AD-A261 617

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED FINAL/01 DEC 88 TO 30 APR 92

4. TITLE AND SUBTITLE COMPLEXITY OF CONNECTIONIST AND CONSTRAINT- SATISFACTION	5. FUNDING NUMBERS 2304/CS AFOSR-89-0151
6. AUTHOR(S) DR. KASIF	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) JOHNS HOPKINS UNIVERSITY HOMEWOOD RESEARCH ADMINISTRATION CHARLES AND 34TH STREET BALTIMORE MD 21218	8. PERFORMING ORGANIZATION REPORT NUMBER AFOSR-TR- 93 0074
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR/NM 110 DUNCAN AVE, SUTE B115 BOLLING AFB DC 20332-0001	10. SPONSORING/MONITORING AGENCY REPORT NUMBER AFOSR-89-0151
11. SUPPLEMENTARY NOTES	
12a. DISTRIBUTION/AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED	12b. DISTRIBUTION CODE

DTIC ELECTE MAR 0 5 1993 S B D

13. ABSTRACT *(Maximum 200 words)*

Since the beginning of the funding of the grant, we established a substantial effort in the area of connectionist optimization algorithms, relaxation networks, and geometrical learning algorithms. All of the above are highly interconnected research projects. We have achieved several significant results that have increased our understanding of the computational capabilities and limitations, of connectionist and constraint network. Our most significant contributions thus far are in the area of parallel complexity of constraint networks, comparative experimentation with learning algorithms and geometric concept learning. Our results in the area of parallel constraint networks are the subject of several publication in first rate journal and conferences. Our experimental research achieves the best results on several well established benchmarks. Most notably our group achieved the best results (in terms of prediction accuracy) in the area of protein folding. The technical result of research investigations are summarized in the following sections.

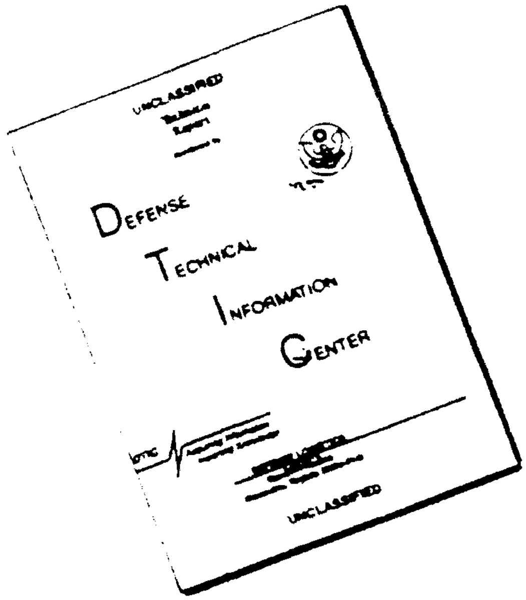
14. SUBJECT TERMS			15. NUMBER OF PAGES 16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAR(SAME AS REPORT)

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. Z39-18 298-102

93 3 4 019

93-04649

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

JOHNS HOPKINS
UNIVERSITY

Homewood Research Administration

105 Ames Hall / 3400 N. Charles Street
Baltimore MD 21218-2686
(410) 516-8668 / FAX (410) 516-7775

30019
P.25

FAX COVER SHEET

TO: Marilyn Ramsey

COMPANY: ONR

FAX NUMBER: (202) 404-7951

FROM: Joan Warfield JW

TITLE: Sponsored Projects Officer

FAX NUMBER: (301) 338-7775 TELEPHONE NUMBER: (301) 338-8668

DATE: January 25, 1993 TIME: 4:15 P.M. OPERATOR: JW

TOTAL NO. OF PAGES (INCLUDING THIS COVER SHEET) 9

SHOULD YOU ENCOUNTER DIFFICULTY IN RECEIVING THIS DOCUMENT, PLEASE CONTACT US AT (301) 338-8668.

MESSAGE: Marilyn, Attached is the requested copy the final report for Grant No. AFSOR-89-0151. As stated in Dr. Kasif's (P.I) letter,his report was forewarded to his program director along with his proposal in June. The hard copy of this document will be sent by first calss mail tomorrow. It is my understanding that the pending NCE will be released as soon as possible.

If you have questions, please contact me at (410) 516-8668.

Prof. Simon Kasif
Department of Computer Science
The Johns Hopkins University
Baltimore, MD 21218
410-516-8296
KASIF@CS.JHU.EDU

January 25, 1993

Dr. Mark Jacobs
Program Director
AFOSR-NM
Bolling Airforce Base
Washingon, DC 20332-6448

Dear Dr. Jacobs

Enclosed please find the final technical report for the research on "Computational Complexity of Connectionist and Constraint Networks" supported by grant AFOSR-89-0151. I sent a similar document in June as part of a new proposal, but apparently it was misplaced. We have several on-going experimental validation projects to test the programs developed with partial support from AFOSR. When this work is completed we shall send you an update. Our research had a strong impact on a number of areas, for example we effectively closed the study of parallel complexity of local consistency in constraint networks and we developed the best known methods for constructing non-axis parallel decision trees. I would appreciate any comments you may have on our research.

Sincerely,

Simon Kasif

Final Report: 1992
Complexity of Connectionist and Constraint-Satisfaction Networks

Simon Kasif
Department of Computer Science
The Johns Hopkins University
Baltimore, Md 21218
KASIF@CS.JHU.EDU
301-338-8296

1 General Progress Summary and Overall Impact

Since the beginning of the funding of the grant, we established a substantial effort in the area of connectionist optimization algorithms, relaxation networks, and geometric learning algorithms. All of the above are highly interconnected research projects. We have achieved several significant results that have increased our understanding of the computational capabilities and limitations of connectionist and constraint networks. Our most significant contributions thus far are in the area of parallel complexity of constraint networks, comparative experimentation with learning algorithms and geometric concept learning. Our results in the area of parallel constraint networks are the subject of several publications in first rate journals and conferences. Our experimental research achieves the best results on several well established benchmarks. Most notably our group achived the best results (in terms of prediction accuracy) in the area of protein folding. The technical results of our research investigations are summarized in the following sections.

2 Theoretical Analysis of Relaxation Networks

Relaxation networks are a special case of constraint satisfaction networks and have been used in optimization, truth maintenance systems, and computer vision. These networks utilize local constraint propagation techniques to achieve local stability (local consistency in propositional constraint networks). One of the stated objectives of the research proposal was to increase our understanding of local optimization (search) techniques in such networks. In [KD92] we provide a complete characterization of achieving local stability (consistency) in symbolic relaxation networks. The results of these investigation are reported in several papers Our results are summarized in a collection of papers [Kas90, Kas86, Kas89, KD91, KD92, KD91]

3 Theoretical Analysis of Discrete Hopfield Nets

We completed our investigation of Discrete Hopfield Nets. Our results are summarized in [KBDS91, KBDS93]. In addition to obtaining several lower bounds for the complexity of finding local minima (stable state) in Hopfield networks, we also studied the complexity of finding a new stable when the network is perturbed by modifying one of the weights by a tiny amount (one bit).

4 Parallel Problem Solving and State Space Search

Our group at Johns Hopkins recently obtained a substantial result on very large state-space search. The first part of this work attacked the problem of 6-piece chess end-game analysis, a long-time outstanding problem in computational chess and AI. The chess end-game analysis was done by Lewis Stiller, a graduate student in our AI group, who implemented an extremely large search-space exploration on the Connection Machine. His program has found 6-piece chess positions that require more than 240 moves to achieve a forced win. The methodology is also applicable to other search problems in symmetry groups. The results are published in [Sti91a, Sti91b, Sti92a, Sti92b]. This work is considered a fundamental breakthrough in the computational chess community, and has received wide publicity in the popular press (*e.g.*, Scientific American, London Times, Washington Post).

As a side project we have discovered an efficient parallel algorithm to compute string statistics for molecular-dynamics simulations [Sti92b]. The algorithm is being used currently on the CM2 in Los Alamos.

Our research program in parallel AI programming also had an important educational side-effect. The PI developed a regular programming workshop on parallel programming the Connection Machine. Graduate and undergraduate students implemented many diverse applications on the CM such as semantic networks, constraint networks, path planning with obstacles, geometric algorithms, a cellular automata model of the heart, psychophysical models and more. Subsequently, many of these students spent summers at National Laboratories such as NRL, SRC, and Los Alamos working on important applications.

5 Desing and Analysis of Algorithms for Machine Learning

With partial funding from the grant we started a set of new investigations of machine learning algorithms. The results of these studies are documented below.

5.1 Comparative Studies of Learning Methods

We are conducting a comprehensive set of experiments that compare (on a a range of real world applications) the performance of backpropagation algorithms to nearest neighbor methods learning and other methods. The experiments are performed by David Heath (funded by the project) and Scott Cost and are co-supervised by Steven Salzberg and the PI. Salzberg studied the comparative effectiveness of the backpropagation algorithm and an instance based learning algorithms on the problem of predicting protein folding. The results are documented in [CS90b]. A comprehensive comparative study of connectionist learning and instance based methods has been completed and is a subject of a of a journal paper [CS90a].

5.2 Experimental Analysis of Backpropagation

We completed experimentation with two heuristics approaches to improve the efficiency and accuracy of backpropagation in connectionist networks. The first approach is based on a scaling (multiple resolution) approach to learning. Instead of teaching the network the exact concept, we train the network on a series of approximations (refinements) of the concept. The first approximation is very crude, and therefore we allow a very large error. As the approximations become more accurate, we require increasing accuracy from the network. The preliminary experiments using this strategy were relatively disappointing in terms of improving performance for various classes of problems. While we found applications where the method improves performance (measured as the number of iterations to learn the concept class) in a majority of cases the new algorithm was comparable in speed and accuracy to the old one. Our findings which will appear in a forthcoming report.

5.3 Geometric Concept Learning

We (as well as many others) observed that there are fundamental links between geometric partitioning algorithms and machine learning. We used combinatorial techniques and computational geometry to study basic properties of learning algorithms. This investigation is the main topic of a Ph.D. thesis by David Heath. David Heath's research has been supervised by Kasif and was fully funded by AFOSR. Preliminary results from his thesis have been published in [HKK+91, SHDK91, Hea91, HKS92] Below we briefly describe these topics.

5.4 Limited Memory Learning

The majority of learning algorithms (including connectionist algorithms) use limited memory (a fixed size network) during learning. We addressed the problem of the complexity of learning when the algorithm is allowed to store only the generalization. It cannot store the entire set of examples and process them off-line. One way to abstract this restriction is by enforcing a limited memory requirement. That is, the algorithm is restricted to some small number of memory locations. We recently completed a paper where we establish fundamental bounds on the number of steps necessary and sufficient to learn concepts when the learning module is not allowed to store all examples. The paper is the first of its kind to establish a fundamental trade off in the number of steps required to learn a concept and the size of the generalization used by the algorithm. Our results are summarized in [HKK+91]

5.5 Learning with Helpful Teacher

Research in theoretical machine learning focuses on the complexity of learning a concept class independent of particular learning architectures. In general, this research does not address the following important questions. What are the strengths and weaknesses of a learning algorithm when applied to a particular learning problem? How do particular algorithms compare with each other? What is the right way to teach a given concept to a particular learning machine? Does providing additional examples always help learn a given concept?

We address some of these questions in the context of nearest neighbour and other learning methods. A theoretical model is introduced where the teacher knows the learning algorithm and chooses examples in the best way possible. We call this the Helpful Teacher model, and note that it often requires a very small number of examples to learn a concept. We prove some lower and upper bounds for a variety of geometric concept classes, This is joint work with Art Delcher, David Heath, and Steven Salzberg Our initial findings on this topic are summarized in [SHDK91] where we also discuss the implications of our results for current experimental research.

We also addressed the problem of the complexity of being a Helfpul teacher. We show that the complexity of presenting the best possible set of examples for geometric problems is NP-hard. In the process of solving this problem we produced several new computational geometry results [Hea91, D.92].

5.6 Learning Oblique Decision Trees

We recently began investigating the following natural generalization of perceptrons. Given two sets of points (e.g., blue/red) in n-dimensions and an integer K, find a hyperplane that partitions the set into two categories such that at most K points are misclassified. We proved this by transformation from the dual problem, namely given a set of N linear inequalities, is there a feasible solutions that excludes at most K inequalities. This problem is NP-complete by an easy transformation from 3SAT.

This result should be contrasted with the existing complexity results on the scalability of learning in networks. The majority of the NP-completeness results such Judd's and Blum and Rivest's that address the representation issue. That is, find a boolean function (threshold circuit) of a certain architecture which is consistent with the training set. But when the network does backpropagation, it usually stops when a certain accuracy is achieved, or alternatively in the case of boolean functions, the network is partially consistent with training set. The question above directly addressed the problem of learning a functions with a given accuracy. Our complexity result shows that this problem is difficult even for a single hyperplane.

We got interested in learning with given accuracy when we were thinking about geometric versions of decision trees, where the decision surfaces are allowed to be linear combinations of features. Such decision trees generate space partitioning by convex polyhedra. The majority of practical work on decision trees seem to address axis-parallel decision trees (e.g., ID3). We recently developed a new algorithm for decision tree induction based on the approximate constraint solving paradigm. We developed a general constraint solving technique that allows us to solve approximate linear programming problems. A simple instance of this class of problems is finding a solution to a set of linear inequalities such that the largest number of such inequalities is satisfied. Our algorithm allows us to synthesize the smallest known decision trees for several applications such as breast cancer diagnosis and classification of astronomical data while preserving the prediction accuracy of previous methods. Our results are described in [HKS92].

6 Summary

To summarize, we have developed a strong research program in several areas. Our main contributions thus far are in the area of parallel complexity of constraint networks, and basic properties of geometric partitioning algorithms with applications to learning.

Our results in the area of parallel constraint networks are the subject of several publications in first rate journals and conferences. Our group's experimental research achieves the best results on several well established benchmarks. Most notably we achived the best performance (in terms of prediction accuracy) in the area of protein folding, improving on the back propagation algorithm. We studied several fundamental geometric partitioning problems that have direct implications on lower and upper bounds on the number of examples needed to learn geometric concepts. We introduced a new model that allows us to derive lower bounds on the number of examples required for learning geometric concepts. The model also allows us to study the computational complexity of teaching rather than learning. We also derived the simplest known complexity result for the scalability of learning in networks, namely approximate learning of a perceptron. We developed a set of programs that construct non-axis-parallel decision trees which are significantly smaller than previously constructed trees for the same domains.

References

[CS90a] S. Cost and S. Salzberg. Exemplar-based learning to predict protein folding. In *Symposium on Computer Applications to Medical Care*, pages 114–118, November 1990.

[CS90b] S. Cost and S. Salzberg. A weighted nearest neighbor algorithm for learning with symbolic features. *Machine Learning Journal (accepted for publication)*, 1990.

[D.92] Heath D. Ph.d. thesis. Technical report, Johns Hopkins University, expected August 1992.

[Hea91] D. Heath. On voronoi covers with applications to machine learning (submitted to a journal). Technical Report JHU/CS-91/08, Johns Hopkins University, 1991.

[HKK+91] D. Heath, S. Kasif, S. R. Kosaraju, S. Salzberg, and G. Sullivan. Learning nested concept classes with limited memory. In *IJCAI*, August 1991.

[HKS92] D. Heath, S. Kasif, and S. Salzberg. Induction of oblique decision trees (submitted for publication). Technical report, Johns Hopkins University, 1992.

[Kas86] S. Kasif. On the parallel complexity of some constraint satisfaction problems. In *Proceedings of the 1986 National Conference on Artificial Intelligence*, August 1986.

[Kas89] S. Kasif. Parallel solutions to constraint satisfaction problems. In *Principles of Knowledge Representation and Reasoning*, May 1989.

[Kas90] S. Kasif. On the parallel complexity of discrete relaxation in constraint networks. *Artificial Intelligence*, pages 275–286, 1990.

[KBDS91] S. Kasif, S. Banerjee, A. Delcher, and G. Sullivan. A note on the complexity of symmetric connectionist networks. In *Symposium on Foundations of Artificial Intelligence*, June 1991.

[KBDS93] S. Kasif, S. Banerjee, A. Delcher, and G. Sullivan. A note on the complexity of symmetric connectionist networks. *Annals of Mathematics and Artificial Intelligence (to appear)*, 1993.

[KD91] S. Kasif and A. Delcher. Analysis of local consistency in parallel constraint networks. In *AAAI Stanford Symposium on Constraint-Based Reasoning*, April 1991.

[KD92] S. Kasif and A. Delcher. Analysis of local consistency in parallel constraint networks. *Artificial Intelligence Journal (accepted for publication)*, 1992.

[SHDK91] S. Salzberg, D. Heath, A. Delcher, and S Kasif. Learning with a helpful teacher. In *IJCAI*, August 1991.

[Sti91a] L. Stiller. Groups graphs and computational symmetry on massively parallel architecture. *Journal of Supercomputing*, pages 99–117, 1991.

[Sti91b] L. Stiller. Some results from a massively parallel retrograde analysis. *ICCA Journal*, pages 129–134, 1991.

[Sti92a] L. Stiller. Massively parallel exhaustive state space search (in preparation). *J. of Adv. Comput. Lab (invited article)*, 1992.

[Sti92b] L. Stiller. Parallelizatio. of a string sequence statistic with application to molecular dynamics (in preparation). Technical report, Johns Hopkins University, 1992.