubin and Reddy 1977] use this incremental aggregation approach. Errors and uncertainty in input data and from incomplete or incorrect knowledge are resolved as an integral part of the interpretation process by testing for inconsistency among partial results. This is in contrast to more conventional techniques for error recovery in which errors are handled as exceptional conditions, requiring processing outside the normal problem solving strategy.

The ability to handle errors within the normal problem solving process itself creates the possibility that these systems will also handle the additional uncertainty introduced by a distributed decomposition: the restricted and inconsistent local view of the overall problem solving data base at each node. To the degree that systems can handle these problems, the need for a complete and consistent local view of the overall data base and for explicit synchronization to maintain such a view is reduced.

Preliminary research in the development of functionally accurate cooperative distributed problem solving techniques (discussed in Section 3) has produced encouraging results which we feel justify further explorations. As outlined in Section 4 on current research directions, we feel generic research in the areas of distributed search and distributed planning will be important for the development of a general theory of functionally accurate cooperative distributed systems.

2. STATE OF THE ART AND OUR APPROACH

Most distributed processing systems work on
local databases which contain exact copies of appropriate portions of the overall problem solving data base. We call this type of distributed processing decomposition nearly autonomous, because each processor usually has the information it requires to complete processing; a processor rarely needs the assistance of another processor in carrying out its problem solving function.

In nearly autonomous systems, the problem solving data base is distributed in such a way that the centralized problem solving algorithms are not significantly perturbed. Therefore, the basic algorithms and control structures which are appropriate in a centralized environment do not need to be modified (decomposed) for the distributed environment, but rather can be replicated or partitioned based on the distribution of the problem solving data base.

Where more complex control structures have been introduced in these distributed systems, their use has been directed at maintaining data base consistency (synchronization, deadlock detection and avoidance) and error recovery. This results in a two-level structure in which the more complex distributed control structures are superimposed on the basic computational algorithms needed to perform the application processing [Eswaran et al. 1976, Thomas 1976, Lampson and Sturgis 1977, Peebles 1977].

An alternative and new approach to structuring distributed problem solving systems is the development of functionally accurate cooperative distributed systems. In cooperative distributed systems, additional complexity over the nearly-autonomous structure is required because the algorithms and control structures are decomposed to operate on local data bases which are incomplete and possibly errorful. In order to resolve these uncertainties in these incomplete local data bases, processors must exchange partial results with each other. Since new information may also be based on processing which used incomplete or incorrect data, an iterative, coroutine type of processor interaction is required. If a processing element does not receive an appropriate partial result in a given amount of time, it must be able to continue processing, utilizing whatever data are available at that time.

A side effect of this type of interaction is that local problem solving tasks in the cooperative decomposition are not disjoint as in the nearly-autonomous structure. Rather, the tasks overlap in the data they need and produce, forming a "cooperative network of tasks" which collectively define the processing needed to solve the overall task. This perspective shifts the view from that of a system distributed over a network to that of a functionally accurate network of cooperating systems, each of which can perform significant local processing on incomplete and inconsistent information. From this viewpoint, the key issue becomes how to introduce new communication paths to handle the inter-node interactions needed for sharing the information required to correct for incomplete and inconsistent local views.

This perspective is essentially identical to the theory of "nearly decomposable systems" devised by Simon (1962, 1969) to describe complex organizational structures. The term "nearly decomposable" emphasizes the fact that systems can be decomposed naturally into clusters which have a high degree of intra-cluster interaction and a lower degree of inter-cluster interaction.

Another effect of non-localized processing in cooperative systems is that reliability and flexibility issues cannot be addressed solely at the hardware and communication level (syntactically) [1], but must also be dealt with as an integral part of the basic problem solving process (semantically). In a nearly-autonomous system, it is easier to detect, isolate, and recover from an error since propagation of the error can be more easily determined (due to the simpler interaction structure) and appropriate recomputation performed. This is in contrast with cooperative systems where incorrect partial results might be propagated extensively through the system before detection if not corrected as part of the normal problem solving process.

No software technology for building cooperative distributed problem solving systems which exhibit functionally accurate behavior has yet been developed. Current methodologies for distributed applications, such as distributed data bases, are not directly applicable to cooperative applications due to the relatively complex interaction required for cooperative tasks. These methodologies also fail to deal with some issues crucial to cooperative distributed computation, such as working with incomplete and inconsistent data.

3. PAST AND PRESENT WORK

Under previous support from ARPA [2] and current support from NSF, a number of functionally accurate computational techniques (e.g., relaxation and distributed hypothesis-and-test ala Hoare's II) developed in the framework of know-

---

[1] Work has also been performed on developing communication protocols that are robust in the face of communication and site failures. [Cerf and Kahn 1974, Kimbleton and Schneider 1975, Tajbapis 1976] and on embedding message passing protocols into existing programming languages and operating systems [Farber et al. 1973, Feldman 1977, Posdick, Schantz, and Thomas 1977, Jefferson 1977]. When viewed from the perspective of the routing task alone, some algorithms used to determine message paths in a communication network have a cooperative problem solving character to them [Gerla 1973, Tajbapis 1977].

[2] This was joint work with Lee Erman at Carnegie-Mellon University.
logic-based AI systems are being suitably modified for use in distributed problem solving environments. In order to test whether these modified techniques have relevance to distributed problem solving, a number of distributed applications using these techniques are being developed.

1.1 DISTRIBUTED AUTOMOTIVE TRAFFIC LIGHT CONTROL

One study has concentrated on investigating the suitability of relaxation as the basis of a cooperative distributed approach to automotive traffic light control. Parallel relaxation has been explored in previous work in image processing, but to our knowledge it has not been explored in distributed domains which have ill-behaved compatibility relationships between nodes and non-localized interactions among nodes in the network.

A test program was written which simulates a distributed relaxation version of traffic control in which the knowledge applied at each node is similar to that used in a standard centralized traffic control system called SICOP-21 [Leiberman and Yoo 1976]. The simulations have produced encouraging results for a number of test cases. However, in a number of cases the parallel scheme does not perform as well as the centralized version, resulting in slow convergence, convergence to near optimal solutions, or oscillation.

These problems are directly attributable to changes in the centralized version, a global node ordering for the relaxation is precomputed using a maximal spanning tree based on the traffic volume at each node. This ordering is very important in reducing the effects of ill-behaved compatibility relationships between nodes and non-localized interactions among nodes. Current research is aimed at introducing, in a distributed way, non-local coordination between nodes to eliminate these problems. We are examining such techniques as multi-level relaxation and distributed versions of the maximal spanning tree heuristics.

A more detailed discussion of this research is contained in a forthcoming report by Brooks and Lesser [1978].

1.2 AN EXPERIMENT IN DISTRIBUTED INTERPRETATION

A second study concentrated on applying the Hearsay-II architecture to the distributed interpretation problem, where each processor is mobile, has a set of (possibly non-uniform) sensing devices, and interacts with nearby processors through a packet-radio communication network. It is desired that processors communicate among themselves to generate a consistent interpretation of "what is happening" in the environment being sensed.

Our approach to decomposing Hearsay-II in the packet-radio network environment is based on the following premise: the only cost-effective way to decompose Hearsay-II is to modify the problem solving strategy to work with processors operating on partial and possibly inconsistent views of the current interpretations and system state.

There are two major ideas necessary to implement the distributed version of the Hearsay-II architecture. The first comes from the relaxation paradigm, where the current state (result) of a processor node is accessed only by nodes that are neighboring in the node network. If the current state of the node is "important" or "relevant", the state will be transmitted gradually to other areas of the node network, resulting in what can be thought of as a spreading excitation of important information. This same approach may be used to decentralize the Hearsay-II blackboard. Not all processor nodes have to immediately receive all information that is pertinent to them; if the information is really important, it will eventually be spread through the network to all nodes.

The second major idea comes from viewing a processor node as a generator function which can be successively re-invoked as needed to generate additional (and possibly less credible) hypotheses for a particular portion of the interpretation. Generator function re-invocation is achieved through the use of threshold values (part of the focus of control data base) which are modified by the same spreading excitation scheme used to transmit important information.

Using these two ideas, we have developed a system which permits effective inter-node cooperation without the high communication bandwidth required to support a completely centralized data base and controller. This lower bandwidth requirement is accomplished by transmitting only a limited subset of the results generated at each node to only a small number of other processing nodes.

A set of experiments to determine how the distributed decomposition of Hearsay-II affects its problem solving behavior is now being analyzed. Aspects of behavior studied include the accuracy of the interpretation, the time required for interpretation, the amount of inter-node communication required, and the robustness of the system in the face of communication hardware malfunction (i.e., communication without positive acknowledgment).

These experiments were only in part simulations, since they used an actual interpretation system analyzing real data: the Hearsay-II speech understanding system [Erman and Lesser 1978]. Tentative results of these experiments indicate that a simplified version of the distributed Hearsay-II architecture as applied to speech data (i.e., different nodes, each with a complete set of knowledge sources, processing overlapping segments of the acoustic data) performs as well as the centralized system. The distributed system produces these results using a low degree of interprocessor communication (fewer than 50% of the locally generated hypotheses need to be transmitted). In addition, the distributed system con-
continues to function, although in a somewhat degraded manner (i.e., lower recognition accuracy), with a noisy interprocessor communication channel (a randomly selected 25% of transmitted messages lost).

A forthcoming paper by Lesser and Erman [1978] further details this research.

4. CURRENT RESEARCH DIRECTIONS

Our exploratory research in the development of functionally accurate cooperative distributed problem solving techniques has produced promising results. This work has highlighted many key design issues that must be dealt with in the development of a general theory for functionally accurate cooperative distributed systems. We believe that sufficient intuition has been obtained in the initial stages of this research to direct the investigation to a set of more generic research topics in knowledge-based AI problem solving and in organizational theory. The results of this research will be important in shaping a general theory.

Further investigation of specific applications remains appropriate to the continuing development of intuitions regarding the key design issues. However, the selection of these should be done in a more directed way, relating directly to generic research topics. The emphasis of our current research is on an investigation of the following two generic topics.

4.1 DISTRIBUTED SEARCH

Experiences with the two distributed interpretation applications discussed in Section 1 has led us to the conjecture that all functionally accurate cooperative distributed structures have at their heart distributed search. Distributed search involves the integration of partial results emanating from multiple, semi-independent (disconnected) loci of search control. An adequate model for describing and analyzing distributed search techniques has not been developed. This lack of a developed model of distributed search exists even in centralized environments such as the Mesaray-II speech understanding system.

We are developing a model for distributed search that provides a common framework for addressing the following questions:

*What is the nature of the search space in a given application?

*How is the search space represented in the system (levels of abstraction, data structure, etc.)?

*What is the nature of a partial result in this representation?

*How are partial results extended and merged together?

*What is the nature of information that must be communicated between semi-independent loci of search control?

*How are certain types of error resolved as part of the search process?

*How are the semi-independent loci of search control coordinated?

*How is the search space searched [3]?

We are developing this model by generalizing four distributed search techniques: relaxation, the locus model, cooperating experts, and hypoth-

esis-and-test (ala Heasay-II). We hope this model will have a taxonomic character that will provide a framework in which new, alternative search techniques become apparent. The model should also lead to the development of a small set of control primitives and data structures which are appropriate for all types of distributed search techniques and applications.

4.2 DISTRIBUTED PLANNING

Experiences with the two distributed interpretation applications has also lead us to understand that distributed focus of attention is a crucial aspect of all functionally accurate cooperative distributed systems. Distributed focus of attention involves the dynamic allocation of processing power, memory, data, and communication capability. We feel distributed focus of attention is part of the larger issue of distributed planning (focus of attention is planning which is directed at one's own immediate internal processing).

To approach distributed planning and distributed focus of attention in the same way as we approach distributed search (i.e., the development of a generic model) seems premature. There is much less experience with distributed focus of attention and distributed planning, and sufficient intuitions have not yet been developed. Therefore, a more directed, experimental approach in this area seems appropriate.

We are engaged in two investigations to help develop these needed intuitions. The first investigation involves determining the implications of reorganizing, for a distributed environment, two existing focus of attention techniques: shortfall density scoring [Woods 1977] and the Hayes-Roth and Lesser focus of attention mechanism [Hayes-Roth and Lesser 1977]. The shortfall density

[3] Current characterizations of search using such terms as breadth-first and depth-first are inadequate even in centralized environments.
scoring technique is an admissible focus of attention strategy and the Hayes-Roth and Lesser mechanism is a heuristic strategy based on a number of general focus of attention principles.

The second investigation approaches the more general issue of distributed planning. The analogy between most current planning techniques and distributed planning is similar to that between heuristic search and distributed search (i.e., in most planning systems there is only one locus of control).

We are taking existing AI planning techniques and generalizing them to more than one locus of control in which the planning processes in a functionally accurate cooperative manner. We are also taking models of group planning developed in Organizational Theory and placing them in a computational framework.

5. CONCLUSION

We feel methodologies can be developed for functionally accurate cooperative distributed systems in which the distributed algorithms and control structures function with both inconsistent and incomplete data. These methodologies are necessary in order to extend the range of applications that can be effectively implemented in distributed environments. Our current generic research studies in the areas of distributed search and distributed planning as outlined in the paper will aid in the development of a general theory for functionally accurate cooperative distributed systems.

6. REFERENCES


Annual Brown University Workshop on Distributed Computation.


