

Unlucky Explorer: A Complete non-Overlapping Map Exploration

Mohammad Sina Kiarostami*, Saleh Khalaj Monfared*, Mohammadreza Daneshvaramoli*, Ali Oliayi†, Negar Yousefian†, Dara Rahmati*, Saeid Gorgin*

*School of Computer Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

† K.N.Toosi University of Technology, Tehran, Iran

{skiarostami, monfared, daneshvaramoli, dara.rahmati, gorgin}@ipm.ir

{ali.oliayi, negaryousefian}@email.kntu.ac.ir

Abstract—Nowadays, the field of Artificial Intelligence in Computer Games (*AI in Games*) is going to be more alluring since computer games challenge many aspects of AI with a wide range of problems, particularly general problems. One of these kinds of problems is *Exploration*, which states that an unknown environment must be explored by one or several agents.

In this work, we have first introduced the *Maze Dash* puzzle as an exploration problem where the agent must find the a *Hamiltonian Path* visiting all the cells. Then, we have investigated to find suitable methods by a focus on *Monte-Carlo Tree Search* (MCTS) and *SAT* to solve this puzzle quickly and accurately. An optimization has been applied to the proposed MCTS algorithm to obtain a promising result. Also, since the prefabricated test cases of this puzzle are not large enough to assay the proposed method, we have proposed and employed a technique to generate solvable test cases to evaluate the approaches. Eventually, the MCTS-based method has been assessed by the auto-generated test cases and compared with our implemented *SAT* approach that is considered a good rival. Our comparison indicates that the MCTS-based approach is an up-and-coming method which could cope with the test cases with small and medium sizes with faster run-time compared to SAT. However, for certain discussed reasons, including the features of the problem, tree search organization, and also the approach of MCTS in the *Simulation* step, MCTS takes more time to execute in Large size scenarios. Consequently, we have found the bottleneck for the MCTS-based method in significant test cases that could be improved in two real-world problems.

Index Terms—Monte-Carlo Tree Search, MCTS, Maze Dash, Exploration, Hamiltonian Path, SAT.

I. INTRODUCTION AND BACKGROUND

Graph Traversal is an important and famous problem in computer science with many applications in memory and storage systems [1], network flow [2], as well as computer games [3]. Advances in different aspects such as insuring the security [4], efficient path fining [5], and *Graph Exploration* in the robotics are also addressed by the researchers [6]. As the conventional approaches to solving this problem and other variations like *Tree Traversal*, Depth-first search (DFS), and Breadth-first search (BFS) are known to be effective in general. On the other hand, random-based approaches such as Monte-Carlo Tree Search (MCTS) are demonstrated to be efficient in many search-based problems and games as well as [7].

The fundamental problem of many simple computer games lays on solving specific computer or mathematical puzzles.

The solution methodology used in many of these games is very relevant to fundamental approaches. For instance, *Flow-Free* is a variant of a known mathematical puzzle named *Numberlink*, and interestingly, the problem could be addressed as a Multi-Agent Path Finding [8]. In this context, *Icosian Game* [9] as an old mathematical game invented by W.R Hamilton could be considered as a modified version of a Graph Traversal problem. The objective in Icosian is finding a *Hamiltonian cycle* along the edges of a dodecahedron, visiting all the vertexes of the graph by ending at the same point as the starting vertex. The Hamiltonian Path problem as an NP-Complete problem [10] has its own applications in various fields [11] with many solution methods [12].

In this article, we investigate the foundation of the *Maze Dash* game. We demonstrate that the constraints involved in solving the game, inevitably minimize the number of *Turning Movement* in the grid exploration procedure. Satisfying this particular condition applied in this game could be interesting in terms of real-world robot exploration since an extra cost is often associated with the *Turning Movement* in smart explorer vehicles [13]. Hence, finding an efficient and effective solution to the focused *Maze Dash* game could lead to faster *Grid Traversal* approaches where realistic restrictions in the robots are considered.

Furthermore, We mathematically define the underlying primary problem of the game as a particular case of a Hamiltonian Path problem. Then, by studying the problem's specifications, we tackle the problem with different possible approaches, including the MCTS, as one of the promising methods. We examine the unique characteristics of the involved tree search in detail and study the exclusive attributes of Hamiltonian Path in this problem.

A. Maze Dash Game

Maze Dash is a puzzle game with a single agent and a 2-Dimensional grid map. The map might have some obstacles or blocking cells. The agent moves in the map and marks the cells after visiting them by changing their color and can not return to the marked cells. So, each cell must be visited just once. Eventually, the puzzle aims to visit all the cells or to explore the whole of the map.

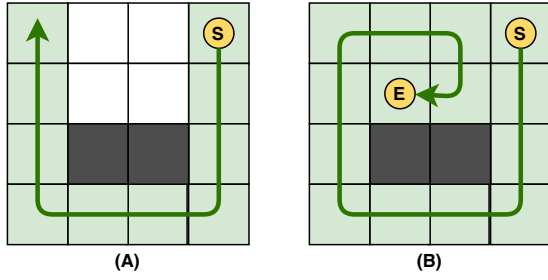


Fig. 1. *Maze Dash* game solving process

The essential rule in this puzzle is that if the agent chooses to go to the one the quad directions, it will continue to move until it reaches an obstacle or wall. As shown in Figure 1, the agent starts to move from the initial cell (S) and decides to go down to reach the wall or the border of the grid. Then, it keeps moving to explore all of the possible cells, finishing the traversal at the last cell (E).

B. Monte-Carlo Tree Search

MCTS is a best-first search algorithm with four main steps which are *Selection*, *Expansion*, *Simulation*, and *Backpropagation*. This algorithm uses Monte-Carlo methods to sample steps and create the search tree indeterminately to solve problems in their particular domain [7]. Like other tree-based approaches, the algorithm requires considering an initial state as its root to construct the search tree. In the context of our problem, the initial state is the state that the agent craves to move from its current cell.

The first step of the MCTS is the selection that the algorithm chooses the best node, which is a leaf at the moment based on the *Tree Policy*. Then, at the expansion point, all non-terminal children of the selected node, if exist, will be expanded. In the next step, simulation, MCTS strides in the search tree aimlessly based on a policy until it reaches a leaf. The obtained result will be evaluated and measured how much is this result is analogous to the desired result, and how many of the rules and conditions of the problem are satisfied. Finally, in the backpropagation, the results are propagated back through the tree, and all related node values are updated. After that, the next rounds will be iterated to find the suitable solution.

II. PROPOSED METHOD

A. Problem Definition

The exact definition of the problem as a modification of a *Hamiltonian Path* in a 2-D grid is described below. The set of $O = \{o_1, o_2, \dots, o_m\}$ is demonstrated as the *obstacle* set which determines the coordination of the obstacle cells in the grid. By considering an $N \times N$ grid, the function π identifies the movement path in the grid:

$$\pi : \mathcal{N} \rightarrow \mathcal{N} \times \mathcal{N} \quad (1)$$

The input of π function rises incrementally to represents the movement path in the grid. The output represents the

coordinates with the constrain of moving a single cell at each step to ensure the consistency of the solution path:

$$\begin{aligned} \pi(0) &= S \\ \text{Direction} &= \{(1, 0), (0, 1), (-1, 0), (0, -1)\} \\ \forall (N^2 - |O|) > i > 0 : \\ \pi(i+1) - \pi(i) &\in \text{Direction}, \pi(i+1) \notin O \\ \forall i, j : \pi(i) &\neq \pi(j) \end{aligned} \quad (2)$$

In Equation 2, S is the initial coordinate of the beginning cell. The Direction set D is the set of all possible movements that can be used in this context. As for the constraints regarding the minimum turning movement restrictions in the game, the π function falls into either one of the two *Straight Movement*, *Turning Movement* conditions as defined in Equation 3, respectively:

$$\begin{aligned} \forall (N^2 - |O|) > i > 1 : \text{oneofThree} : \\ \text{Straight} : \pi(i) + (\pi(i) - \pi(i-1)) &= \pi(i+1) \\ \text{Turning} : \pi(i) + (\pi(i) - \pi(i-1)) &\in O \\ \text{Turning} : \exists j < i : \pi(i) + (\pi(i) - \pi(i-1)) &= \pi(j) \end{aligned} \quad (3)$$

Note that the turning is occurred either by a blocking obstacle or a previously occupied cell by the earlier path. Hence, the i^{th} step in a cell must be as the same previous direction, or a turning movement happens.

Eventually, the solution of the game comprise of the adequate assignment for the π function satisfying all the constraints presented in Equations 2 and 3.

B. Promising Approaches

1) *SAT*: The accurate mathematical definition discussed, as the foundation of the problem, could be fed to a Satisfiability solver (SAT Solver, e.g., Z3 [14]) as a simple solution approach by converting the conditions into boolean constraints. In this respect, the boolean assignments of $t, s : \mathcal{N} \rightarrow \{0, 1\}$, should be defined to determined whatever a cell falls into the *Turning Movement* or the *Straight Movement* sets. Then, by creating the same constraints in the function π for each cell of the grid, the Satisfiability check could be performed over the described assignments of $\{t, s, \pi\}$.

2) *Backtracking*: As a naive approach to solve the problem, one could apply the *Backtracking* method to find the correct solution. The backtracking process is very similar to a *DFS* method. Considering the conditions of the game, a single branch of the possible solution is pursued until all the cells are traversed or a deadlock occurs. In the case of a failure, the movement path backtraces itself to the previous state and changes the branch by choosing another possible path. A simple backtrack demonstration is shown in Figure 2. The algorithm has to eliminate its current path to correct it since one of the cells is not visited.

3) *MCTS*: As mentioned in the background section, one of the particular and pronounced features of the game is that the agent must explore the grid with the minimum number of turns. More precisely, the agent only could change its direction

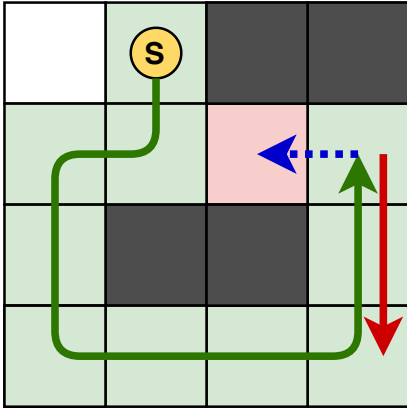


Fig. 2. Procedure of Backtracking in solving the puzzle

when it reaches the end of the current path. This constraint intensely affects the tree search of the puzzle. As shown in Figure 3, each non-terminal node of the tree search has only one or two children. Therefore, the *Branch Factor* of the tree is equal to or smaller than two making the tree significantly long in-depth. One of the approaches that could solve the problem is MCTS. As explained earlier, MCTS runs its four steps iteratively to construct the tree search and find the solution.

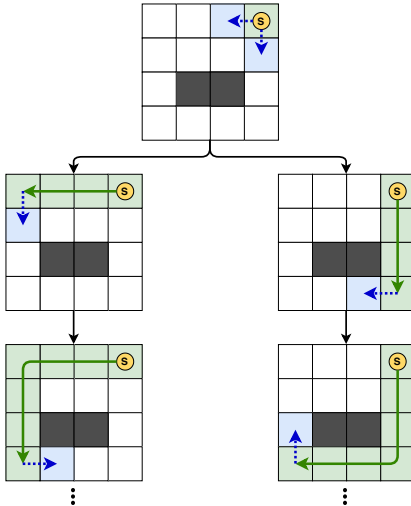


Fig. 3. Construction of the tree search of the puzzle

To solve the problem more efficiently, we have disregarded the nodes with a certain failure at their final stats to prune the tree. More precisely, the Expansion function was amended to block a terminal state with an inadequate value of the desired final state. Then, the Selection function blocks a node if all of its children were blocked before, by assigning 0 as the value of the node. Therefore, gradually wrong paths could be blocked efficiently. Nevertheless, even by employing these modifications to the algorithm, the execution time increases immensely due to the computational overhead caused by huge depth in the tree. Another optimization has been employed by applying the *Fast Rollout Function*, where instead of

constructing a new state for each *Child Node*, the state of the parent is updated each time. This optimization reduces the memory consumption of $\mathcal{O}(N^2)$ to $\mathcal{O}(1)$ for each node [15]. These details will be discussed in the result section.

Note that, another approach to the problem could be a BFS method where all of the branches are searched at the time without traversing deep into the final state. Not surprisingly, this method burdens considerable memory usage to store the branch states. This memory overuse could be predicted due to the tremendous depth of the tree search. Consequently, a BFS approach would not be a suitable solution compared to other possible methods.

III. EVALUATION

As the test cases of the Maze Dash game are not large enough to assess the methods correctly, we decided to utilize a method to generate large random test cases. First, an empty grid with the desired size is assumed. Then an agent starts to move through the grid randomly. Whenever the agent turns or changes its direction, an obstacle is placed at the next cell of the current cell. It means that there was a hypothetical obstacle in the path, so the agent decided to change the direction. For avoiding creating a unique path for solving the test case, we defined a variable as the number of obstacles. After generating the test cases, all of them have been tested by the Backtracking approach to make sure that they are solvable.

All of the compiled files were executed on a machine running Ubuntu 16.04 equipped with two Intel XEON E5 2697V3 CPUs clocked at 2.6 GHz and 128 GB of DDR3 RAM.

As shown in Table I, we have implemented and compared possible approaches to solve the defined problem. Each test case is executed by each method, 50 times, and the average values are presented. It is worth mentioning that the *Backtracking* algorithm is used to indicate our worst-case scenario, not to be a good rival. The MCTS approach performs well in small and medium-size test cases but could not cope with large ones. The results of the MCTS method could be discussed. First, most of the execution run-time of the algorithm is spent in the simulation step. Assume that I defines the number of iterations of the simulation in each MCTS traversal. In each simulation, the algorithm would traverse the tree down to N^2 depth. After selecting each node and adding one depth to the MCTS tree, the algorithm would be repeated. Thus, the search would be performed for another N^2 times. Ultimately, simulations process enlarges and will have an immense cost, calculated as follows:

$$SimulationsCost = I \times N^4 \quad (4)$$

This order is a huge cost for the problem since the agent could not determine or predict a complete solution until it tests all possible movements. For instance, the *Evaluation Function* could return 0.98 as a value of a final state, meaning that only 2 cells are not visited among 100. We know that this state is not the accurate answer but the algorithm would recognize

TABLE I
EVALUATION OF DIFFERENT ALGORITHMS FOR SOLVING *Maze Dash* GAME

Grid Size(N^2)	No Obstacle	SAT		Backtrack(DFS)		Randomize(MCTS)	
		Run-Time(S)	Memory(MB)	Run-Time(S)	Memory(MB)	Run-Time(S)	Memory(MB)
5x5	4	0.45	6.01	0.010	7.7	0.011	5.94
6x6	10	0.65	8.12	0.012	8.1	0.013	10.61
10x10	32	1.79	12.83	0.45	10.4	0.07	11.5
15x15	66	4.18	13.27	5.24	13.5	0.09	14.2
20x20	133	11.28	13.86	9.62	19.2	1.16	16.9
30x30	378	17.94	15.5	43.7	26.1	5.42	20.1
50x50	776	126.6	49.9	Failed	Failed	Failed	Failed

it as a promising state in the previous steps. So, the algorithm would never prune these kinds of states. Furthermore, owing to the non-overlapping feature of this modification of the exploration, after each movement, the agent would be faced by new obstacles. Moreover, previously visited cells are considered as dynamic obstacles and walls, which makes the algorithm unable to be optimized by any pre-process methods. Furthermore, as explained, the branch factor of the problem is equal or smaller than 2, constructing a tree with a huge depth without many branches. This kind of tree produces an arduous circumstance for MCTS in its simulation step.

IV. DISCUSSION AND CONCLUSION

In this article, we first introduced the Maze Dash puzzle as a modification of exploration problem or, more precisely, a non-overlapping exploration that could be solved by Hamiltonian Path. This means that the primary purpose of this problem is to visit all of the cells in 2-D grid. We have investigated promising approaches to find a proper solution. Although four methods implemented to compare with each other, the main focus was on MCTS and SAT. As expected, our implemented SAT could solve the auto-generated test cases accurately. Nevertheless, by reducing the number of obstacles in the grid, the execution time increased exponentially.

In small and medium-size test cases, MCTS could outperform SAT. In the large test cases, as explored, the simulation time increased uncontrollably since the algorithm recognizes the failure early in the simulation. As our observation, due to the non-overlapping feature of the problem, the agent considers its previous path as dynamic walls or obstacles. Therefore, the algorithm could not be optimized by pre-processing methods, but the SAT performs more beneficial in these test cases.

Investigating the introduced intricacy, the authors of this article came to realize two more practical and pronounced problems that should be considered. The first one is reasonably analogous to the current problem but differs in the non-overlapping constraint. It means that the aim of the problem is that an agent must explore all of the environment fast and accurately. However, the agent's previous path would not be defined as a new obstacle. Thus, the agent prefers not to use the visited cells but is not forced to do this. The second problem is the exploration of a grid to find a goal with minimum numbers of turns, by assuming that turning movement has an additional

cost, since the agent must reduce its velocity, stop, and then start to move again [13], [16]. By this constraint, the agent prefers to choose a path with lesser turnings. In our future studies, we would concentrate on these two problems, which would be useful in the real world applications, such as 2-D Robotic soccer.

REFERENCES

- [1] A. Aggarwal and S. Vitter, Jeffrey, "The input/output complexity of sorting and related problems," *Communications of the ACM*, vol. 31, no. 9, pp. 1116–1127, 1988.
- [2] T.-Y. Cheung, "Graph traversal techniques and the maximum flow problem in distributed computation," *IEEE Transactions on Software Engineering*, no. 4, pp. 504–512, 1983.
- [3] A. Plaat, J. Schaeffer, W. Pijls, and A. De Bruin, "Exploiting graph properties of game trees," in *AAAI/IAAI, Vol. 1*, 1996, pp. 234–239.
- [4] M. S. Karvandi, S. K. Monfared, M. S. Kiarostami, D. Rahmati, and S. Gorgin, "A way around umip and descriptor-table exiting via tsx-based side-channel attack," *arXiv preprint arXiv:2005.10333*, 2020.
- [5] J. E. Bobrow, S. Dubowsky, and J. S. Gibson, "Time-optimal control of robotic manipulators along specified paths," *The international journal of robotics research*, vol. 4, no. 3, pp. 3–17, 1985.
- [6] P. Fraigniaud, D. Ilcinkas, G. Peer, A. Pelc, and D. Peleg, "Graph exploration by a finite automaton," *Theoretical Computer Science*, vol. 345, no. 2-3, pp. 331–344, 2005.
- [7] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton, "A survey of monte carlo tree search methods," *IEEE Trans on Comput Int and AI in Games*, vol. 4, no. 1, pp. 1–43, March 2012.
- [8] M. S. Kiarostami, M. Reza Daneshvaramoli, S. K. Monfared, D. Rahmati, and S. Gorgin, "Multi-agent non-overlapping pathfinding with monte-carlo tree search," in *2019 IEEE CoG*, 2019, pp. 1–4.
- [9] E. Pegg Jr, "The icosian game, revisited," *The Mathematica Journal*, vol. 11, no. 3, pp. 310–314, 2009.
- [10] J. Hartmanis, "Computers and intractability: a guide to the theory of np-completeness (michael r. gary and david s. johnson)," *Siam Review*, vol. 24, no. 1, p. 90, 1982.
- [11] J. Cooper and R. Nicolescu, "The hamiltonian cycle and travelling salesman problems in cp systems," *Fundamenta Informaticae*, vol. 164, no. 2-3, pp. 157–180, 2019.
- [12] A. Bjorklund, "Determinant sums for undirected hamiltonicity," *SIAM Journal on Computing*, vol. 43, no. 1, pp. 280–299, 2014.
- [13] M. Davoodi, F. Panahi, A. Mohades, and S. N. Hashemi, "Clear and smooth path planning," *Applied Soft Comput*, vol. 32, pp. 568–579, 2015.
- [14] L. De Moura and N. Björner, "Z3: An efficient smt solver," in *International conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer, 2008, pp. 337–340.
- [15] M. Daneshvaramoli, M. S. Kiarostami, S. K. Monfared, H. Karisani, H. Khashehchi, D. Rahmati, S. Gorgin, and A. Rahmati, "Decentralized cooperative communication-less multi-agent task assignment with monte-carlo tree search," *arXiv preprint arXiv:1910.12062*, 2019.
- [16] H. Robots. Go 2015 spl finals: Nao-team htwk vs. b-human 1st half. Youtube. [Online]. Available: <https://www.youtube.com/watch?v=NFNEOooEQX4>