Heuristic Search for Large Problems with Real Costs

Matthew Hatem

Department of Computer Science
University of New Hampshire
Durham, NH 03824 USA
mhatem@cs.unh.edu

Abstract

Heuristic search is a fundamental technique for solving problems in artificial intelligence. However, many heuristic search algorithms, such as A* are limited by the amount of main memory available. External memory search overcomes the memory limitation of A* by taking advantage of cheap secondary storage, such as disk. Previous work in this area assumes that edge costs fall within a narrow range of integer values and relies on uniformed search order. The goal of this dissertation research is to develop novel techniques that enable heuristic search algorithms to solve problems with real values using a best-first search order while exploiting external memory and multiple processors. This work will be organized into four components. The first component will discuss external memory search and present a novel technique for incorporating real-valued edge costs. The second component will present a novel algorithm for solving problems with large branching factors with application to the challenging problem of Multiple Sequence Alignment (MSA). The third component will cover bounded suboptimal external search for practical MSA applications. The final component of this research will be the development of a novel distributed search framework; allowing parallel and external memory heuristic search algorithms to run cooperatively on a commodity computing cluster. Together these four components will enable heuristic search to scale to large problems in practical settings while exploiting modern hardware.

External Memory Search with Real Costs

Best-first graph search algorithms such as A* (Hart, Nilsson, and Raphael 1968) are widely used for solving problems in articial intelligence. A* maintains an *open list*, containing nodes that have been generated but not yet expanded, and a *closed list*, containing all generated nodes, in order to prevent duplicated search effort when the same state is generated via multiple paths. As the size of a problem increases, however, the memory required to maintain the open and closed lists makes A* impractical.

External memory search algorithms take advantage of cheap secondary storage, such as hard disks, to solve much larger problems than algorithms that only use main memory. A naïve implementation of A* with external storage has poor performance because it relies on random access and

disks have high latency. Instead, great care must be taken to access memory sequentially to minimize seeks and exploit caching.

Previous approaches achieve sequential performance in part by dividing the search into layers. For heuristic search, a layer refers to nodes with the same lower bound on solution cost f. Many real-world problems have real-valued costs, giving rise to a large number of f layers with few nodes in each, substantially eroding performance.

This first component of the dissertation extends previous work with a new technique for external memory search that performs well on graphs with real-valued edges. Recently we presented a new general-purpose algorithm, Parallel External search with Dynamic A* Layering (PEDAL) (Hatem, Burns, and Ruml 2011), that uses external memory and parallelism to perform a best-first heuristic search capable of solving large problems with real costs. We showed theoretically that PEDAL is I/O efficient and empirically that it gives far superior performance on problems with real costs. This work was published in the Proceedings of the Twenty-Fifth Conference on Artificial Intelligence (AAAI-2011).

External Search with Large Branching Factors

Next we consider the challenge of problems with large branching factors. One real-world application of heuristic search with practical relevance (Korf 2012) is Multiple Sequence Alignment (MSA). MSA can be formulated as a shortest path problem and solved using heuristic search, but the large branching factor makes algorithms such as A* impractical for all but the smallest problems.

Partial Expansion A* (PEA*) reduces the space complexity of A* by generating only the most promising successor nodes. However, even PEA* exhausts available memory on many problems. Another alternative is Iterative Deepening Dynamic Programming, which stores only the nodes along the search frontier but uses an uninformed search order. However, it too cannot scale to the largest problems.

The main contribution of this second component of research is a new general-purpose algorithm, Parallel External PEA* (PE2A*) (Hatem and Ruml, under review), that combines PEA* with Delayed Duplicate Detection (DDD) (Korf 2008) to take advantage of external memory and multiple processors to solve large MSA problems. In our experiments, PE2A* is the first algorithm capable of solving

the entire Reference Set 1 of the standard BAliBASE benchmark using a biologically accurate cost function. This work suggests that external best-first search can effectively use heuristic information to surpass methods that rely on uninformed search orders.

Bounded Suboptimal External Search

The third component investigates techniques that forgo optimal solutions in order to solve problems in less time. It is commonly appreciated that solving search problems optimally can overrun RAM. One alternative is to use a bounded suboptimal search algorithm such as Weighted-A* (wA*) (Pohl 1970). The wA* algorithm allows one to trade solution quality for computational effort. One attractive property of wA* is that the solutions generated are guaranteed to be within a pre-specified factor w of the optimal solution. Bounded suboptimal search does not completely eliminate the RAM constraint and it is possible for wA* to overrun RAM for large problems.

While external memory search algorithms can solve large problems such as MSA optimally, they may require days or weeks of running time. One obvious solution is to combine bounded suboptimal search with external memory search. In our experience, life scientists are willing to tolerate suboptimal solutions in exchange for a significant reduction in running time. The main contribution of this third component of the research is an investigation in bounded suboptimal external memory heuristic search with application to the MSA domain (in progress).

One side effect of forgoing optimality in MSA is reduction in the biological plausibility of the alignments produced. We can increase plausibility by using biologically relevant cost functions. Historically the MSA domain has consisted of edges with rounded integer costs but recently it has been shown (Edgar 2009) that real-values can produce alignments with a higher degree of biological plausibility. Another aspect of this work is to combine the real-valued handling of PEDAL with PE2A* to solve challenging MSA problems using a biologically accurate real-valued cost function (in progress).

Distributed External Search

Finally, we investigate techniques for exploiting multiple processors. To take advantage of the increasing number of cores on the modern CPU and the computing power of distributed systems it is imperative that we develop parallel versions of heuristic search algorithms. Niewiadomski, Amaral, and Holte (2006) present Parallel Frontier A* with Delayed Duplicate Detection (PFA*-DDD), a distributed search algorithm capable of solving challenging MSA problems. PFA*-DDD combines Frontier A* search (Korf et al. 2005) with DDD but does not exploit multiple CPU cores or external storage.

An ideal distributed search algorithm will take advantage of all of CPU cores available at each node as well as all primary and secondary memory. We believe PEDAL and PE2A* are strong candidates for this. The main contribution of this final component of the research is to extend our research in external memory heuristic search to develop a general purpose distributed search framework that can take advantage of all computing resources available at each node in the cluster. We are considering extensions of the Hash Distributed A* (HDA*) (Kishimoto, Fukunaga, and Botea 2009) and Structured Duplicate Detection (SDD) (Zhou and Hansen 2004) techniques as well as other novel approaches for localizing memory references and workload distribution.

Summary

Though there has been a good deal of research in external memory and distributed heuristic search, previous work has emphasized integer cost domains and do not fully exploit the computing and storage resources available. MSA is a compelling real-world domain for heuristic search. The large branching factor makes in-memory algorithms impractical and applying real costs to this domain is troublesome for previous work. This dissertation extends previous work to a general purpose framework that takes advantage of all computing resources available to solve large problems in practical settings.

References

Edgar, R. C. 2009. Optimizing substitution matrix choice and gap parameters for sequence alignment. *Bioinformatics*.

Hart, P. E.; Nilsson, N. J.; and Raphael, B. 1968. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions of Systems Science and Cybernetics* SSC-4(2):100–107.

Hatem, M., and Ruml, W. 2013. External memory best-first search for multiple sequence alignment. *Submitted to Proceedings of the Twenty-Seventh Conference on Artificial Intelligence (AAAI-2013)*.

Hatem, M.; Burns, E.; and Ruml, W. 2011. Heuristic search for large problems with real costs. In *Proceedings of the Twenty-Fifth Conference on Artificial Intelligence (AAAI-2011)*.

Kishimoto, A.; Fukunaga, A.; and Botea, A. 2009. Scalable, parallel best-first search for optimal sequential planning. In *Proceedings of the Nineteenth International Conference on Automated Planning and Scheduling (ICAPS-09)*.

Korf, R. E.; Zhang, W.; Thayer, I.; and Hohwald, H. 2005. Frontier search. *Journal of the ACM* 52(5):715–748.

Korf, R. E. 2008. Linear-time disk-based implicit graph search. *Journal of the ACM* 55(6).

Korf, R. 2012. Research challenges in combinatorial search. In *Proceedings of AAAI-2012*.

Niewiadomski, R.; Amaral, J. N.; and Holte, R. C. 2006. Sequential and parallel algorithms for frontier A* with delayed duplicate detection. In *Proceedings of the 21st national conference on Artificial intelligence (AAAI-06)*, 1039–1044. AAAI Press.

Pohl, I. 1970. Heuristic search viewed as path finding in a graph. *Artificial Intelligence* 1:193–204.

Zhou, R., and Hansen, E. A. 2004. Structured duplicate detection in external-memory graph search. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)*.