# Thinking Too Much: Pathology in Pathfinding

**Mitja Luštrek** [1] and **Vadim Bulitko** [2]

## 1 INTRODUCTION

Incomplete single-agent search methods are often better suited to real-time pathfinding tasks than complete methods (such as A*). Incomplete methods conduct a limited-depth lookahead search, i.e., expand a part of the space centered on the agent, and heuristically evaluate the distances from the frontier of the expanded space to the goal. Actions selected this way are not necessarily optimal, but it is generally believed that deeper lookahead increases the quality of decisions. However, in two-player games, where similar methods are used, it has long been known that this is not always the case [7, 1]. This phenomenon has been termed minimax *pathology*. More recently pathological behavior was discovered in single-agent search as well [3]. Some attempts to explain it have been made [5, 6], but the pathology in single-agent search is largely still not understood.

In this paper we investigate lookahead pathology in real-time pathfinding on maps from commercial computer games. First, we present an empirical study showing a degree of pathology in *over 90%* of the problems considered. Second, we give four explanations for such wide-spread pathological behavior.

## 2 THE PATHOLOGY OBSERVED

We study the problem of an agent trying to find a path from a start to a goal state in a two-dimensional grid world. The agent plans its path using the Learning Real-Time Search (LRTS) algorithm [2]. LRTS conducts a lookahead search centered on the current state and generates all the states up to $d$ moves away. It heuristically estimates the distances from the frontier states to the goal state and moves to the most promising frontier state. Upon reaching it, it conducts a new search. The initial heuristic is the shortest distance assuming an empty map. After each search, the heuristic of the current state is updated to the estimated distance through the most promising frontier state, which constitutes the process of learning.

We conducted two types of experiments: *on-policy* and *off-policy*. In the first type the agent follows a path from the start state to the goal state as directed by the LRTS algorithm. In the second type the agent appears in a (randomly selected) state and selects the first move towards the goal state. If the move does not lie on the shortest path to the goal state, it is erroneous. The *error* $e(S_d)$ is the fraction of erroneous moves taken in the set of states $S_d$ visited using lookahead depth $d$. The degree of *error pathology* in the sequence of sets $S_1, \ldots, S_{d_{\max}}$ is $k$ iff $e(S_{d+1}) > e(S_d)$ for $k$ different $d < d_{\max}$.

We generated 1,000 problems on maps from a commercial role-playing game. The lookahead depth ranged from 1 to 10 = $d_{\max}$. First we conducted the basic on-policy experiment: the agent solved the problems, we measured the degree of error pathology for each problem and counted the number of problems with each of the possible degrees. The on-policy row in Table 1 shows that *over 90%* of the problems are pathological.

**Table 1.** Pathology in the basic on and off-policy experiments.

| **Degree** | 0 | 1 | 2 | 3 | $\geq 4$ |
|---|---|---|---|---|---|
| **Pat. problems on-policy [%]** | 6.3 | 13.1 | 24.8 | 29.0 | 26.7 |
| **Pat. problems off-policy [%]** | 83.1 | 14.9 | 2.0 | 0.0 | 0.0 |

The first possible explanation of the on-policy results in Table 1 is that the maps contain a lot of states where deeper lookaheads lead to suboptimal decisions, whereas shallower ones do not. If this were the case, the basic off-policy experiment, where the pathology is measured in randomly selected states, should yield comparable pathology. However, the off-policy row in Table 1 shows much less pathology. In the rest of the paper, we will investigate the reasons for this.

## 3 EXPLANATIONS OF THE PATHOLOGY

The **first explanation** is that the LRTS algorithm's behavioral policy steers the search to pathological states. The explanation was verified by computing off-policy pathology from the error in the states visited during the basic on-policy experiment instead of randomly selected ones. The results in Table 2 do show more pathology compared to the basic off-policy experiment in Table 1 (23.2% vs. 16.9%), but they are still far from the basic on-policy experiment (23.2% vs. 93.7%).

**Table 2.** Pathology measured off-policy in the states visited on-policy.

| **Degree** | 0 | 1 | 2 | 3 | $\geq 4$ |
|---|---|---|---|---|---|
| **Patological problems [%]** | 76.8 | 13.8 | 5.7 | 2.3 | 1.4 |

The basic on-policy experiment involves learning, but no learning takes place in the basic off-policy experiment. It is harder to find the path to the goal when the lookahead depth is small. Consequently the agent backtracks more, encountering updateted states more often when the lookahead depth is large. This leads us to the **second explanation**. Smaller lookahead depths benefit more from the updates to the heuristic. This can be expected to make their decisions better than the mere depth would suggest and thus closer to larger depths. If they are closer to larger depths, cases where a deeper lookahead actually performs worse than a shallower one should be more common.

The first test of the second explanation is an on-policy experiment where the agent is directed by the LRTS algorithm that uses learning (to prevent infinite loops), but the error is measured using only the initial, non-updated heuristic. The results in Table 3 suggest that learning is indeed responsible for the pathology, because the pathology in the new experiment is markedly smaller than in the basic on-policy experiment shown in Table 1: 70.4% vs. 93.7%.

**Table 3.** Pathology on-policy with error measured without learning.

| **Degree** | 0 | 1 | 2 | 3 | $\geq 4$ |
|---|---|---|---|---|---|
| **Pathological problems [%]** | 29.6 | 20.4 | 19.3 | 18.2 | 12.5 |

[1] Jožef Stefan Institute, Department of Intelligent Systems, Jamova cesta 39, 1000 Ljubljana, Slovenia, email: mitja.lustrek@ijs.si

[2] University of Alberta, Department of Computing Science, Edmonton, Alberta T6G 2E8, Canada, email: bulitko@ualberta.ca

The second test is to measure the volume of heuristic updates, which reflects the benefit of learinng. This volume is the sum of the differences between the updated and the initial heuristics in the states generated during search. Figure 1 shows the results for the basic on-policy experiment and for the basic off-policy experiment (where no learning takes place). We see that in the on-policy experiment the volume of updates decreases with lookahead depth (unlike in the off-policy experiment), which confirms our explanation.

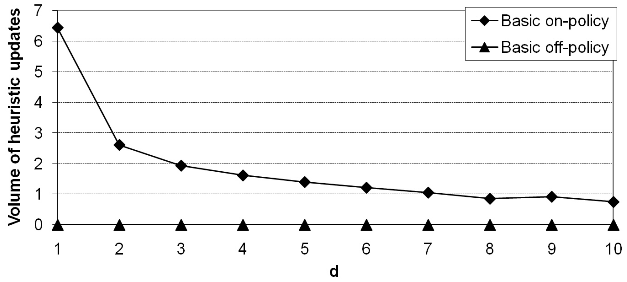


**Figure 1.** The volume of heuristic updates encountered per move with respect to the lookahead depth in the basic on- and off-policy experiments.

The results in Table 3 still show more pathology than in the basic off-policy experiment, so there must be a **third explanation**. Let $\alpha_{\text{off}}(d)$ and $\alpha_{\text{on}}(d)$ be the average number of states generated per move in the basic off-policy and on-policy experiments respectively. In off-policy experiments a search is performed every move, whereas in on-policy experiments a search is performed every $d$ moves. Therefore $\alpha_{\text{on}}(d) = \alpha_{\text{off}}(d)/d$. This means that in the basic on-policy experiment fewer states are generated at larger lookahead depths than in the basic off-policy experiment. Consequently the depths in the basic on-policy experiment are closer to each other with respect to the number of states generated. Since the number of states generated can be expected to correspond to the quality of decisions, cases where a deeper lookahead actually performs worse than a shallower one should be more common.

The first test of the third explanation is an on-policy experiment where a search is performed every move instead of every $d$ moves. The results in Table 4 confirm the explanation. The percentage of pathological problems is considerably smaller than in the basic on-policy experiment shown in Table 1: 34.7% vs. 93.7%. Since LRTS that searches every move is very similar to LRTA* [4], LRTA* can also be expected to be less pathological.

**Table 4.** Pathology on-policy when searching every move.

| **Degree** | 0 | 1 | 2 | 3 | $\geq 4$ |
|---|---|---|---|---|---|
| **Pathological problems [%]** | 65.3 | 14.6 | 8.6 | 7.1 | 4.4 |

The second test is to measure the number of states generated per move. Figure 2 shows that in the basic off-policy experiment and in the on-policy experiment when searching every move, the number increases more quickly with lookahead depth than in the basic on-policy experiment. The depths are thus less similar than in the basic on-policy experiment, which again confirms our explanation.

Experiments with eight-puzzle [8] showed that pessimistic heuristics can prevent the pathology. This inspired the **fourth explanation** of the pathology. During lookahead search, states with low heuristic values are favored. If the heuristic values are optimistic (as in our case), the lowest heuristic value is likely to be particularly far from the true value. With deeper lookahead, more states are considered and the chances of selecting a state with an especially inaccurate heuristic increase. If the heuristic values are pessimistic, the opposite is true: the states with accurate heuristic values are favored and the more states are considered, the more likely a state with a very accurate heuristic value will be selected.

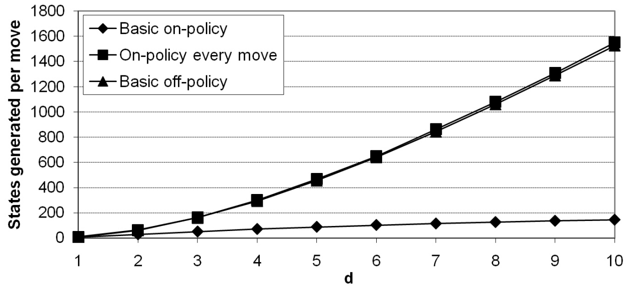


**Figure 2.** The number of states generated per move with respect to the lookahead depth in different experiments.

We verified the forth explanation with an on-policy experiment with pessimistic heuristic values. If the regular heuristic value of a state $s$ is $h(s) = h^*(s) - e$, where $e$ is the heuristic error, then the pessimistic heuristic value is $h_{\text{p}}(s) = h^*(s) + e$. Such a heuristic is unrealistic, but it should give us an idea of what to expect from realistic pessimistic heuristics, should we be able to design them. The results in Table 5 do show a decrease in pathology compared to the basic on-policy experiment shown in Table 1: 86.1% vs. 97.7%.

**Table 5.** Pathology on-policy with pessimistic heuristic.

| **Degree** | 0 | 1 | 2 | 3 | 4 | $\geq 5$ |
|---|---|---|---|---|---|---|
| **Pat. problems [%]** | 13.9 | 4.1 | 8.3 | 22.9 | 27.7 | 23.1 |

## 4 CONCLUSION

The first two expanations of the pathology do not seem to offer practical ways for avoiding the pathology. When ivestigating the third explanation, we learned that searching every move the way LRTA* does brings the pathology from 93.7% to 34.7%. It also generates up to 2.6 times shorter solutions. However, it increases the number of states generated per move roughly by a factor of $d$. This means that the number of states generated per problem when searching every move is up to 4.5 larger (at $d = 10$) than with the regular LRTS. A promising direction of research therefore seems to be a method for dynamically selecting the point at which a new search is needed.

Finally, the fourth explanation suggests that pessimistic heuristics may be less prone to the pathology. In addition, the solutions found using the pessimistic heuristic were nearly optimal (3.8–7.2 times shorter than with the regular heuristic), so pessimistic heuristics deserve further attention.

## REFERENCES


[1] Donald F. Beal, 'An analysis of minimax', in *Advances in Computer Chess*, volume 2, pp. 103–109, (1980).

[2] Vadim Bulitko and Greg Lee, 'Learning in real time search: A unifying framework', *Journal of Artificial Intelligence Research*, **25**, 119–157, (2006).

[3] Vadim Bulitko, Lihong Li, Russell Greiner, and Ilya Levner, 'Lookahead pathologies for single agent search', in *Proceedings of IJCAI, poster section*, pp. 1531–1533, Acapulco, Mexico, (2003).

[4] Richard E. Korf, 'Real-time heuristic search', *Artificial Intelligence*, **42**(2, 3), 189–211, (1990).

[5] Mitja Luštrek, 'Pathology in single-agent search', in *Proceedings of Information Society Conference*, pp. 345–348, Ljubljana, Slovenia, (2005).

[6] Mitja Luštrek and Vadim Bulitko, 'Lookahead pathology in real-time path-finding', in *Proceedings of AAAI, Learning for Search Workshop*, pp. 108–114, Boston, USA, (2006).

[7] Dana S. Nau, *Quality of Decision versus Depth of Search on Game Trees*, Ph.D. dissertation, Duke University, 1979.

[8] Aleksander Sadikov and Ivan Bratko, 'Pessimistic heuristics beat optimistic ones in real-time search', in *Proceedings of ECAI*, pp. 148–152, Riva del Garda, Italy, (2006).