REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE June 4, 2001	3. REPORT TYPE AND DATES COVERED Final Report 01Aug96-30Sep97

4. TITLE AND SUBTITLE
Exploiting Decision and Policy Models More Effectively through Intelligent Search, Data Mining, and Visualization

5. FUNDING NUMBERS
DAAH04-1-0391
96

6. AUTHOR(S)
Steven O. Kimbrough

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Pennsylvania, Philadelphia, PA

8. PERFORMING ORGANIZATION REPORT NUMBER
20010604-OPIM-1

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
36193.6-CI

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

RECEIVED
JUN 07 2001
By

2001 JUN -6 PM 12:21

20010618 121

13. ABSTRACT (Maximum 200 words)

Candle-lighting analysis encompasses a series of automated techniques for supporting intelligent examination of decision and policy models. Such examinations are well-suited for model validation and for post-evaluation analysis, and can be particularly useful for providing decision makers with insight. The primary alternative to candle-lighting analysis is manually-directed what-if questioning. Such what-if analysis is a cumbersome process, appropriate mainly for small-scale models; it is a process forever at risk of being undermined by human biases and cognitive limitations.

Candle-lighting analysis concepts hold the promise of greatly enhancing the value and effective use of management science models in the U.S. Army, and DoD generally.

Essential to the concept of candle-lighting analysis (CLA) is the creation of multiple databases containing information gleaned from particular formal or mathematical models using intelligent heuristics (including, but not limited to, genetic algorithms). CLA treats models as complex, hyperdimensional structures in decision space. The resulting (hyperdimensional) surfaces need to be explored intelligently and mapped for human decision makers. These explorations yield large data sets, which are stored in a database environment. The databases are then mined in order to give decision makers insights in the problem at hand, and to suggest promising policy options to them.

14. SUBJECT TERMS management science, operations research, model, modeling, validation, verification, decision making	15. NUMBER OF PAGES 5
	16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used for announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements.***

MASTER COPY: PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

MEMORANDUM OF TRANSMITTAL

U.S. Army Research Office
ATTN: AMSRL-RO-BI (TR)
P.O. Box 12211
Research Triangle Park, NC 27709-2211

☐ Reprint (Orig + 2 copies)

☐ Manuscript (1 copy)

☐ Technical Report (Orig + 2 copies)

XX☐ Final Progress Report (Orig + 2 copies)

☐ Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER: DAAH04-1-0391

REPORT TITLE: Exploiting Decision and Policy Models More Effectively through Interlligent Search, Data Mining, and Visualization

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

[signature]

FINAL REPORT DOCUMENTATION PAGE [1]

This report pertains to "Exploiting Decision and Policy Models More Effectively through Intelligent Search, Data Mining, and Visualization," funding numbers DAAH04-1-0391.

1 List of Manuscripts

- "On Heuristic Mapping of Decision Surfaces for Post-Evaluation Analysis." Completed September 30, 1996; appeared in *Proceedings Hawaii International Conference on System Sciences* (a refereed conference), IEEE Press (published electronically on CDROM), January 1997.
- "On Relevance and Two Aspects of the Organizational Memory Problem," Garett Dworman, Steven O. Kimbrough, Stephen Kirk, and Jim R. Oliver. Version: orgmems5 (December 15, 1997). technical report.
- "MOTC: An Aid to Multidimensional Hypothesis Generation," by K. Balachandran, J. Buzydlowski, G. Dworman, S.O. Kimbrough, T. Shafer, and W. Vachula, in Nunamaker and Sprague, eds., *Proc. Thirty-First Hawai'i Int'l. Conf. on System Sciences,* IEEE Computer Press, 1998.
- "MOTC: An Interactive Aid for Multidimensional Hypothesis Generation," Balachandran, K., Buzydlowski, J., Dworman, G., Kimbrough, S., Shafer, T. and Vachula, W. *Journal of Management Information Systems,* **16** no. 1 (Summer 1999), pp. 17–36. File: motcjmis5.pdf.
- *Pattern-Oriented Access to document Collections,* University of Pennsylvania, Ph.D. thesis, 1999, by Garett Dworman.

2 Scientific Personnel

- Bill Branley (Ph.D. student)
- Russel Fradin (undergraduate student)
- Steven O. Kimbrough (PI, University of Pennsylvania Professor)

[1] File: aro-final.tex

- Tate Shafer (undergraduate student)
- Garett Dworman (Ph.D. student)
- Balachandran, K., Buzydlowski, J., and Vachula, W. (Ph.D. students)

3 Inventions

None during this period. Subsequently (winter 1999-2000) two inventions, disclosed to the University of Pennsylvania Center for Technology Transfer, pertaining to fast creation of indexes and to incremental updating of index files.

4 Scientific Progress and Accomplishments

4.1 Background

Candle-Lighting Analysis (CLA) is the name for a series of techniques for automated support for intelligent examination of formal models, particularly management science models. Further, these same techniques are valuable for validation of such models. Such examinations are well-suited for model validation and for post-evaluation analysis, and can be particularly useful for providing decision makers with insight into the system being modeled. Candle-lighting analysis holds, and has demonstrated, much promise for reducing users' time-to-insight as they work with decision and policy models.

The primary alternative to candle-lighting analysis is manually-directed *what-if* questioning. Such what-if analysis is a cumbersome process, appropriate mainly for small-scale models; it is a process forever at risk of being undermined by human biases and cognitive limitations. While valuable, manually-directed *what-if* questioning is widely recognized to have serious limitations, yet until the development of *candle-lighting analysis* concepts and the CLAP-NT software system, there existed no general, principled, analyst-friendly way to *systematically* and *thoroughly* perform *what-if* questioning.

Candle-lighting analysis concepts hold the promise of greatly enhancing the value and effective use of management science models in the U.S. Army, and DoD generally. Candle-lighting analysis modules, operating in a real-time, distributed environment, have an essential rôle to play in supporting decision making in a wide variety of functional areas—including logistics, personnel, and operations.

Essential to the concept of candle-lighting analysis (CLA) is the creation of multiple databases containing information gleaned from particular formal or mathematical models (including, but not limited to, management science models) using intelligent heuristics (including, but not limited to, genetic algorithms). Very roughly, CLA treats models as (creating or entailing) complex, hyperdimensional structures in decision space. The resulting (hyperdimensional) surfaces need to be explored intelligently and mapped for human decision makers. These explorations yield large data sets, which are stored in a database environment. The databases are then *mined* in order to give decision makers insights in the problem at hand, and to suggest promising policy options to them.

In short, CLA works with *decision models,* which decision makers use to estimate the consequences of making decisions. Typically, these models apply for problems of optimal allocation of resources. In such situations, decision makers need—and require—more than a simple recommendation regarding the "best" alternative course of action. Faced with a series of possible courses of action, we may characterize the face of decision as a multidimensional hypersurface representing the quality of the outcomes for the various possible courses of action. Ideally, decision makers (or the staff supporting them) require:

1. A comprehensive map of the (very complex) decision surface presented to them by the choices at hand.

2. Tools and concepts for interpreting and visualizing this decision surface, in a way that is conducive to effective decision making.

Part 1 of this project was aimed at developing concepts, principles, and usable software for meeting these requirements. Part 2 was aimed at developing effective information retrieval concepts, especially as they pertain to candle-lighting data. Here the project took a surprising and very successful direction. Broadly we developed the mathematics of the DCB representation and we extended it to use in pattern-oriented retrieval. These accomplishments are documented in the aforementioned papers and technical reports.

4.2 Summary of Accomplishments

1. Redesign and implementation of the CLAP-NT (Candle-Lighting Analysis Program in NT).

   The work under this contract continues work previously supported by the Army AI Center. Our previous work produced various key concepts for candle-lighting analysis and a demonstration prototype. That prototype yielded important proof-of-concept demonstrations, but time and resource contraints prevented it from being implemented in a fully general fashion. (For example, it was limited to handling only about 30 models.)

   During this phase of the contract we performed considerable redesign and implementation of the CLAP-NT system. While much of the "look and feel" remained constant, the internal design was changed radically. As a consequence of this, for example, the database design is now quite general and can handle an indefinitely large number of models. Also, models are now implemented as objects. This considerably simplifies integrating new models into the system and allows for use of distributed models, models residing on other computers and accessed over the Internet, for example.

2. Location, evaluation, and implementation of a series of new models in the CLAP-NT system.

   The most important of these models are described in our paper (see the paper cited in §1). We think that the range (e.g., including integer programming models, knapsack models, goal programming models) and scope of these models (e.g., applications in logistics, federal funding allocations, and construction design) amply demonstrate the usefulness of CLA concepts and the CLAP-NT system.

3. Implementation of additional heuristic search methods.

   At the start of this contract we had three heuristic search methods for generating the map of the decision surface (see requirement (1) in §4.1): search by genetic algorithm, random search (for benchmarking purposes), and greedy hill-climbing. During this period we reimplemented these methods as objects (see item (1), above) and we added new methods (as objects): simulated annealing and branch-and-bound.

4. Performance of computational experiments.

   (These are partly reported in our paper, see §1.)

5. Perform research and analysis on (a) what is known about the structure of decision surfaces, and (b) general statistical methods for KDD in the context of CLA.

5 Technology Transfer

None during this period, but Practical Reasoning, Inc., of Bala Cynwyd, PA, a recent startup, will be using these techniques extensively.