

Sokoban: Search in a complex domain

Yann Chazallon, Nicolas Dossou-Gbété, Tony Chan Ki Hong and Michal Staniaszek

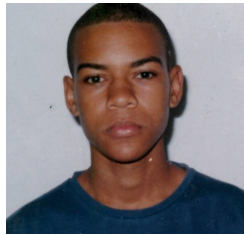
October 18, 2013

Abstract

In this report we describe our implementation of a *Sokoban* solver. We use a two stage method to find the solution. A bi-directional best-first search with manhattan distance as a heuristic is used to find the path of boxes on the map. This disjoint path is then linked by using A* to find the path between successive box moves. The solver is currently able to solve 47 of the 100 evaluation maps on Kattis, and 81 of the 100 maps in the simpler test set within the 11 second time limit.



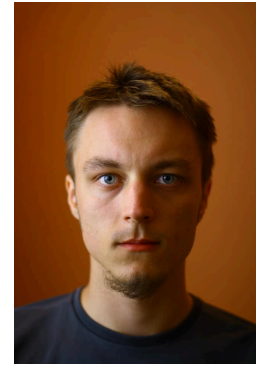
Yann Chazallon
91/01/16



Nicolas Dossou-Gbété
89/09/26



Tony Chan Ki Hong
92/07/25



Michal Staniaszek
90/12/07

1 Introduction

Sokoban is a puzzle game which was written in 1981 by Hiroyuki Imabayashi. First published in 1982, it is now a very popular game, with many clones available on the internet. The player controls a warehouse keeper (for which *sokoban* is the Japanese word), whose job it is to push boxes onto goal locations on the map. The player can move in four directions (up, down, left or right) on the map, which is split into discrete cells. The player can push boxes, but is unable to pull them. To be able to push a box, the player must be adjacent to it, and there must be an empty space behind the box into which it can be pushed. Only one box can be moved at a time; if two boxes are contacting each other, pushing one box does not move the other. While there are many graphical implementations, there are also many text-based implementations which use symbols for representing parts of the board, an example of which is shown in Figure 1.

The game has garnered some interest in the artificial intelligence (AI) community due to the difficulty of finding automatic solutions. Games are an ideal platform for the development and testing of AI techniques, as game environments are much simpler and less (or not at all) noisy, and also have sets of simple rules for interaction with the world. Working in such environments

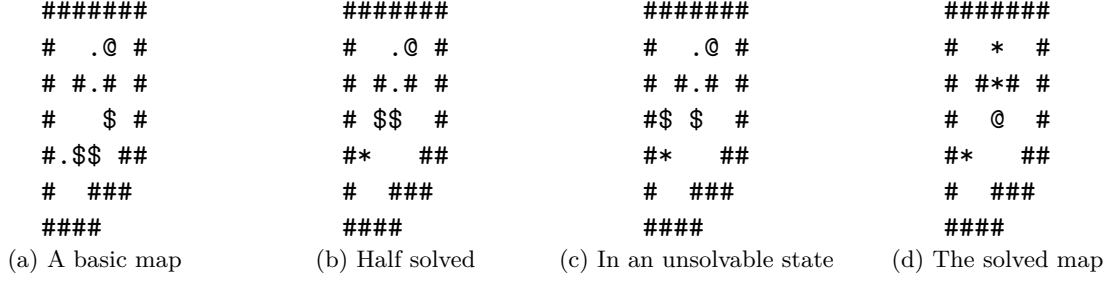


Figure 1: A typical *Sokoban* map. The player is represented by @, or + when on a goal, boxes by \$ or * on a goal, walls by #, goals by a period, and empty space by a blank. The map in (c) is unsolvable, as the two boxes on the left are both against a wall. This is called a *deadlock*.

allows for more control and closer investigation of the relevant parts of the problem that is being considered. *Sokoban* in particular is an interesting problem due to its high branching factor, and the depth of the search trees that are generated when attempting to solve a problem. Even relatively simple problems can take upwards of 100 moves to solve, and more complex problems can exceed 500 moves, even in the optimal case. Additionally, it is possible to extend the search tree indefinitely—there is no time limit, and boxes can be pushed without restraint, so long as the push is valid. The branching factor is the total number of valid actions that can be applied to any box reachable from the player’s current position. In the worst case, if there are N boxes, all of which are accessible and can be pushed in any direction, the factor is $4N$ (each box can be pushed in 4 directions). Although in practice it is not possible to access all boxes, or push them in an arbitrary direction, but most maps have a large enough number of boxes that the branching factor has a large impact on the number of states expanded.

2 Method

The solver should be able to read string-based map representations into a manipulable object. It must be possible to move boxes and the player on this map representation. In addition, the solver must be able to deal with both static and dynamic deadlocks. A heuristic is required to drive the search towards the goal. Finally, a search method is required, which, given a start and goal state, must be able to find a path between the states, or indicate that a path does not exist.

2.1 Map Representation

To be able to solve a *Sokoban* map, it is necessary to have a map representation which can be manipulated and used to gain information about the state of the board. The representation of the state is required by any search method in order to be able to generate successor states and find a path to the goal state. Our map representation uses a two-level model, where dynamic objects (boxes and the player) are separated from the static part of the map, that is, the goals, empty space, and walls. By separating these two parts of the map, we are able to reduce the storage space required for each board, as the static map is represented only once. When move validity must be checked, the dynamic objects are “collapsed” onto the static map.

The static map also stores a costmap which relates each position on the board to a set of costs to reach each goal. The costmap is calculated by doing a flood fill which attempts to pull a single box from each goal to all points on the map, with no other boxes present. This gives an optimistic estimate of the number of steps required to move a box from any position on the

map to any goal when the map is populated with boxes, as in practice the path of the box may be obstructed by other boxes.

Functionality for checking board equality is also required to be able to perform search, as it is necessary to check whether states have been visited previously in order to prevent unnecessary re-evaluation of states. Our map representation has different types of equality based on the state expansion procedure being used.

The state expansion defines how successors of a given state are generated. The player space expansion is based on the motion of the player alone; the player can move up, down, left or right, and so the maximum number of successors is 4. In this expansion, exact equality is used—the position of all dynamic objects on the board is checked, and the player and all boxes must be in exactly the same locations for the boards to be equal.

The board space expansion is based on the motion of all boxes in the state. With this expansion, the maximum number of successors is $4M$, where M is the number of boxes which can be reached by the player, as each box could be pushed in one of four possible directions. To determine the successor states of this state, it is first necessary to find all the positions which are accessible by the player. This is done by starting a flood fill from the current player position and finding points at which a box can be accessed. With this expansion, the board equality no longer checks the exact location of the player. Instead, the top leftmost position accessible by the player is used to represent the player. This means that states in which the player has the same accessible area, and where boxes are in the same positions are considered equal.

hashing!

2.2 Deadlock Detection

A deadlocked state is a state from which further expansions can never give a solution. If they are detected early, they can be removed from the search tree, pruning big sections of it and making the game much easier to solve. We identified two classes of deadlocks:

2.2.1 Static deadlock

It's a deadlock caused by a wrong position of the box relative to the walls of the map. For example, pushing a box in a corner makes it impossible to be moved again.

We detect them using the static costmap, created and filled at the launch of the program. The positions for which we don't have any cost are the ones out of reach from the goals, and therefore the ones where a box will be in a deadlock situation.

2.2.2 Dynamic deadlock

They are deadlock created by the position of a box relative to all the other ones. They have many possible manifestations and thus are much harder to detect. We only considered the most basic ones

YANN EXPLAIN HERE!

2.3 Heuristics

2.4 Search Method

The search methods form the core of the solver. Using them, we find a path between a starting board state and the goal state, where all boxes are on goals. We implemented three different search methods; A*, best-first search and bi-directional search. A* expands nodes in a specific order based on the value of a state evaluation function $f(s) = g(s) + h(s)$, where s is some board

state. The value $h(s)$ is the heuristic value of the state, and represents an estimate of how far the state s is from the goal. The value of $g(s)$ depends on the path that was taken to reach this node—the longer the path, the higher the value. A* places states into a priority queue and expands them starting from nodes with a lower f value; a shorter distance to the goal. This evaluation function means that so long as $h(s)$ is an underestimate (an *admissible* heuristic) of the true cost to the goal state, A* is both optimal and complete, that is, it will always find a solution if one exists, and that solution will be optimal[2].

In best-first search, the cost to reach the current node is not considered in the expansion, and so the evaluation function simplifies to $f(s) = h(s)$. This results in the search always expanding the state with the lowest cost to the goal, regardless of how far it is from the start state. This results in a loss of optimality and completeness [2], but can lead to a goal being found much faster than in the case of A*. Since we do not care about finding an optimal solution, and speed is of the essence in this problem, best-first search is more appropriate than A*. Our implementation is shown in Algorithm 1.

To further improve the speed of the search, we implemented a bi-directional search. While both A* and best-first search have a complexity of $\mathcal{O}(b^d)$, running two searches, one from the start state to the goal, and one from the goal state to the start, reduces this to $\mathcal{O}(b^{d/2})$ [2], where b is the branching factor, and d is the solution depth. The search terminates when the frontiers of the two searches meet. This can be understood intuitively by considering a circle centred on the front node, with a radius d , with the goal node on the edge of the circle. The area inside this circle is very large compared to two smaller circles centred on the start and the goal meeting in the middle. Our implementation of bi-directional search is shown in Algorithm 2.

Our approach involves splitting the problem into two distinct search problems. First, we find the actions required to move the boxes from the start state to the goal state using a best-first bi-directional search. Once this intermediate path has been found, we run a series of A* searches to find the path of the player between each successive box movement, which gives us the final path from the start to the goal. When finding the player path, we consider boxes as immovable objects.

3 Implementation

In our implementation, we use several objects which allow for cleaner code and easier access to required data. The most basic part of the implementation is the **Board** class, which holds methods to deal with moving the player and accessing various information about the state of the board. The **Board** object holds a list of dynamic points, as well as a reference to the static map. The static map is implemented as a singleton, so all objects created have a reference to that same object. This reduces memory usage, as the static part of the map is not duplicated for each **Board** object.

As we are using bi-directional search, the search methods have been implemented in a way that allows their usage in both a bi-directional and a single directional search. This is important as we do not perform only a single search to find the path to the goal. To achieve this, we use a **step** method to perform a single step of the search. This step, in the case of A* and best-first search, is to expand the front of the priority queue, and add the successors to the open lists, performing any relevant checks. The best-first search implementation is shown in Algorithm 1. Algorithm 2 shows our implementation of bi-directional search, which makes use of best-first search.

The search methods make use of a custom **SearchNode** class, which stores a parent node, a **Board** representing the state of that node, and a **BoardAction** object, which represents the

Algorithm 1: Best-first search

input : An open list O , containing a start node S . A goal node G . A closed list C .

output: A list of action, point tuples (A, P) , containing the path from S to G .

Function *findPath()*

```
    while  $O \neq \emptyset$  do
        successors = step()
        if  $G \in \text{successors}$  then return successors.get( $G$ ).unwind()
    end
```

end

Function *step()*

```
    front = open.front()
    successors = front.expand()
    closed.add(front)
    foreach  $n \in \text{successors}$  do
        if  $n \notin C \wedge n \notin O$  then
            // nodes in  $O$  sorted by heuristic value
             $n$ .costEstimate =  $h(n.\text{state}, G)$ 
             $O$ .add( $n$ )
        end
    end
    return successors
```

end

action used to generate the board by a point object (the location of the pushed box) and an associated **Action** enum, which is one of $\{up, down, left, right\}$, and indicates the direction of the push. The **unwind** method of the **SearchNode** recurses up to the start node, and reconstructs the path from there to the node on which the method was called.

The dynamic lock algorithm we've implemented allows the detection of cases where boxes are locking each other. Since this algorithm is used for pruning search branches, it's important to avoid false positives that could lead to losing solution states. The basic idea is that every box has two degrees of freedom. If any of the boxes loses its two degree of freedom permanently then it's considered locked, and thus the state is not worth to be expanded. We consider a degree of freedom locked if the box is locked on one of its side, fox example it's vertically locked if

4 Evaluation

experimental evaluation and a thorough analysis of your results. This section should include: How well did your agent do on the different maps?

54: a box must be placed in a key position (x) before putting other boxes on their goals. 66: not many box movements to do, with a limited number of possibilities at a time, can be easily bruteforced. 93: No goal-box matching is done, so the boxes are stacked around the area with 3 and we run out of time.

Algorithm 2: Bi-directional search

Data: Search implementations forwards and backwards, initialised with start and goal.

Result: A list of tuples of (point, int) containing motion of boxes from start to goal.

```
begin
  current  $\leftarrow$  forwards
  opposite  $\leftarrow$  backwards
  new  $\leftarrow \emptyset$ 
  key  $\leftarrow$  null
  repeat
    swap current and opposite
    new  $\leftarrow$  current.step()
    if  $node \in new \wedge node \in opposite.open$  then key  $\leftarrow$  node
  until  $key \neq null$ 
  fkey  $\leftarrow$  forwards.open.get(key)
  bkey  $\leftarrow$  backwards.open.get(key)
  fActions  $\leftarrow$  fkey.unwind()
  bActions  $\leftarrow$  bkey.unwind()
  return merge(fActions, bActions.reverse())
end
```

Algorithm 3: Single box dynamic lock detection

input : box position B .

output: boolean representing whether or not the box at point B is locked.

if B is on goal **then**

 | **return** false

end

closed $\leftarrow \{\emptyset\}$

return isMovable(B , closed)

Function isMovable(Point P , Set closed)

if $P = WALL$ **then**

 | **return** false

else if $P = BOX$ **then**

if $P \in closed$ **then**

 | closed.add(P)

 | **return** false

else

 | **return** ((isMovable($P.left$, closed) \wedge isMovable($P.right$, closed)) \vee

 | ((isMovable($P.up$, closed) \wedge isMovable($P.down$, closed)))

end

else

 | **return** true

end

end

Search Method	Time limit		
	5 sec	11 sec	15 sec
A*	12	15	16
Best First	56	60	64
Bi-directional Best First	76	81	82
Bi-directional A*	39	41	43

Figure 2: The number of test cases solved out of 100 available in the offline test set, running various search implementations with the manhattan distance heuristic. Note the lack of large differences in the number of solved maps with different time limits.

Search method	Avg expanded	Avg opened	Avg rejected	Avg time
A*	9149	33382	17929	3.4
Best First	10741	40473	19499	2.0
Bi-directional Best First	3801	12727	4136	1.0
Bi-directional A*	4591	16462	4340	2.0

Figure 3: Some statistics for the search nodes, and the average solution time in the tests in Figure 2. Expanded nodes are those nodes whose successors have been generated. Opened nodes is the total number of nodes placed onto the open list. Rejected nodes are successors not added to the open list because they would generate locked states.

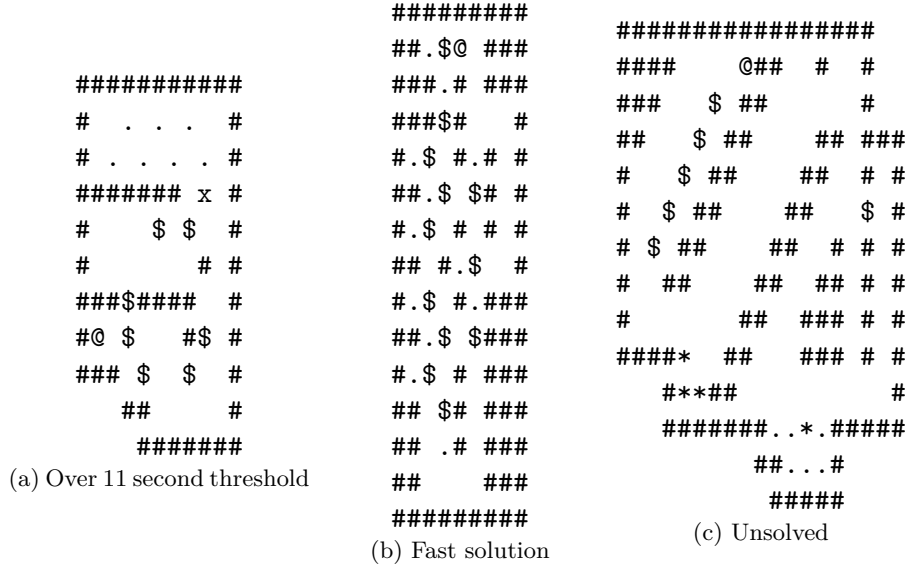


Figure 4: Examples of edge cases for our solver. Maps 54 (a), 66 (b) and 93 (c) from the test set. (a) is a border case, solvable within 15 seconds, but not 11. x indicates a critical location. (b) is solved in only 0.09 seconds. (c) is not solved even with a 15 second time limit.

5 Discussion

During our first meeting, we discussed several ideas about how to solve the problem, based on brief readings of some previous work [3, 1]. We decided that in the ideal case we should attempt to implement some of the techniques mentioned in these papers to produce an effective solver. The first step we too was to create the foundations for the solver, by implementing a board representation and some search algorithms, which we knew we would need based on our discussions. Work was distributed by each group member choosing something to work on, and then working on the parts of the solver related to that. This worked reasonably well, but meant that some members did not fully understand the code that was being used, because they had no part in writing it. We attempted to amend this by commenting methods with enough information to understand what the method was *supposed* to do. In order to check that important methods were doing what they were supposed to be doing, we used JUnit tests, running them whenever there were significant changes to the code to ensure that everything was working as expected.

Our approach to developing the solver led to some problems. Because each member was working on a somewhat different task, we did not meet directly very often, which resulted in too little communication and only a vague idea of what we were aiming for. As such, we spent the first couple of weeks adding unnecessary code to the solver, which would be removed later. This left us with a solver which, at the milestone deadline, could solve the four basic maps, but nothing else. Once we started meeting more often, we were able to make progress on the solver, but most of the progress was fixing errors that were caused by misunderstandings of the intended use of methods, and various other bugs which could not be directly detected by the unit tests. Fixing these problems left us with little time to consider any more complex techniques that we could use to improve the solver, which negatively impacted on the final result.

Our first approach to the problem was to use the motion of the player as the state expansion in an A* search implementation. This solver was able to complete very simple maps, but the size of the search space was far too large to solve any of the maps in the test set. Our second attempt was to use an expansion which considered all valid moves of boxes in a given board state. This reduced the search space to only those moves which were relevant to solving the problem. Once all bugs had been fixed, this version of the solver could solve a small number of maps in the test set, but continued to be hampered by slowness. This version was improved by using a more efficient equality check based on a string representation of the board, and a best-first search implementation. With this, we were able to solve approximately 15 of the maps from the online test set. Our current approach was reached by our need to improve the speed of the algorithm, and we thought that rather than using improved heuristics, changing the search algorithm would be more effective. We implemented a minimum matching heuristic, but the additional complexity resulted in the search taking longer despite the improved heuristic estimates. The performance measure used during the process was the number of maps of the Kattis test set solved by the solver. Profiling was also used to judge the effects of more efficient implementations of parts of the solver.

How would you solve it if you were asked to do it again given what you know now?

References

- [1] Andreas Junghanns and Jonathan Schaeffer. “Sokoban: Enhancing general single-agent search methods using domain knowledge”. In: *Artificial Intelligence* 129.1 (2001), pp. 219–251.

- [2] S.J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall series in artificial intelligence. Prentice Hall, 2010.
- [3] Timo Virkkala. “Solving Sokoban”. MA thesis. University of Helsinki, 2011.