Identifying spatially defined model representation templates in large data sets using MOTORS

Steven Woods[1]

[1]Defense Research Establishment Valcartier, Courcelette, Quebec, Canada, G1R 2L5

Abstract

A template is defined as an expected configuration of (possibly loosely) spatially related objects. Each object in a template can have constraints which describe it. In an object-rich environment it can be difficult to recognize these templates as a result of clutter, or as a result of the complex interactions of many object constraints. The problem becomes even more difficult if we consider many templates, or a dynamic environment. Our earlier work in this domain focused on providing a set of tools which would enable a human to orient a template defined in terms of a set of expectations upon a spatial workspace and then to be notified if a certain degree of correspondence was detected between the template components and the environment. A method requiring less human interaction would be desirable, particularly in cases where many templates could exist or co-exist. This paper is intended as an evaluation of the applicability of known constraint satisfaction and model pattern recognition methods to the solution of the template recognition problem. The system MOTOR (MOdel Template Object Recognition system) presented in this paper has been constructed by merging constraint propagation techniques with search strategies, and controlled via an interactive preference structure based on abstraction.

Contents

List of Tables

List of Figures

Chapter 1

Executive Summary

1.1 Project Identification

During January in conjunction with the members of my group and group leader Marlene Gauvin, I identified "Recognition of Spatial Deployment Templates from Temporal Data" as the preliminary topic for my research during my time at DREV. The following is an extract of this work proposal of February 3, 1993.

> "A deployment template represents information about a particular, pre-determined, set of relative unit positions. Each element of a template has a particular (ideal) constitution and relationship (spatial or otherwise) with the other units of the template. One method intelligence analysts use in the effort to construct a better "hypothesis" of an enemy situation is recognition of these common patterns or templates. The temporal nature of the messages describing sightings further complicates the analyst's task. I propose to investigate the possibility of implementing a decision aid subsystem that would be capable of assisting an intelligence analyst recognize the existence of these templates by examining the past positions, past movements, and constitution of units described in intelligence messages. Of particular interest in my research will be the movement of sets of units over time and in determining whether or not any pattern can be detected indicating evidence for or against a particular template or set of templates.

1.2 Work Summary

Initially, I researched existing DREV documents and made use of DREV personnel with knowledge of templates in order to form a better model in my mind of what the "essential" elements of a doctrinal template were in the view of an intelligence analyst. The result of this research is captured in the simplified model of templates I make use of throughout this document. Essentially I adopted a constraint based system where individual enemy elements are composed into a template of a particular type based on inter-element relationships of various types.

As the next major step of my preliminary work, I undertook a literature survey across several broad areas searching for models or discussions of systems of a similar nature. In particular I was interested in determining what approach to use in finding a practical, implementable system which could deliver the minimum requirements of serving as a useful tool for the analyst in the pursuit of identifying templates. I focused primarily on military related materials and conference proceedings for specific mention of this problem type, on prominent artificial intelligence publications for discussion of practical applications of modelling and reasoning with constraints in this fashion, and as well attempted to broaden my outlook by examining several alternative solution approaches such as massively parallel model recognition which demonstrates similarity in a restricted manner. Each of these tasks is discussed later in this paper, however, I was unable to find published discussion of similar work or implementations. Published documentation on constraint satisfaction implementation lead me to believe that a combined search and constraint propagation approach could solve the problem if a model could be found which appropriately handled analyst interaction in the search process.

At this time I established several major goals that I believed would be critical in the development of a working subsystem. These goals are listed below, along with a comment summarizing the state of the work toward these each goal.

1. Restrict the scope of the entire problem statement even further so as to manageable handle a static model first, before attempting to address the more complex temporal aspects. To this end, identify the critical assumptions and processes that must be made to time-standardize the message data.

   I have assumed throughout my work a "snapshot" model of a particular point in time. Templates as conceived by the military currently represent a picture of unit deployment at one time point. Identification of these unit patterns in message data is itself a difficult problem, and a strategy towards this static identification needs to be understood more fully before a temporal approach could be attempted. The intelligence messages containing enemy unit or equipment sighting information are of varying accuracy and completeness with respect to perceived structure, location and time. Messages describe a sighting at some point in time T1, are sent for collation at some other time T2, and arrive later at T3. In attempting to organize hundreds of these messages at some still later time T4 into a coherent picture of the world at T4, the analyst faces a daunting task. Any perception of the world at T1 decays over time is such a dynamic situation, and as such, sightings lose their information value over time. As a result, analysts commonly attempt to form a picture for some past time TP somewhere before T4 and after T1 based on all reports available at T4. This single-time model created of the past will be considered to be the input to our system, essentially a blurry snapshot of a moment in the past.

2. Identify major limiting assumptions about the nature of the incoming message traffic itself so as to better understand the complexity of the problem. In particular, the spatial scope of messages needs to be better characterized in terms of the analyst's own perceptions, and "typical" volume and distribution values need to be determined.

   Since one major aspect of the constraint satisfaction problem of template recognition is spatial relationships of sighted units, it is necessary to characterize what locational information is considered implicit in a sighting message for some particular unit with type, size and other attributes. During May of 1993 I was assigned a summer student, and after investing much time explaining the problem itself to him, I instructed him to visit Valcartier directly and discuss the problem with an intelligence officer. Some progress was made in characterizing the scope of certain messages, and these results are captured in this paper. However, more work is required in this area, in particular in characterizing the way in which the analysts themselves perceive and use the information they receive.

   Volume and distribution of information is very important and in fact represents "noise" that confuses the situation and makes the discovery of template instances difficult. While both myself and my summer student have spent time examining the sample situation model present in the TIF demo, I believe that further examination of real situation message data is necessary to better characterize "typical" message distribution in terms of the location and type of sightings. Throughout the course of our experiments we have assumed largely random distributions of message sightings over the field of focus.

3. Model the problem solving subsystem as a stand-alone, non-interactive tool which could identify the templates in "typical" data. From this experiment, determine the deficiencies and highlight the value of the analyst experience in solving these problems with the intent of capturing the analyst's function in the complete, interactive model.

   We have modelled stand-alone doctrinal deployment template recognition as a *constraint satisfaction* problem (CSP). In this paper a toolset is described called MOTORS (MOdel Template Object Recognition System)which provides for the solution of CSPs utilizing a configurable hybrid search and constraint propagation strategy. Some very promising results have been obtained by experimenting with various heuristics and search strategies, and in fact the nature of spatial templates as modelled has allowed for discovery of all instances of a particular example template in very noisy data in only a few seconds. Specific results are detailed later in this paper.

4. From the stand-alone model, identify problem-dependent heuristics that may simplify the problem solution. In addition, attempt to identify problem characteristics that will suggest what type or mode of hybrid is appropriate. In all constraint satisfaction approaches, a trade-off must be made between search and constraint propagation

techniques during search, and the problem properties which might indicate the balance need to be examined. Earlier work in simplistic domains has suggested several possible approaches, but no work has yet been published with so complex a problem as template recognition.

Throughout this paper, work done by myself and by my summer student is presented which outlines the various approaches and tradeoffs. Implemented versions of several possible algorithms are available and may be easily integrated into the MOTORS stand-alone prototype. Additionally, some experiments have been carried out with spatial heuristics in an effort to restrict the size of the search space based on the concept of template-instance locality within the body of message data.

5. Develop a model accommodating the application of the analyst's experience and expertise together with the constraint reasoning approach of the stand-alone tool. Pay particular attention to forming a model that is capable of flexibly shifting between interaction and stand-alone so as to better suit a particular analyst's changing needs and available time. A best-so-far approach is required that will give some solution "anytime" if the analyst needs an answer immediately.

   In [69], a model has been developed which will allow for the user interaction in the search process for templates. This model is based upon the conception of expressing templates in terms of "key" or "critical" elements and additional "supporting" evidence which lends credence to the acceptance of a set of enemy sightings as a template instance. The model accounts for making explicit use of the new problem specification and change of representation information, is capable of providing best-so-far answers where unlimited time is unavailable, and has a configurable amount of user interaction.

6. Implement the full model as a prototype independent of the Tactical Information Fusion (TIF) system. In this way, the TIF experts can better identify exact application technology or interface requirements and take advantage of TIF specific data structures and implementation quirks of which I am not aware and cannot expect to master in so short a time period.

While the stand-alone MOTORS system is written in LISP and is fully operational, the interactive model prototype has not yet been implemented.

1.3 Future Work

Much of the implementation and experimentation work I have undertaken has been towards developing interesting constraint satisfaction algorithms which function as hybrids of search and constraint propagation. Examination of how to best dynamically carry out this search is not well understood, even for simpler problems, however, our implementation has suggested that model template recognition is restricted enough to allow for reasonable time solutions.

The model presented for interactive template recognition is consistent with other ongoing efforts in abstract constraint satisfaction work aimed at developing scheduling and planning systems elsewhere [71, 72]. I believe that the model presented in [69] can be utilized to create an interesting, useful and functional decision aid tool for application in an intelligence collation system such as TIF, and would provide a flexible testbed for new heuristics, propagation methods and search strategies.

There are many related interesting research issues that deserve attention. Problem-specific issues include:

- Message data characterization, particularly in terms of spatial distribution, and data point "resolution" indicating the area of possible location of a sighted element.
- Message volume characterization, and the time-range of data considered in template recognition must be determined more precisely.
- Integration of additional information such as terrain can significantly reduce the complexity of locating template instances by restricting even further the possible locations for enemy units.
- Extension of the model for template recognition to include temporal message values and extended fused data (such as tracks and merges).
- Make use of partial template matches supported by the model to identify probable or possible locations of unsighted enemy units based on acceptance of proposed templates. This application was discussed at some length at the TIF demo in Ottawa and was universally seen by the participants as a valuable additional tool. Further work needs to be undertaken to better characterize how the analysts themselves view these hypotheses in terms of likelihood and acceptance.

Additionally, constraint satisfaction provides a rich environment for future domain independent work which can significantly improve solution time for template recognition as well as other problems such as scheduling and planning:

- Template recognition is a very spatial problem, and more though must be given to how to best take advantage of this spatial nature in terms of limiting the scope of search for certain templates. In fact, all constraint satisfaction problems of a spatial nature can benefit from new approaches to propagation using spatial heuristics.
- Further expand upon the hybrid model so as to better understand where to make use of propagation effort and where to take advantage of search of certain types. "Smarter" hybrid search strategies such as that shown in Chapter 5 can make significant improvements in overall performance.

- Model constraint satisfaction so as to take advantage of parallel implementations of some aspects of constraint propagation where applicable.

- The tradeoffs between finding "one", "some" and "all" solutions are many, particularly in an interactive system. While we have adopted the "all" approach with an eye towards providing them each as quickly as possibly (one at a time), other slightly different requirements would suggest a different approach altogether. For instance, a scheduling problem requires only one solution of a certain "goodness". Similar work can be applied to template recognition in attempting to find a template instance that has a certain degree of "confidence".

- Expression of constraint satisfaction problems in terms of abstraction hierarchies with explicit representation change information as presented in our model suggests further work in profitably applying these representation changes during search. It is well known that problem representation has a direct and dramatic affect on search performance, and this new representation is only understood at a very high level currently. Other representations may exist for template recognition and other constraint satisfaction problems that suggest modified or entirely new solution approaches.

1.4 Looking Ahead

Commencing September, 1993 I will be commencing my Ph.D. work under the supervision of Dr. Qiang Yang at the University of Waterloo as a member of the Logic Programming and Artificial Intelligence Group. The exact nature and topic of my Ph.D. thesis is not completely determined at this time, however, I have agreed in principle with Dr. Yang that the problem of military doctrinal template recognition is sufficiently interesting and complex to base a thesis upon. I intend to focus on some aspects of these spatial and temporal constraint satisfaction problems and various solution possibilities. There are many domains for these problems, and the template recognition one is a rich and interesting one with good clear application possibilities.