

Heuristic Monte-Carlo Tree Search for Mancala

Owen Siljander

University of Minnesota

December 18, 2019

Abstract

Depth-limited minimax with alpha-beta pruning has been, and still is, a very strong algorithm for two player games such as chess, where it has had great success. Monte-Carlo Tree Search (MCTS) emerged as a strong competitor, and largely gained interest due to its success in the previously human-dominated game, Go [3]. MCTS performed particularly well within Go due to its domain independence and the sparse search tree it constructs that can take long-term rewards into account. This has made MCTS a strong strategic player, but it still fails against minimax in some domains such as chess. Since the search tree that MCTS constructs is highly selective, it can miss critical moves, ones that lead to a quick win or defeat; this often leads to its defeat by human players or minimax. Improvements upon MCTS have been made in this respect, namely variants such as MCTS-Solver (which uses proven wins and proven losses) and hybrids between MCTS and minimax, which often lead to better results in search spaces with search traps. However, MCTS-Solver isn't quite sufficient enough to overcome a minimax search using strong heuristics. The hybrids are also quite weak in the sense that minimax adds computational overhead to MCTS, which can lead to its failure in search spaces that are devoid of search traps. In these search spaces, it is often better to use regular MCTS. The *mancala* family of board games is one domain where MCTS and minimax are roughly equally competitive due to its lack of search traps [4]. Variants of MCTS that use heuristic functions and modifications to minimax are explored here. One particular variant of minimax, known as forward search sparse sampling minimax (FSSS-Minimax), is a rollout based search, much like MCTS, that has been shown to outprune traditional alpha-beta minimax [9]. This paper judges the efficacy of FSSS-Minimax as a contender to MCTS and its heuristic variant in search spaces without search traps.

1 Introduction

Monte-Carlo Tree Search (MCTS) is a sampling based search algorithm that recently came to light because of its success over other popular algorithms such as Alpha-Beta Minimax ($\alpha\beta$ -Minimax). Since MCTS is a sampling based search that conducts game simulations, it can be more accurately described as a general format for an algorithm rather than a single algorithm in and of itself. Most commonly, the selection policy of MCTS is modified to use a solution to *exploitation-exploration* problems. These are problems in which the player (or agent) must attempt to balance how much it *explores* to gain new knowledge and how much it *exploits* to increase the accuracy of current knowledge. The popular variant of MCTS known as Upper Confidence Bound applied to Trees (UCT) uses a solution to these problems and has been shown to be optimal for simulation rewards in the range $[0, 1]$ [7]. UCT is optimal in the sense that given an infinite number of rollouts (game simulations), it will converge to the value yielded by minimax. Although UCT is a modification to MCTS, literature commonly uses the term MCTS and assumes that the particular selection policy is that of UCT. The same is assumed here.

Up until quite recently, the game Go was dominated by human players. MCTS was used to beat some of the top Go players, gaining it widespread attention. In general, the particular reason that other popular algorithms, such as Alpha-Beta Minimax ($\alpha\beta$ -Minimax), have failed to conquer Go is because there are no known heuristics that can consistently characterize the strength of the game board [3]. But since MCTS is a sampling based search, conducting game simulations on the actions it thinks are most promising, it is domain independent and can perform well within Go, even given that the search tree is quite non-trivial. This led to some research into MCTS as both a general game playing algorithm as well as its usage in other domains mostly conquered by $\alpha\beta$ -Minimax. Unfortunately, MCTS failed in a lot of domains, such as chess, largely because the search tree that MCTS constructs through game simulations is highly selective. It often can miss critical moves, meaning it either quickly loses or fails to quickly win the game. This has been described as a tactical weakness for MCTS, a more strategic algorithm.

[4] characterized search spaces that can be problematic for MCTS, namely search spaces that are relatively dense with *search traps*. A search trap (sometimes referred to as a *shallow trap* or a *shallow loss*) is a node in a search tree from which the opposing player has a guaranteed winning strategy in a finite number of moves (i.e. the opponent can win regardless of what actions the agent chooses). More descriptively, a node m within a search tree is called a *level- k search trap* if the opponent has a guaranteed winning strategy from node m that consists of at most k actions. Search traps have been indirectly referenced to be prominent in games such as chess, though usually described through some other terminology such as “sudden-death” games [10]. Search traps were first informally defined in [4] and this characterization of search spaces allows a more direct analysis of why MCTS is weaker than full-width minimax within some domains. Many variants upon MCTS and hybrids between it and minimax have been made to alleviate this weakness while maintaining domain independence. However, some of these algorithms tend to only perform well within search spaces that are relatively dense with search traps. This is often because the additional computation time added by the modifications aren’t worth it for the amount of information they yield. There are still many ways to attempt to improve MCTS within its problematic domains. Dropping the requirement for domain independence is the approach that this paper takes. Other algorithms, such as ones that can automatically generate heuristics based on game rules, can be used to bring back in domain independence. Further research into alleviating the pitfalls of MCTS are especially interesting as it may allow for the use of MCTS as a general game playing agent.

2 Related Work

[7] examines several different algorithms on several different variations of Mancala. Those of interest here are minimax, alpha-beta pruning minimax, advanced heuristic minimax (AHMM), and Monte-Carlo tree search (MCTS). All minimax searches were limited to search depths of 1, 4, and 8. AHMM, which uses better heuristics than other variations

for evaluating end states, is stronger than all other variations of minimax considered (with equivalent search depths). MCTS (up to 4000 iterations) only wins against AHMM when minimax is limited to search depths of 1 and 4. This demonstrates just how strong good heuristics are for searching. This motivates looking into combinations of MCTS and heuristics.

[9] extends forward-search sparse sampling from [8] to include minimax search. This algorithm, FSSS-Minimax, expands a subset of the nodes expanded by traditional alpha-beta pruning. Much like MCTS-Minimax hybrids, FSSS-Minimax attempts improve upon UCT by diminishing or removing issues of shallow traps from MCTS. As it may be apparent, [9] is very similar to [2], but it only considers the selection phase of MCTS as opposed to other useful hybrids of minimax and MCTS (e.g. in back propagation, rollout, selection of nodes). This algorithm also does not conduct game rollouts and instead assumes all states within the computation budget are terminal states. An evaluation function is run on the leaf node and this value is backpropagated much like in MCTS.

Search traps are largely problematic for MCTS since it poorly estimates the utility of a search trap and often gives it a high probability of being selected [4]. This isn't because MCTS simply doesn't have enough time to sample the search trap, even expanding 10 times the number of nodes as minimax, the utility it estimates for the search trap is still higher than that of minimax [4]. This pitfall of MCTS seems largely due to its backpropagation phase where it averages expected values back up the search tree [1]. A particular variant of MCTS called MCTS-Solver partially alleviates this weakness by augmenting MCTS to prove game theoretic values (e.g. proven wins, proven losses) [10]. However, this algorithm still doesn't outperform the best $\alpha\beta$ -Minimax algorithms, and so [2] explores MCTS and minimax hybrids to further close the gap between $\alpha\beta$ -Minimax and MCTS.

[2] tested three different variations of MCTS and minimax hybrids. The occurrence of search traps within the search space was found to generally be a good performance indicator of these hybrids. These hybrid algorithms use both of the modifications made by MCTS-Solver and UCT. The three different combinations of minimax as a sub-search of MCTS that will be discussed here are minimax in the rollout phase, selection and expansion phase, and the backpropagation phase. Minimax within the rollout phase of MCTS attempts to make more informative rollouts using a depth-limited minimax search, as opposed to uniformly random moves, to better estimate the utility of a node. While this does lead to a higher quality rollout, this adds significant computation time to the algorithm and the performance of the algorithm drops significantly in some domains [2]. The performance of this hybrid wasn't entirely consistent as it did perform strongly in other domains [1]. Minimax within the selection and expansion phase of MCTS runs a minimax search from a particular node after some arbitrary number of visits to that node. There are also other possible criteria for running a minimax search, but they were not tested. The theory here is to strongly verify the results of a node that looks promising. This hybrid was found to perform significantly better than MCTS in some domains. The last variant is minimax within the backpropagation phase of MCTS. Here minimax is used to better propagate game theoretic values back up the search tree to avoid re-sampling by MCTS. This variant performs well in some domains but in others,

it's neither strong nor weak.

These hybrid algorithms aren't that consistent among different domains though minimax within the selection and expansion phase seems to perform the best in general. This is because the density and depth of search traps within these games have a significant effect on the performance of MCTS and the hybrids. It's important to note that high level search traps (i.e. search traps with large k values) are largely ignored in literature. This is because any human player would not be able to detect these moves consistently (or at all) and they're often outside the search depth of minimax [4]. The use of minimax adds computational overhead but it doesn't yield much knowledge to MCTS if no search traps are present [1]. The presence of search traps even in synthetic search spaces is also shown to have the same affect upon MCTS [5]. These same characteristics are also present in MCTS-minimax hybrids that assume the existence of heuristics. Namely, [6] finds that minimax within the backpropagation phase can significantly improve the performance of MCTS in the game Mancala.

3 Approach

This paper looks at another vector by which one can improve the performance of MCTS within domains that are largely devoid of search traps. Particularly, combining strong heuristics into MCTS. To keep this fairly even, this will be compared to $\alpha\beta$ -Minimax and a modification of it, known as forward search sparse sampling minimax (FSSS)¹, that is also a sampling based search algorithm. Some modifications were made to this algorithm to make it suitable for this problem. In particular, FSSS was modified to be depth-limited and the bounds U and L that it calculates (much like α and β) were not compared for equality as in [9]. Instead, a sufficiently small difference between the two bounds would end its search.

The heuristics for minimax and FSSS are the same as in [7] for Advanced Heuristic Minimax. These are as follows:

Heuristic	Value
H1	Counters in left most pit
H2	Counters in all pits
H3	Number of non-empty pits
H4	Number of captured pieces
H5	Number of pieces captured by opponent

This has a slight discrepancy compared to the heuristics in [7]. Notably, heuristic number 5 is originally heuristic number 6. One of the original heuristics relied on the previous move made by the agent, which is not tracked and thus is not used. These heuristics make up the reward function for $\alpha\beta$ -Minimax and FSSS:

¹Not to be confused with regular forward search sparse sampling. The acronym "FSSS" is used for the sake of brevity.

$$\sum_{i=1}^5 H_i \cdot W_i$$

The weights, W_i , are the same as in [7] for their respective heuristic:

W1	W2	W3	W4	W5
0.198649	0.190084	0.370793	1	0.565937

MCTS will use the selection policy of UCT for all the variants following in this discussion. This selection policy uses the following function:

$$\frac{\text{Total reward}}{\text{Total visits}} + 2C\sqrt{\frac{\log(\text{Visits to root})}{\text{Total visits}}}$$

The “Visits to root” is the number of times the parent node has been visited. This occurs every time MCTS conducts a game simulation. The nodes on the path from the root node to the node being simulated from, are all visited as well. Here, the “Total reward” and “Total visits” are based on the current node. The former being the sum of backpropagated values of the reward function. The reward function in standard UCT is generally 1 if the game was a win, 0 otherwise. Occasionally, the value 0.5 is used for game draws. MCTS here will follow the same convention. C is an arbitrary constant, the optimal value of which is $\frac{\sqrt{2}}{2}$ for a reward function in the range of $[0, 1]$. When a game simulation finishes, the reward function is evaluated on the final state and this is backpropagated up the tree. Each node adds the reward to their total reward.

In the first variant, plain MCTS will be used with a reward function in the range of $[0, 1]$. In the second variant, the reward function of MCTS will use the same advanced heuristic functions as previously described. Although this reward modification may cause the selection policy to be sub-optimal, the same selection policy as UCT will still be used. This variant of MCTS will simply be called Heuristic MCTS (HMCTS).

4 Empirical Results

MCTS was used with heuristics in its rollout phase. After a game was simulated, a more advanced heuristic function, rather than the plain reward function, was run on the terminal state. This value was then backpropagated up the search tree. Both FSSS and $\alpha\beta$ -Minimax used the same heuristic function. Not much is to be said about the runtime of these algorithms as often a majority of the time was spent updating player actions. The number of nodes expanded is not covered in depth either as MCTS is arbitrarily limited in the number of nodes it expands by the person running the algorithm. Additionally, the number of nodes expanded by FSSS was consistent with [9] in that it expanded less nodes than $\alpha\beta$ -Minimax on average. A direct comparison was not made. MCTS was limited to 25, 50, and 100 iterations for two reasons. Any higher number of iterations resulted in an excessive time calculating a move. Additionally, both FSSS and $\alpha\beta$ -Minimax expanded between 50 to 300 nodes on average when running

with depths 1, 3, and 5. These depths were chosen as these are common depths within literature. Note that at depth 1, both algorithms degenerate into a variation of a greedy algorithm. FSSS, $\alpha\beta$ -Minimax, and MCTS using a heuristic function are all compared against each other. $\alpha\beta$ -Minimax will serve to represent a baseline and will be compared against regular MCTS.

FSSS was tested with depth limits of 1, 3, and 5, similar to $\alpha\beta$ -Minimax. The number of nodes expanded were consistent with [9]. For example, one test yielded an average of 4,000 expansions for FSSS and an average of 5,000 for $\alpha\beta$ -Minimax. This is merely to establish some consistency between the modified depth limited FSSS used here compared to FSSS used in [9]. Further results are omitted as this is not the focus of the tests. FSSS was run against MCTS using the simple reward function with iterations of 25, 50, and 100. These numbers were chosen as $\alpha\beta$ -Minimax and FSSS commonly expanded anywhere from 60 to 300 nodes.

4.1 Data

It should be noted that “Average # of moves” refers to the total of moves until the game ended. Games were arbitrarily restricted to have at most 300 moves to avoid infinite loops. Each configuration was run for 50 games. Much longer than this was mostly impractical or infeasible due to runtime issues. The win rates are how often the first listed algorithm won against the second. First, $\alpha\beta$ -Minimax against regular MCTS:

$\alpha\beta$ Depth	MCTS Iterations	Win Rate (%)	Average # of moves	Draws (%)
1	25	96	41.96	0
1	50	100	48.13	0
1	100	100	46.6	0
3	25	96	39.9	0
3	50	100	41.26	0
3	100	100	34.4	0
5	25	100	34.93	0
5	50	100	39.266	0
5	100	100	42.93	0

Secondly, FSSS against HMCTS:

FSSS Depth	HMCTS Iterations	Win Rate (%)	Average # of moves	Draws (%)
1	25	80	79.53	10
1	50	83	73.2	13
1	100	67	78.433	7
3	25	60	88.1	13
3	50	67	81.56	7
3	100	60	83.93	10
5	25	67	54.766	3
5	50	73	57.833	7
5	100	50	57.833	17

Thirdly, FSSS against $\alpha\beta$ -Minimax:

FSSS Depth	$\alpha\beta$ Depth	Win Rate (%)	Average # of moves	Draws (%)
1	1	100	63	0
1	3	0	142	0
1	5	0	216	0
3	1	100	189	0
3	3	100	117	0
3	5	0	252	0
5	1	0	270	0
5	3	0	107	0
5	5	0	295	0

Lastly, $\alpha\beta$ -Minimax against HMCTS:

$\alpha\beta$ Depth	HMCTS Iterations	Win Rate (%)	Average # of moves	Draws (%)
1	25	83	77.86	3
1	50	80	78.5	7
1	100	87	85.1	0
3	25	60	79.866	7
3	50	47	86.56	3
3	100	47	83.7	7
5	25	77	67.366	0
5	50	57	55.933	10
5	100	67	64	13

5 Analysis

Overall the results were very inconclusive. $\alpha\beta$ -Minimax against regular MCTS had very inconsistent results. Normally in literature, MCTS would be given on the order of several thousands of iterations to calculate its move. However, this was entirely infeasible in this situation. The results are not given here, but even with up to 5,000 iterations, the performance of MCTS against $\alpha\beta$ -Minimax held the same. This indicates that there

may be an issue with the underlying implementation of MCTS or $\alpha\beta$ -Minimax. The only somewhat conclusive data from this is that as the depth of $\alpha\beta$ -Minimax increased, the game ended in generally fewer moves. FSSS against HMCTS also had inconsistent results. Its win rate appeared to decrease as its depth increased. This test shared some similarities with the previous one in that the game ended in shorter moves as the depth limit increased. FSSS against $\alpha\beta$ -Minimax yielded no useful information at all. The results of the games were very deterministic. Attempts were made to alleviate this by introducing a small amount of randomness into $\alpha\beta$ -Minimax. Particularly, if a move for $\alpha\beta$ -Minimax was within some negligibly small value of the current known best, $\alpha\beta$ -Minimax would select this new action with a 50% probability. This did introduce results that were more believable: FSSS and $\alpha\beta$ -Minimax were often head to head with equal depth limits, whichever had the higher depth limit often won more often. However, this caused $\alpha\beta$ -Minimax to perform significantly worse against MCTS, dropping its win rate far below rates that were consistent with existing data in literature. $\alpha\beta$ -Minimax against HMCTS likewise had inconsistent results.

There was an overall trend of the average number of moves to game completion decreasing as the computational limit on the algorithms was extended. This indicates to some degree that the algorithms, as pairs, often got better at simply capturing the pieces on the board. However, this didn't produce any usable data about the algorithms themselves. Some of these results were very inconsistent with existing tests in the literature which indicates that there is an issue with the underlying implementation. It's worth noting that the underlying mechanism that the algorithms operated upon was verified for correctness. FSSS was also tested to ensure that it performs exactly the same as in [9] given a sufficient depth limit.

The original intuition was that HMCTS is sub-optimal (due to its limited number of iterations). This could explain why the performance of $\alpha\beta$ -Minimax and FSSS decreases as they increase their look ahead, since they assume the opponent is playing optimally. However, this doesn't make sense given that $\alpha\beta$ -Minimax performs exceedingly well against MCTS with the same restrictions as HMCTS. HMCTS did perform better than MCTS in comparison despite the individual tests not producing accurate results.

6 Conclusion

The results of the tests didn't yield much useful information. Largely, it just indicated that either the underlying implementation was faulty, or that using MCTS with a relatively small number of iterations is not sufficient to gain any insight. The data points to the latter as HMCTS did perform better than MCTS in comparison so it remains to see if this would change as the number of iterations increased. Interestingly, FSSS and $\alpha\beta$ -Minimax were both very deterministic. The reason for this is very unclear. The initial choice perhaps has a very large effect on the outcome of the game. The fact that the initial choice is almost always the same may indicate that the two algorithms essentially play the same game every time. This is also supported by the fact that some randomness introduced into $\alpha\beta$ -Minimax caused the results to be more believable (though not

included here). FSSS also performed somewhat as expected. The number of nodes it expanded were on average fewer than the number by $\alpha\beta$ -Minimax.

Characterizing search spaces for algorithms is still an ongoing area of research, namely understanding exactly how the level of the search trap can influence the performance of MCTS. Further testing of these algorithms within other domains that do have search traps could be done. For further research, one could combine heuristics into MCTS in different ways. This could be within the expansion phase to select which child node to simulate from. This could also be within the selection phase to perhaps force a higher probability of selecting nodes which hold a higher heuristic value, as opposed to using a pure UCT selection policy.

References

- [1] Hendrik Baier and Mark HM Winands. “MCTS-minimax hybrids”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 7.2 (2014), pp. 167–179.
- [2] Hendrik Baier and Mark HM Winands. “Monte-Carlo tree search and minimax hybrids”. In: *2013 IEEE Conference on Computational Intelligence in Games (CIG)*. IEEE. 2013, pp. 1–8.
- [3] Cameron B Browne et al. “A survey of monte carlo tree search methods”. In: *IEEE Transactions on Computational Intelligence and AI in games* 4.1 (2012), pp. 1–43.
- [4] Raghuram Ramanujan, Ashish Sabharwal, and Bart Selman. “On adversarial search spaces and sampling-based planning”. In: *Twentieth International Conference on Automated Planning and Scheduling*. 2010.
- [5] Raghuram Ramanujan, Ashish Sabharwal, and Bart Selman. “On the behavior of UCT in synthetic search spaces”. In: *Proc. 21st Int. Conf. Automat. Plan. Sched., Freiburg, Germany*. 2011.
- [6] Raghuram Ramanujan and Bart Selman. “Trade-offs in sampling-based adversarial planning”. In: *Twenty-First International Conference on Automated Planning and Scheduling*. 2011.
- [7] Gabriele Rovaris. “Design of artificial intelligence for mancala games”. In: (2017).
- [8] Thomas J Walsh, Sergiu Goschin, and Michael L Littman. “Integrating sample-based planning and model-based reinforcement learning”. In: *Twenty-Fourth AAAI Conference on Artificial Intelligence*. 2010.
- [9] Ari Weinstein, Michael L Littman, and Sergiu Goschin. “Rollout-based game-tree search outprunes traditional alpha-beta”. In: *European Workshop on Reinforcement Learning*. 2013, pp. 155–167.
- [10] Mark HM Winands, Yngvi Björnsson, and Jahn-Takeshi Saito. “Monte-Carlo tree search solver”. In: *International Conference on Computers and Games*. Springer. 2008, pp. 25–36.

The algorithms on the following pages serve as reference for the results described above. The algorithms should perform exactly the same as their standard counterparts (there are no known issues), though the data does not agree with this.

Algorithm 1 UCT

```

1 def default_policy(n: Node):
2     m = copy(n)
3     UCT_NODES_EXPANDED += 1
4     while m is not a leaf node:
5         play random move for m
6     return m.reward()
7
8 def back_propagate(n: Node,  $\delta$ : float):
9     while n is not None:
10         n.visits += 1
11         n.total_reward +=  $\delta$ 
12         n = n.parent
13
14 def best_uct_child(n: Node, c: float):
15     best =  $-\infty$ 
16     for i in n.children:
17         res = i.total_reward/i.visits +  $c \cdot \frac{\sqrt{2}}{2} \cdot \sqrt{\log(n.visits / i.visits)}$ 
18         if res > best:
19             best = res
20             max_child = i
21     return max_child
22
23 def expand(n: Node):
24     a = pick an untried action from n
25      $n_0 = T(n, a)$ 
26     n.untried_actions.remove(a)
27     n.children.append( $n_0$ )
28      $n_0$ .parent = n
29      $n_0$ .action = a
30     return  $n_0$ 
31
32 def tree_policy(n: Node, c):
33     while n is not a leaf node:
34         if n.untried_actions:
35             return expand(n)
36         else:
37             n = best_uct_child(n, c)
38     return n
39
40 def uct(n: Node, c):
41     while within computational budget:
42          $n_1 = \text{tree\_policy}(n, c)$ 
43         delta = default_policy( $n_1$ )
44         back_propagate( $n_1$ , delta)
45         n.play(A(best_uct_child(n, c)))

```

Algorithm 2 Depth-Limited Minimax with Alpha-Beta Pruning

```
1 def alphabeta(n,  $\alpha$ ,  $\beta$ , depth, is_root):
2     if n is a leaf node or depth == 0:
3         return n.reward()
4     best_value = None
5     best_action = -1
6     for action in n.get_actions():
7         res_value = alphabeta(T(n,action),  $\alpha$ ,  $\beta$ , depth-1, False)
8         if is_root and res_value > best_value:
9             best_action = action
10        if n.player == 1:
11            best_value = max(best_value, res_value)
12             $\alpha$  = max( $\alpha$ , best_value)
13        elif n.player == 2:
14            best_value = min(best_value, res_value)
15             $\beta$  = min( $\beta$ , best_value)
16        if  $\alpha$  >=  $\beta$ :
17            break
18    if is_root:
19        n.play(best_action)
20        return
21    else:
22        AB_NODES_EXPANDED += 1
23        return best_value
```

Algorithm 3 Depth-Limited FSSS-Minimax

```

1 def traverse(n: Node,  $\alpha$ : float,  $\beta$ : float) -> tuple:
2   for i in n.get_actions():
3     if i not in A(n.children):
4       FS_NODES_EXPANDED += 1
5       n.children[i] = T(n,i)
6       n.Uprime[i] = min( $\beta$ , n.children[i].U)
7       n.Lprime[i] = max( $\alpha$ , n.children[i].L)
8    $\alpha' = \alpha$ 
9    $\beta' = \beta$ 
10  if n.player == 1:
11     $i^* = \text{argmax}(n.Uprime)$  # Pick action that maximizes U
12     $v = \max_{i \neq i^*}(n.Uprime)$  # Pick max U value that isn't by  $i^*$ 
13     $\alpha' = \max(\alpha, v)$ 
14    if  $\alpha' == n.Uprime[i^*]$ :
15       $\alpha' = \alpha' - \epsilon$ 
16    else:
17       $i^* = \text{argmin}(n.Lprime)$ 
18       $v = \min_{i \neq i^*}(n.Lprime)$ 
19       $\beta' = \min(\beta, v)$ 
20      if  $\beta' == n.Lprime[i^*]$ :
21         $\beta' = \beta' + \epsilon$ 
22    return  $i^*, \alpha', \beta'$ 
23
24 def search(n: Node,  $\alpha$ : float,  $\beta$ : float, limit: int) -> None:
25   if n.leaf or limit == 0:
26     n.L = n.U = n.reward()
27     return
28    $i^*, \alpha', \beta' = \text{traverse}(n, \alpha, \beta)$ 
29   search(n.children[ $i^*$ ],  $\alpha', \beta', \text{limit} - 1$ )
30   actions = n.get_actions()
31   for i in actions:
32     if n.player == 1:
33       n.L = max(n.L, n.children[i].L)
34       n.U = max(n.U, n.children[i].U)
35     elif n.player == 2:
36       n.L = min(n.L, n.children[i].L)
37       n.U = min(n.U, n.children[i].U)
38   return
39
40 def fsss(n: Node, limit: int):
41   while abs(n.L - n.U) >  $\epsilon$ :
42     search(n,  $-\infty, \infty, \text{limit}$ )
43   action =  $\text{argmax}(n.children.L)$  # Get action which yields largest L
44   n.play(action)

```
