The Utilisation of Hyper-Heuristics in Parallel Code Generation

B. McCollum, P.J.P. McMullan, P. Milligan and P.H. Corr
School of Computer Science
Queen's University Belfast
Belfast BT7 1NN

B.McCollum@qub.ac.uk
Tel +44 2890274622
Fax +44 028 90683890

The generation of code for parallel processing systems has been, and remains, an important area of research. In particular, the efficient generation of parallel code for multi-processor environments has attracted significant research effort [BAN92, AYG98].

Attempts to automate the processes associated with eliminating, or at least reducing, data dependencies, and the problems associated with data distribution within a multi-processor environment have not been successful. To date, attempts to address these, and other problems, have always resulted in significant input from users. Because of the importance of these topics the automation of these processes remains a practically important research area [McC99a].

In recent years heuristic/meta-heuristic approaches have been used to direct the parallelisation process [MAN94, COR01]. These approaches have been used either in constructing semi-automatic parallelisation environments [McM97] or in providing access to knowledge that enables the user to incrementally drive the parallelisation process forward [McC98a].

The KATT parallelisation environment, developed at Queen's University, is a framework which allows the user varying amounts of involvement with the parallelisation process depending on level of experience [McC99a, b]. At one extreme, a novice may allow a particular piece of code to be automatically parallelised for a target architecture whereas an experienced user may follow the process in a step-by-step manner, making informed decisions at each stage and thereby directing the production of the resultant code.

The work contained in this paper describes the incorporation of expert system and neural network technologies into the KATT framework. . The use of these technologies supports the elicitation, capture and subsequent incorporation of relevant experienced user knowledge into the parallelisation process. Such knowledge utilisation allows the entire parallelisation process to be efficiently automated. The combination of these technologies at different times allows all aspects of the parallelisation process to be controlled and managed without the need for the user intervention. It is argued that this combination of heuristic based approaches in a common framework represents a hyper-heuristic [BUR99b] approach to parallel code generation.

Heuristics (derived from the Greek for discovery) are criteria, methods, or principles for deciding among several alternative courses of action. Computerised techniques, which attempt to incorporate such "rules of thumb", are termed heuristic techniques. The hyper-heuristics approach differs from meta-heuristic approaches, which refer to heuristics, which control simpler heuristics for a narrow range of problems, rather than of choosing between a range of heuristic approaches to solve a wide range of problems. This is a term coined by the ASAP group [PET98, BUR99b] to describe the idea of using a number of different heuristics together, so that the actual heuristic applied may differ at each decision point. In essence, hyper-heuristics are heuristics to choose heuristics. The hyper-heuristics provides a method for managing these simple heuristics, which are poor at producing effective solutions when considered simply, to yield a good solution overall. An example of the use of this concept is

given in [FRC94], in which a genetic algorithm (GA) is applied to open-shop scheduling problems, to evolve the choice of heuristic to apply whenever a task is to be added to the schedule under construction.

The hyper-heuristic approach represents a novel and exciting area of research whose utilisation offers a solution to a long standing problem. The KATT environment, as presented, provides an overview of the development of an automated/semi-automated parallelisation environment and details the approaches at Queen's University to incorporate several heuristic based techniques into a common hyper-heuristic framework.

[AYG98] Ayguade, E., Garcia, J., Kremer, U., "Tools and techniques for automatic data layout", Parallel Computing 24 (1998) 557-578.

[BAN92] Banerjee U, "Loop Transformations for Restructuring Compilers", Macmillan College Publishing Company, 1992.

[BUR99b] E.K. Burke, A.J. Smith, ``Hybrid Evolutionary Techniques for the Maintenance Scheduling Problem", accepted for publication in the IEEE Power Engineering Society Transactions**.**

[COR01] Corr PH, Milligan P and Purnell V,. "A Neural Network Based Tool for Semi-automatic Code Transformation." In VECPAR'2000. Selected Papers and Invited Talks from the 4th International Conference on Vector and Parallel Processing, Springer, Lecture Notes in Computer Science 1981, 2001, pp 142-153, ISBN 3-540-41999-3.

[FRC94] H-L Fang, P.M.Ross and D.Corne, ``A Promising Hybrid GA/Heuristic Approach for Open-Shop Scheduling Problems", in {\em Proceedings of ECAI 94: 11th European Conference on Artificial Intelligence}, A. Cohn (ed), pages 590--594, John Wiley and Sons Ltd, 1994.

[MAN94] Mansour, N., Fox, G., "Allocating data to distributed memory multiprocessors by genetic algorithms", Concurrency: Practice and Experience, Vol. 6(6), 485-504(September 1994).

[McC99a] McCollum B.G.C., Milligan P. and Corr P.H., Knowledge Exploitation for Improved parallel orientated software design, Parallel Computing Technologies, Springer, Lecture Notes in Computer Science 1662, 1999, pp 84-86, ISBN 3-540-66363-0.

[McC99b] McCollum B.G.C., Milligan P. and Corr P.H., "The Structure and Exploitation of Available Knowledge for Enhanced Data Distribution in a Parallel Environment" ,IMACS/IEEE CSCC'99, Athens,1999, pp3301-3307, ISBN 960-8052-00-9

[McC99c] McCollum B.G.C., Milligan P. and Corr P.H., "The Structure and Exploitation of Available Knowledge for Enhanced Data Distribution in a Parallel Environment" ,Software and Hardware Engineering for the 21st Century, Ed. N. E. Mastorakis, World Scientific and Engineering Society Press, 1999, pp139-145, ISBN 960-8052-06-8

[McC98a] McCollum B.G.C., Milligan P. and Corr P.H, "The Improvement of a Software Design Methodology by the Elicitation of Knowledge from Code", IEEE Computer Society Press, 1998. pp913-918, ISBN 0-8186-8646-4

[McM97] McMullan P. J. P, Milligan P, Sage P. P. and Corr P. H.. A Knowledge Based Approach to the Parallelisation, Generation and Evaluation of Code for Execution on Parallel Architectures. IEEE Computer Society Press, ISBN 0-8186-7703-1, pp 58 - 63, 1997.

[PET98] S.Petrovic, "Case-Based Reasoning in the DSS for Multiattribute Analysis", Journal of Decision Systems Vol. 7, pp 99-119, Hermes, 1998.