

Master's Thesis: Which Node to Expand?

Andrew Mitchell

Department of Computer Science
University of New Hampshire
ajx256@wildcats.unh.edu

Abstract

The purpose of this research is to produce a real-time heuristic search algorithm prototype that reasons carefully about which node is best to expand next. Traditionally, nodes are selected for expansion so that the f -value of the selected node is minimized.

However, in real-time search the goal is not to find the minimal path from the start to the goal, it is to find the best path possible, towards the goal, within an allotted time or number of expansions per move. Therefore, in real-time search the node on the frontier with the lowest f -value is not necessarily the best node to expand. The best node to expand in this instance is the node that increases confidence the most about the best action that the agent can immediately take.

I will implement a prototype algorithm that reasons more carefully about which node is best to expand next, and test it on some classic heuristic search benchmarks, as well as against traditional real-time search algorithms.

1. INTRODUCTION

Traditionally, exploration of a search space in heuristic search is done by expanding next the node on the frontier with the lowest f -value. In situations where an algorithm has time to fully explore the search space and find this path before taking an action, this is a provably optimal exploration method, and if used in conjunction with an admissible heuristic, will yield the minimal cost path. [3] [2]

However, this is not the case in every situation. For example, in real-time heuristic search, the agent's exploration is limited to an allotted time or number of expansions before it has to make a decision about which immediate action to take. In this situation the agent may not be able to fully explore the search space and find a minimal cost path before the time or expansion limit is reached. When time and expansions are limited, the question of which node to expand becomes more complicated than simply expanding the frontier node with the lowest f -value.

It might be expected that the exploration method for real-time heuristic search should differ from that of traditional heuristic search because the goals are inherently different. The traditional goal is to expand nodes on the frontier until a complete path from the start to the goal has been found. The real-time goal is to expand nodes to gather information about the environment so that the agent may make an informed decision about which is the best immediate action to take towards the goal from its current state.

In real-time heuristic search, an agent may be presented with several actions that it could take from its current state, which will be referred to as top-level actions. Beneath every top-level action is a subtree that could potentially lead to the goal state. The role of the agent is to explore these subtrees, by expanding nodes within them, and to then decide on which top-level action is the best to take based on the information that it has currently gathered from the subtrees. As exploration is limited, it is important to decide which nodes should be explored, and which should be ignored.

This thesis is based on theory provided by Wheeler Ruml, who has proposed that in the case of real-time heuristic search, the best node to expand is the node which increases the agent's confidence the most about which top-level action to take. Confidence of a top-level action is defined as the probability that no other top-level action will lead to a better path to the goal. Therefore, instead of selecting the frontier node with the lowest f -value, the agent should select for expansion the node which maximizes the expected change in confidence in the best current top-level action. This will ensure that only nodes that yield potentially useful information about the search space are expanded.

2. Related Work

The concepts explored in this thesis can be defined as metareasoning for search. Previous work has been done on metareasoning for real-time heuristic search, and some of the first concepts of it were discussed by Russell and Wefald in their book, *Do the Right Thing* [9].

Mutchler wrote a paper in 1986 [6] that presented a probabilistic model of where limited resources should be spent in a search tree. In the context of this research, node expansions are the limited resource, and we are trying to focus them on only those nodes which result in the maximum change of our expected confidence.

Inspired by the ideas that Mutchler explored in his paper, Pemberton described a decision making strategy, k-best [8], which would approximate the optimal action by taking the top-level action that minimizes expected path cost. This is as opposed to naively moving towards the node with minimum cost on the frontier, as doing so ignores the distribution of potential leaf costs, and can lead to sub-optimal behavior.

More recently, the Mo'RTS algorithm [7] allows agents to reason about when they should take identity actions to allow for additional exploration and to reason about how many actions they should commit to taking during a movement step. Similarly, Slo'RTS [1] is an algorithm that estimates if the expected gain in future decision-making war-

rants taking a top-level action with a longer than optimal duration to allow for additional search time.

Other algorithms such as Lazy A* [10] take into account that the generation of heuristic values can be costly. However, guiding search through the use of multiple different heuristics can often yield better solutions than merely relying on a single heuristic. To avoid large overhead at node generation time, Lazy A* only evaluates a single heuristic function, and evaluates the others each time a node reappears at the front of the open list.

This research was done in loose collaboration with Joerg Hoffmann and Fabian Spaniol. While they are also exploring the problem of which nodes to expand, they are looking into using this theory in the Fast Downward planner.

3. Research Plan

This thesis will include an implementation of the proposed algorithm, currently called Nancy. Once this algorithm is implemented, it will be compared against traditional real-time heuristic search algorithms, such as LSS-LRTA* [5], \hat{f} [4], and an implementation of k-best [8]. For reference, these will also be compared against A* [3] and beam search [11]. All of these algorithms will be compared based on their solution path cost.

These algorithms will be compared on a domain called tree world [8], which is a tree with branching factor b and depth d , where every node at depth d is a goal node and the root is the start node. Edges of the tree represent action costs between nodes, and will be a random number from a uniform distribution drawn from the range $[0,1]$. The algorithms will be compared on this domain with various depths and various lookaheads. The goal is to determine if this new method of node exploration yields better solution costs or goal achievement time. They will also be compared on classic heuristic search benchmarks, such as the sliding tile puzzle and racetrack.

For the time being, Nancy will be implemented with a decision making strategy that utilizes minimin. However, if time permits, Nancy could be implemented using Cserna backups [1].

References

- [1] Bence Cserna, Wheeler Ruml, and Jeremy Frank. Planning Time to Think: Metareasoning for On-Line Planning with Durative Actions. In *Proceedings of the Twenty-Seventh International Conference on Automated Planning and Scheduling, ICAPS 2017.*, pages 56–60, 2017.
- [2] Rina Dechter and Judea Pearl. Generalized Best-First Search Strategies and the Optimality of A*. *J. ACM*, 32(3):505–536, 1985.
- [3] Peter E. Hart, Nils J. Nilsson, and Bertram Raphael. A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Trans. Systems Science and Cybernetics*, 4(2):100–107, 1968.
- [4] Scott Kiesel, Ethan Burns, and Wheeler Ruml. Achieving Goals Quickly Using Real-time Search: Experimental Results in Video Games. *J. Artif. Intell. Res.*, 54:123–158, 2015.
- [5] Sven Koenig and Xiaoxun Sun. Comparing real-time and incremental heuristic search for real-time situated agents. *Autonomous Agents and Multi-Agent Systems*, 18(3):313–341, 2009.
- [6] David Mutchler. Optimal Allocation of Very Limited Search Resources. In *Proceedings of the 5th National Conference on Artificial Intelligence. Volume 1: Science.*, pages 467–471, 1986.
- [7] Dylan O’Ceallaigh and Wheeler Ruml. Metareasoning in Real-Time Heuristic Search. In *Proceedings of the Eighth Annual Symposium on Combinatorial Search, SOCS 2015.*, pages 87–95, 2015.
- [8] Joseph C. Pemberton. k-Best: A New Method for Real-Time Decision Making. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, IJCAI 95*, pages 227–235, 1995.
- [9] Stuart J. Russell and Eric Wefald. *Do the right thing - studies in limited rationality*. MIT Press, 1991.
- [10] David Tolpin, Tal Beja, Solomon Eyal Shimony, Ariel Felner, and Erez Karpas. Toward Rational Deployment of Multiple Heuristics in A. In *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, pages 674–680, 2013.
- [11] Christopher Makoto Wilt, Jordan Tyler Thayer, and Wheeler Ruml. A Comparison of Greedy Search Algorithms. In *Proceedings of the Third Annual Symposium on Combinatorial Search, SOCS 2010*, 2010.