

A Use of publicly available code and data

PuzzleJAX is based on the *PuzzleScript* game engine and Domain-Specific Language, and we further include copies of the original code within our repository for the purpose of validating existing *PuzzleScript* games in our engine. Since *PuzzleScript* is provided with an MIT license, we include the same license in the *PuzzleJAX* repository. We also consulted with *PuzzleJAX*'s author, Stephen Lavelle, during this project's development.

To validate our engine, we used a script to scrape over 800 games from an online database [23], following links to Github gists containing standalone *PuzzleScript* game files. We additionally validated against the games contained in the *PuzzleScript* Gallery², and authored a wide variety of minimal test scenarios during implementation of various features. It may also be possible to scrape games from the (currently active) *PuzzleScript* forum³ (e.g. by seeking out Github gist links in threads with the "[GAME]" tag), or from Itch.io⁴ (with these having the additional benefit of meta-data such as user ratings, comments, and number of plays; though these do not always link to the source code in a Github gist, or do not do so in a consistent way). Searching for *PuzzleScript* game file gists directly through the Github REST API may also be possible, given clever use of search keywords to circumvent pagination limits.

In this work, we do not distribute any curated dataset of actual human-authored *PuzzleScript* games. Instead, our contribution is the *PuzzleJAX* engine itself. The set of *PuzzleScript* games above are used primarily to demonstrate *PuzzleJAX*'s coverage of a vast array of possible games, and to ensure maximum interoperability with the established *PuzzleScript* DSL. Researchers may either use the *PuzzleJAX* engine to run newly designed *PuzzleScript*-style games, or to benchmark the performance of various methods on extant *PuzzleScript* games, potentially drawn from one of the sources above at their own discretion.

The exemplar *PuzzleScript* games presented in the main paper are largely drawn from the *PuzzleScript* Gallery, where they are presented with permission from the game authors. We list these exemplar games below, with links to these games in the *PuzzleScript* IDE (where they are playable and editable), and authorship credits:

- [Sokoban](#) (under Load Example -> Tutorial -> Basic Example) ported by Stephen Lavelle
- [Sokoban Match 3](#) (under Load Example -> Tutorial -> Match 3) by Stephen Lavelle
- [Lime Rick](#) by Tommy Tuovinen
- [Take Heart Lass](#) by Kevin Zuhn
- [Blocks](#) by Liam K Sheehan
- [Kettle](#) by Stephen Lavelle
- [Atlas Shrank](#) by James Noeckel
- [Multi-Word Dictionary Game](#) by Sarah Northway
- [Travelling Salesman](#) by Rabbit from Hell
- [Zen Puzzle Garden](#) by Lexaloffle
- [Notsnake](#) (under Load Example -> Elementary -> Notsnake) by Terry Cavanagh
- [Slidings](#) by Alain Broebecke
- [Constellation Z](#) (under Load Example -> Intermediate -> Constellation Z) by Stephen Lavelle

B Ethical considerations

PuzzleJAX is intended as a benchmark to assist in developing more generally capable and human-like AI agents, in particular by surfacing questions about the role of insight to solve a mechanically and semantically rich space of diverse puzzle games. We acknowledge that the overarching goal

²<https://www.puzzlescript.net/Gallery/index.html>

³<https://groups.google.com/g/puzzlescript>

⁴<https://itch.io/games/made-with-puzzlescript>

of creating generally capable AI agents may present both dangers and benefits to humanity. While these broader questions are out of scope for the present discussion, we believe that benchmarks like *PuzzleJAX* are crucial in understanding AI agents and learning algorithms which appear to have super-human abilities in some domains, but whose limitations are often poorly understood. *PuzzleJAX* is particularly relevant because it brings to the fore a swath of domains in which we expect many state-of-the-art agents and algorithms are likely to fail in surprising and perhaps counter-intuitive ways, even despite the apparent simplicity of the tasks at hand.

PuzzleScript’s DSL makes it easy, for example, to invert the canonical semantics of a game like *Sokoban*, such that with a simple variation to the game’s rules, the player now pushes a crate forward by moving *away* from it (as in [Okosban](#)). We expect that in games with such inverted or otherwise alien semantics, LLMs may have particularly difficulty in generating competent strategies (even supposing a more robust LLM-player pipeline is developed to address their difficulties in solving more canonical puzzles). As such *PuzzleJAX* can serve as an effective test of the abilities of LLMs to reason and problem-solve in the kind of out-of-distribution scenarios they may encounter once deployed into the wild, which situations may ultimately be of high consequence of users and designers.

In terms of *PuzzleJAX*’s impact on game designers, we hope that by fostering the development of more capable puzzle-solving agents, designers of *PuzzleScript* games may eventually be able to automatically playtest their games more effectively. *PuzzleScript*’s creator has recently expressed apprehension around embedding a best-first-search-driven solver agent⁵ into the *PuzzleScript* IDE, given that it might lead designers to create games that are significantly complex or challenging from the perspective of tree search, but potentially un-interesting or less fun or enjoyable for human players⁶. Given that *PuzzleJAX* facilitates the development of a wide variety of AI player agents beyond simple tree search—such as those based on LLMs, or those involving Reinforcement Learning—we hope that developers might ultimately have access to a diverse set of potentially human-like agents, allowing them to automatically measure proxies of human enjoyment or satisfaction (granted, this will likely require significant algorithmic advances, and benchmarking any such proxy benchmarks against actual human playtraces and surveys).

As alluded to in our Conclusion, *PuzzleJAX* also potentially facilitates the use of LLMs or genetic programming to generate new puzzle games automatically (e.g. by leveraging metrics generated by diverse player agents inside an evolutionary loop, as in [\[30\]](#)). Concerns may be raised here around the potential for automating away the process of game design, and burying human ingenuity and artistry in a barrage of AI-generated content that maximizes superficial metrics of player retention or engagement. In this regard, we advocate for the development of design assistant tools that incorporate human feedback and allow designers to intervene in the process of automatic game generation, as in [\[36\]](#), or as in the general paradigm of design through interactive evolution [\[38, 37, 35\]](#).

C Additional implementation and validation details

To validate the fidelity of *PuzzleJAX*, we use breadth-first search to find solutions for each level of each game in our collected dataset. We cap the number of environment steps during search at 100,000 and set a timeout of 1 minute. Where search does not find a winning state, we return the action sequence leading to the highest score, and in case of ties prefer longer action sequences (in hopes of exploring more of the game’s state space and thus ensuring a more robust validation). (The full results of this search procedure on the collected dataset of games is reported in [Table 1](#).) We then initialize each game and level in *PuzzleJAX*, and replay the action sequence, ensuring that it results both in the win conditions being met, and in an equivalent state (in terms of the layout of object in the level).

We report the results of this validation pipeline in [Table 2](#), and find that over 400 existing *PuzzleScript* games are valid in *PuzzleJAX*. Over 250 games are fully valid in *PuzzleJAX* (with each level’s solution in JavaScript resulting in the same outcome in *PuzzleJAX*), among these games with over 50 rules.

⁵Available at <https://github.com/Auroriox/PuzzleScriptPlus/blob/master/README.md>.

⁶<https://x.com/increpare/status/1905568607410532690>

Total Games	955
Valid Games	260
Partially Valid Games	147
Total Levels	7507
Successful Solutions	1781
Compile Errors	102
Runtime Errors	713
Solution Errors	2894
State Errors	325
Unvalidated Levels	808

Table 2: Results of validating *PuzzleScript* games in *PuzzleJAX*, by using breadth-first search to generate solutions for each level in JavaScript, then replaying these solutions in JAX, and ensuring they lead to equivalent end-states.

Of the over 7,000 individual levels in our dataset, 1,781 admit valid solutions in *PuzzleJAX*. Though this already constitutes a wealth of novel tasks for learning and reasoning agents, it means that a large number of levels result in errors (or remain unvalidated—most likely due to timeouts or memory issues during compilation). A large number of compile errors likely result from *PuzzleJAX*’s not yet capturing the extensive permissiveness of *PuzzleScript*. Already, we conduct preprocessing to clean up some of the syntactic errors which *PuzzleScript* affords (e.g. in rule definitions, if the cell boundary token “|” is contained between kernels—i.e. “] | [”—it is ignored; if the line detector token “...”, which can occupy a cell within a kernel to denote that the cells on either side of it may be separated by an arbitrary number of tiles, appears between kernels—i.e. “[...]”—the kernels are joined and the line detector is placed within its own cell in the kernel), but more examination of those games which cause issues with our Lark parser after pre-processing will be necessary to improve interoperability with *PuzzleScript*.

Solution errors—discrepancies between the win-state resulting from the solution found in JavaScript and that resulting from replaying the same solution in *PuzzleJAX*—usually indicate some difference between implementations of mechanics in the JavaScript and JAX engines, and continued development will seek to address them. During development of *PuzzleJAX*, for example, we used such discrepancies to ensure that rules were being broken down into rotational variants in the right order (so that, in *Carnival Shooter!*, for example, when the player “shoots” while next to two enemies, the enemy to the left of the player will be removed before the enemy to their right).

The one major feature which, to our knowledge, remains unimplemented in *PuzzleJAX* is the “rigid” keyword, which is used to simulate rigid-body physics. The use of this keyword appears in only 9 games in our dataset (leading to 9 compilation errors). We omit it for simplicity, given that its implementation appears relatively involved, and would require the use of additional channels in our level-state representation. The *PuzzleScript* documentation stresses this point, in fact (with the author writing that they “kinda regret adding this keyword to the engine”) and strongly advises the user to deploy other strategies to simulate rigid-body physics⁷.

In addition to the features described in the body of this paper, we note that we also implement the “line detector” feature (denoted in the *PuzzleScript* DSL as an ellipsis), which recognizes patterns separated by an arbitrary number of tiles along a row or column. Under the hood, we treat line detectors as a special kind of kernel that detect sub-kernels (the groups of cells on either side of the ellipsis) across the board, then detect if these subkernels’ activations fall in some ordered sequence along a line. These sequences are considered in order of the least to the most space between the sub-kernels. The line projection function then iterates through these detected lines in order, attempting to apply their respective subkernels until this has an effect on the board.

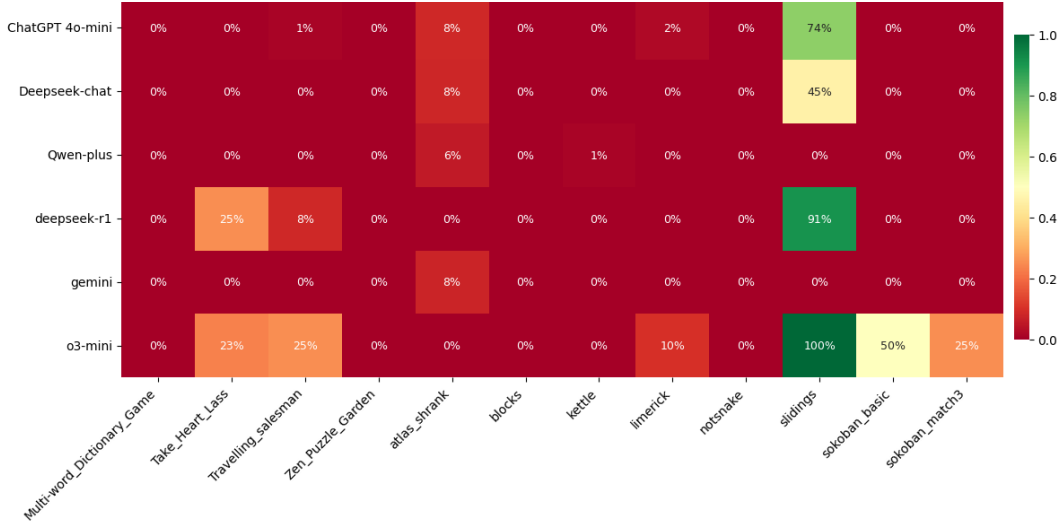


Figure 5: Average win rate comparison across different language models and games. The heatmap shows performance variations where darker red indicates lower performance (0%) and green indicates higher performance (up to 100%). Each cell represents the average win rate of a specific model on a particular game task.

D Additional results

D.1 LLMs

For our LLM experiments, we employed both reasoning-enabled LLMs and non-reasoning LLMs. Based on the experimental results presented in Figure 5, we observe significant performance variations across different LLMs when evaluated on 12 distinct games compiled in *PuzzleJAX*. The findings reveal that all model performance is highly task-dependent, with no single model demonstrating consistent superiority across all evaluated games. Notably, o3-mini achieved perfect performance (100% win rate) on the *Slidings* puzzle task and demonstrated strong capabilities in several other games, including *Sokoban Basic* (50%), *Take Heart Lass* (23%), *Travelling Salesman*, and *Sokoban match 3* (25%). DeepSeek-R1 exhibited exceptional performance on the *Slidings* puzzle task (91% win rate) while showing moderate success in strategic games such as *Take Heart Lass* (25%) and *Travelling salesman* (8%). ChatGPT-4o-mini displayed a more balanced performance profile, achieving its highest success rate on the *Slidings* puzzle (74%) and moderate performance on *Atlas Shrank* (8%) and *Limerick* (2%). In contrast, models such as Qwen-plus and Gemini showed limited success across most tasks, with Qwen-plus achieving only 6% on *Atlas Shrank* and 1% on *Kettle*, while Gemini’s performance peaked at 8% on *Atlas Shrank*. The results suggest that certain games, particularly *Slidings* puzzles, may be more amenable to current language model capabilities, while others such as *Multi-Word Dictionary Game*, *Blocks*, *Notsnake*, and *Zen Puzzle Garden*, remain challenging across all evaluated models.

D.2 Reinforcement learning

For our Reinforcement Learning experiments, we use the fully-jitted training loop written in JAX provided by [14], allowing us to take advantage of *PuzzleJAX*’s jitted environment step function. (We add utilities for saving model checkpoints and rendering episodes intermittently during training.) We use the above repo’s default hyperparameters for PPO, training agents on each level over 5 different random seeds for a total of 5 million environment steps each, with a learning rate of $1e^{-4}$, 128 rollout steps per minibatch, with 4 minibatches and 10 update epochs, with a $\gamma = 0.99$, an entropy coefficient of 0.01 and a value function coefficient of 0.5. We set batch size as large as possible for each game and level combination within the constraints of the VRAM available on the GPUs we use for training.

⁷<https://www.puzzlescript.net/Documentation/rigidbody.html>

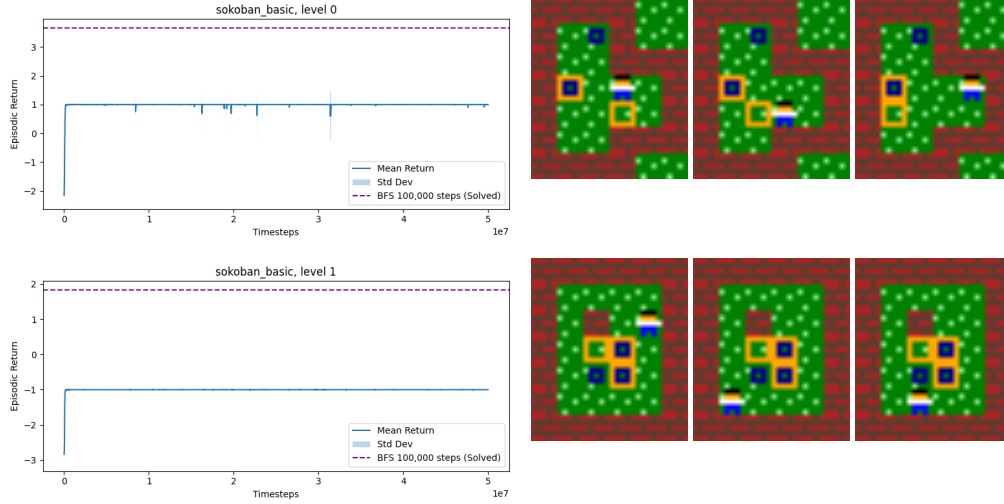


Figure 6: Comparison of RL against breadth-first search in *Sokoban*. Episode rollouts from RL are pictured on the right. Here, the agent greedily maximizes the heuristic (the sum of manhattan distances between targets and their nearest crates), preventing discovery of optimal solutions.

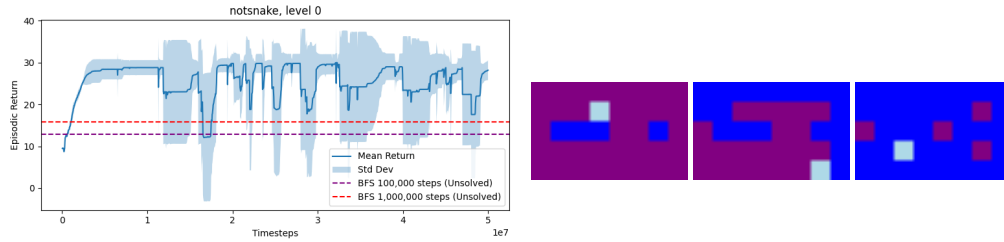


Figure 7: Comparison of RL against breadth-first search in *Notsnake*. Episode rollouts from RL is pictured on the right.

We use our institution’s high-performance computing cluster for training, and include in our code-base scripts for deploying sweeps of training jobs to nodes in this cluster via SLURM (we provide similar scripts in order to parallelize the tree search and JAX episode-rollouts in our *PuzzleJAX* validation pipeline). The GPUs on this cluster include the NVIDIA RTX8000, V100, A100, and H100, and the AMD MI100 and MI250. (We use a separate consumer machine with an NVIDIA 4090 for our speed profiling experiments).

While RL can be deceived by the heuristic functions of *Sokoban Basic* (Figure 6) and *Limerick* (Figure 8), in which positive reward can be sparse and optimal solutions may require first moving circuitously “away” from rewarding states, it does well in games admitting very short solutions such as *Slidings* (Figure 9) and *Kettle* (Figure 10), and games that constitute dense reward combinatorial optimization problems such as *Notsnake* (Figure 7), where it even discovers a better solution than did breadth-first search after 1 million environment steps (though it still does not discover the *exact* solution). (Note however that this does not necessarily constitute a fair comparison, which would arguably require running search for an equal number of environment steps, and/or comparing the wall clock times of each algorithm.) In *Take Heart Lass* (Figure 11), agents perform well in early levels which effectively constitute a simple control task involving running away from the encroaching despair and toward a goal, whereas on later levels that require efficiently pushing blocks to clear paths in the nick of time, or block or undo the propagation of despair tiles, the agent often runs into dead-ends or otherwise winds up trapped by despair tiles while attempting to bee-line toward to goal tile.

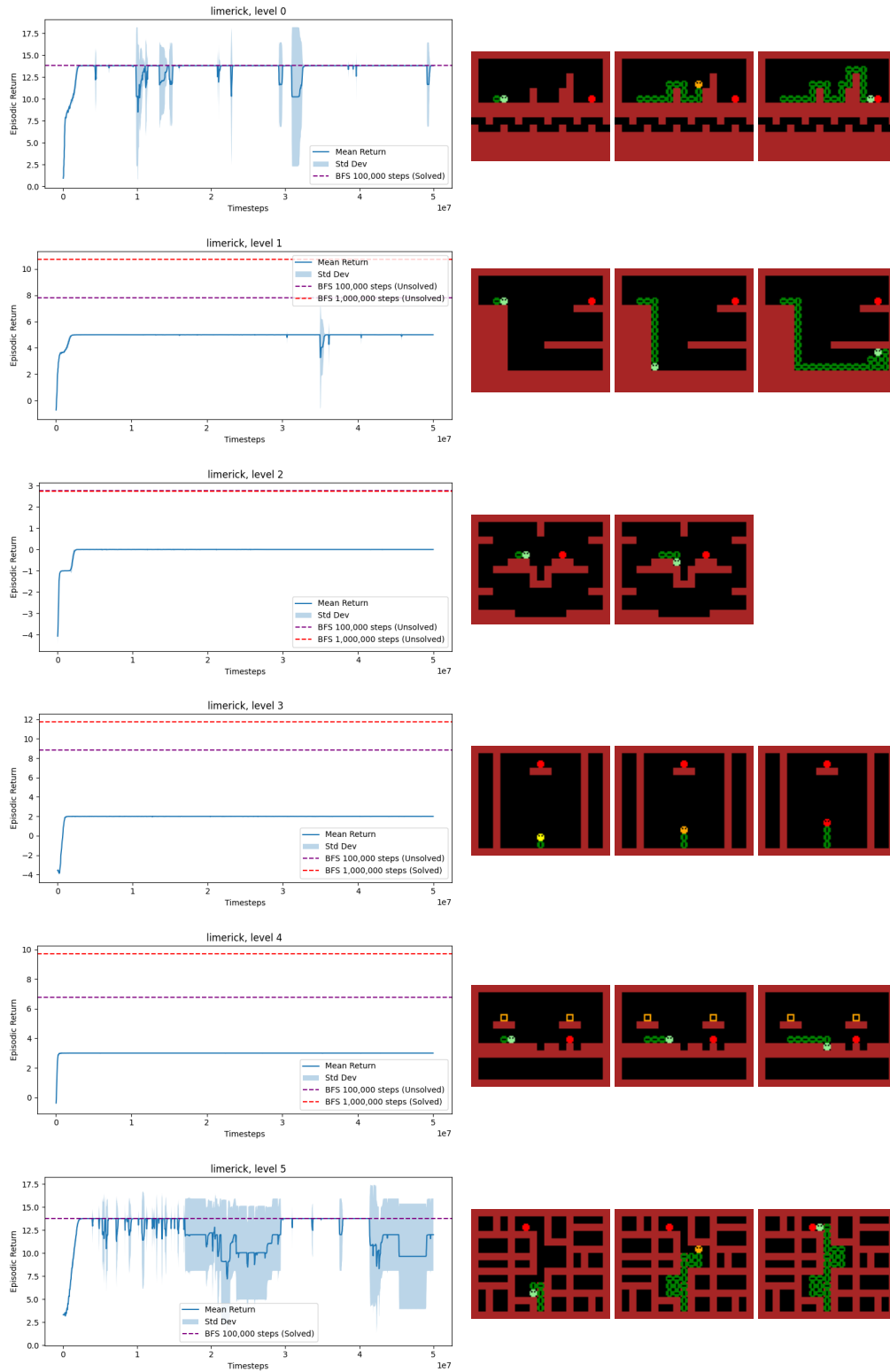


Figure 8: Comparison of RL against breadth-first search in *Limerick*. Episode rollouts from RL are pictured on the right. Agents only master levels with a relatively straightforward path to the goal. They do not generally uncover strategies involving significant roundabouts away from the goal, and can fall prey to “obvious” traps along the more direct path.

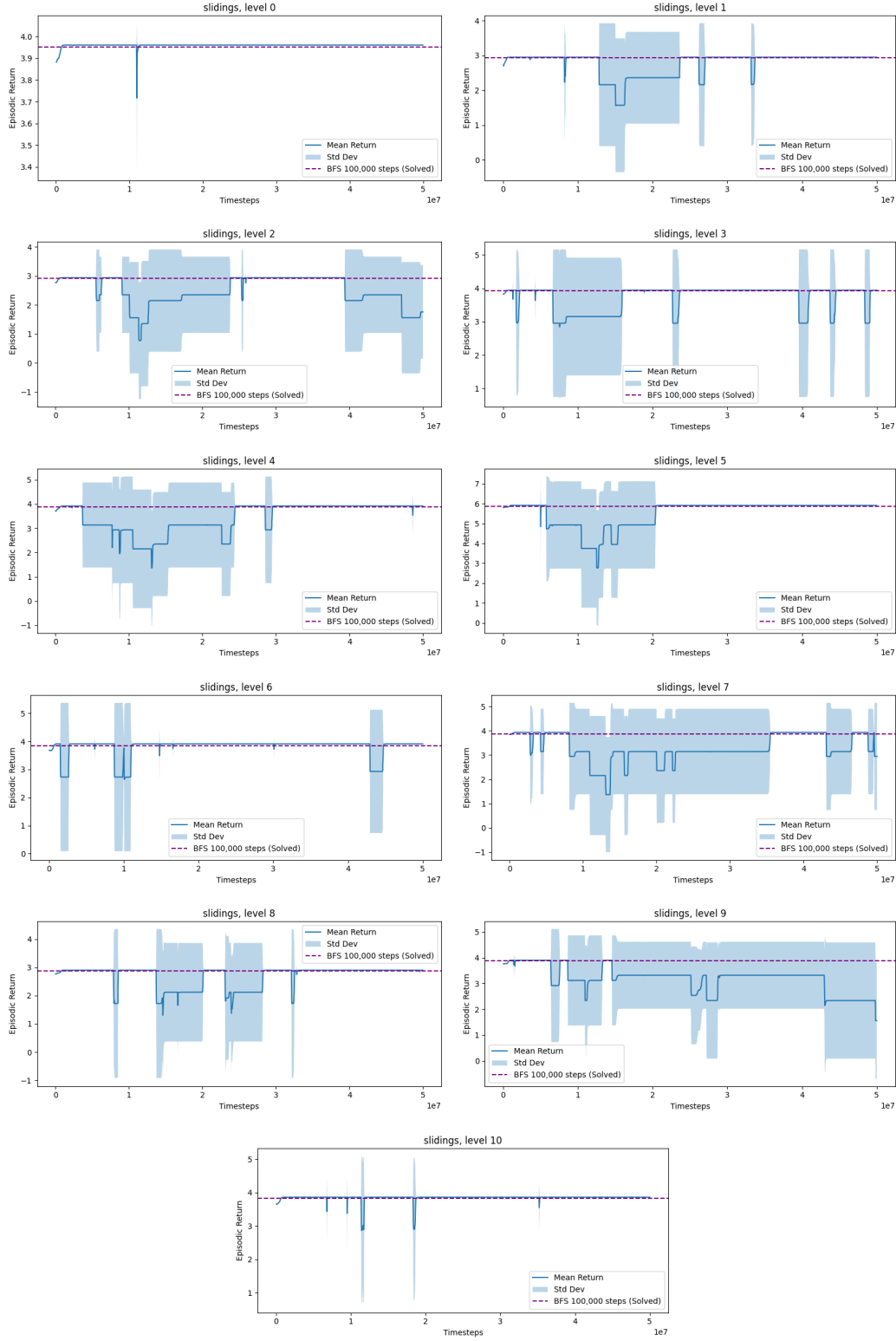


Figure 9: Comparison of RL against breadth-first search in *Slidings*.

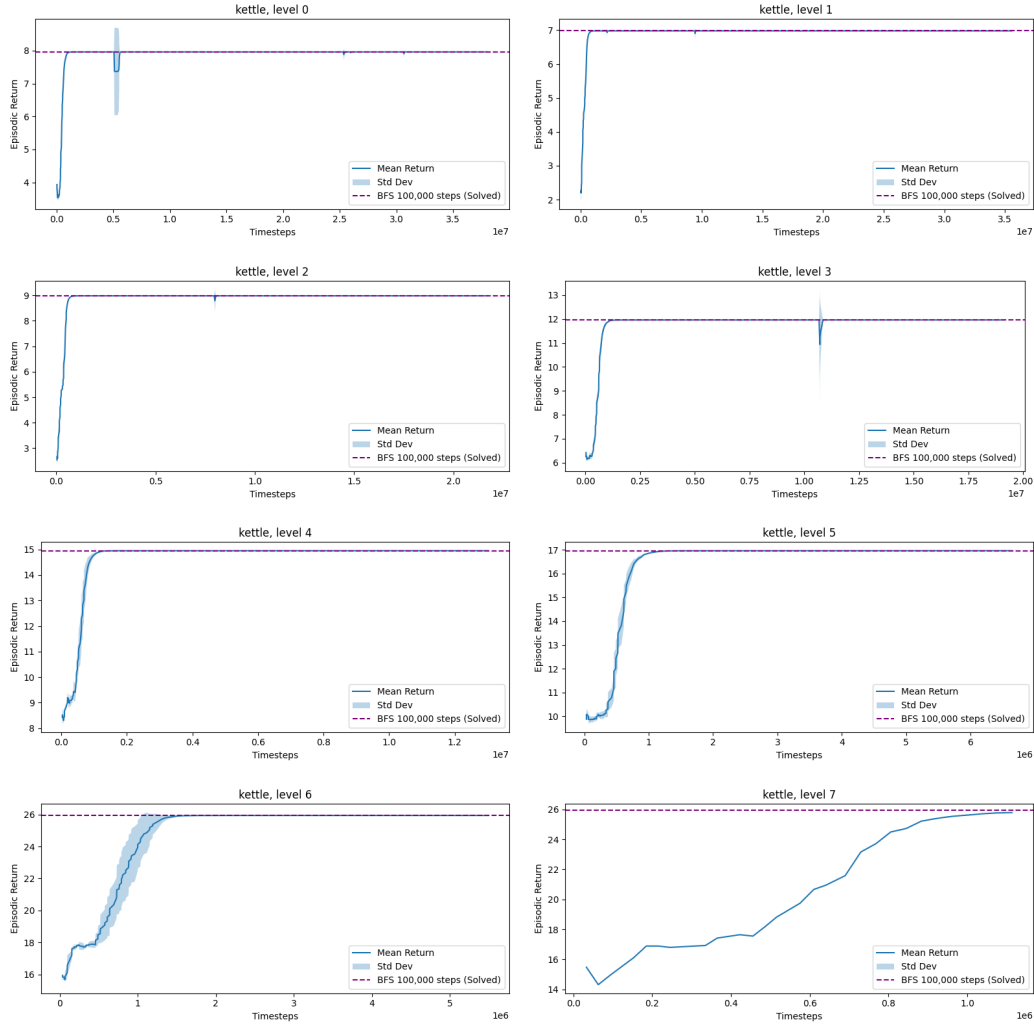


Figure 10: Comparison of RL against breadth-first search in *Kettle*. RL agents are able to find optimal solutions, which involve a short sequence of actions, though the time taken to learn this optimal strategy steadily increases as levels (and optimal action sequences) grow and complexify.

945 D.3 Tree search

946 We run breadth-first search on our full dataset of games, capping the number of environment steps
 947 at 100,000, and setting a timeout of 1 minute in [Table 5](#).

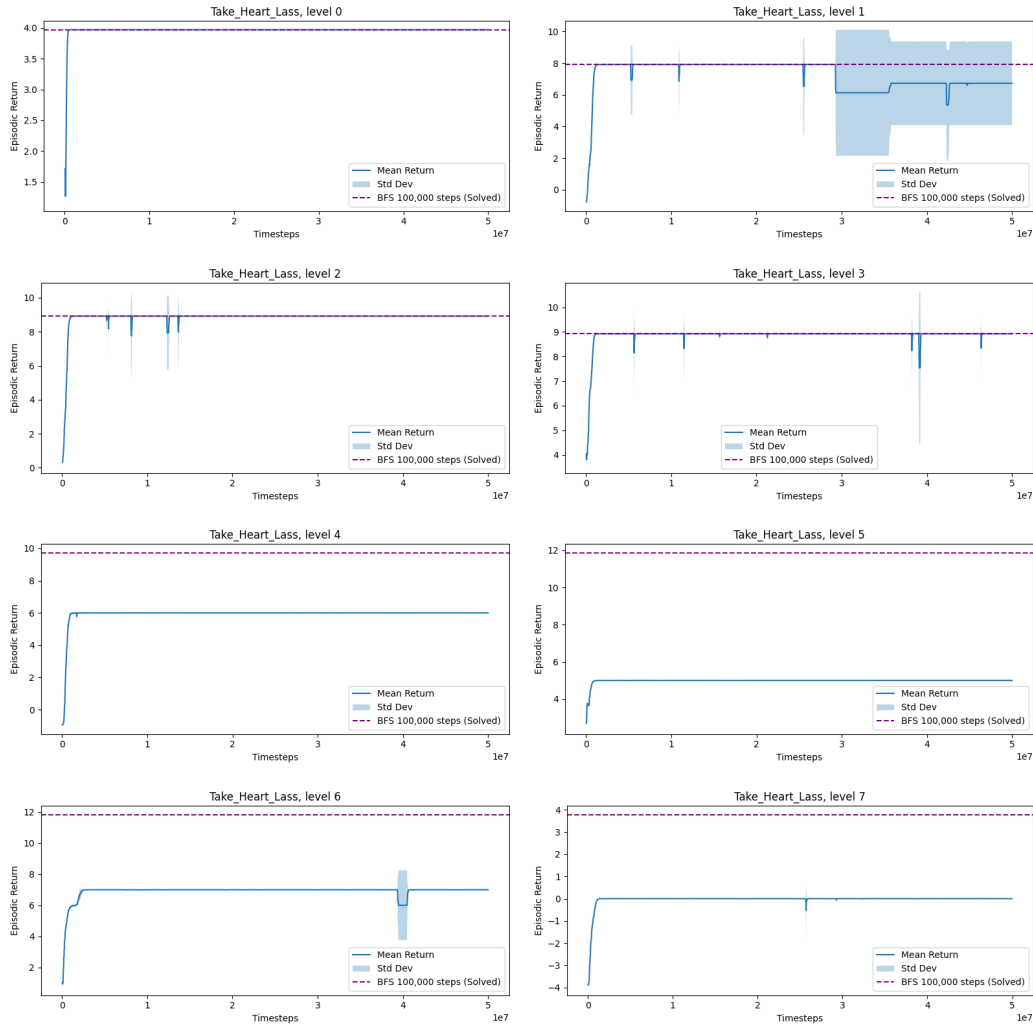


Figure 11: Comparison of RL against breadth-first search in *Take Heart Lass*. RL can handily find solutions to early levels which involve effectively evolve running away from encroaching despair and toward a goal, but it has difficulty in later levels that introduce the use of pushable hearts to strategically block the despair's advance.

Listing 1: Example of a *PuzzleScript* file (*LimeRick*)

```
title Lime Rick
author Tommi Tuovinen
homepage http://www.kissmaj7.com/

(Ported with the very gracious permission of Tommi Touvinen
The first ten levels of a neato game - you can play the full version here
http://www.kongregate.com/games/KissMaj7/lime-rick
The full version includes some mechanics that aren't covered in the levels here,
but they are supported.)

=====
OBJECTS
=====

Background
black
```

```
Exit
red
.000.
00000
00000
00000
.000.

Apple
blue
.000.
00000
00000
00000
.000.

PlayerBodyH
green
.000.
00000
0...0
00000
.000.

PlayerBodyV
green
.000.
00.00
00.00
00.00
.000.

Crate
orange
00000
0...0
0...0
0...0
00000

PlayerHead1
lightgreen
.000.
0.0.0
00000
00000
.000.

PlayerHead2
yellow
.000.
0.0.0
00000
00000
.000.

PlayerHead3
orange
.000.
0.0.0
00000
00000
.000.

PlayerHead4
```

```

red
.000.
0.0.0
00000
00000
.000.

Wall
brown

=====
LEGEND
=====

Player = PlayerHead1 or PlayerHead2 or PlayerHead3 or PlayerHead4
Obstacle = PlayerBodyH or PlayerBodyV or Wall or Crate or Player
PlayerBody = PlayerBodyH or PlayerBodyV
. = Background
P = PlayerHead1
# = Wall
E = Exit
A = Apple
C = Crate

=====
SOUNDS
=====

sfx0 3295707 (player jump)
sfx1 3538707 (player jump to max)
sfx2 42451307 (player move horizontally)
endlevel 96434300
startgame 49875902

=====
COLLISIONLAYERS
=====

Background
Exit, Apple
PlayerBody
Player, Wall, Crate

=====
RULES
=====

(this game handles all the movement stuff itself - it removes all movements before
the movement phase has a chance to tick at all)

UP [ UP PlayerHead4 ] -> [ PlayerHead4 ]
UP [ UP PlayerHead3 | No Obstacle ] -> [ PlayerBodyV | PlayerHead4 ] sfx1
UP [ UP PlayerHead2 | No Obstacle ] -> [ PlayerBodyV | PlayerHead3 ] sfx0
UP [ UP PlayerHead1 | No Obstacle ] -> [ PlayerBodyV | PlayerHead2 ] sfx0

horizontal [ > Player | Crate | No Obstacle ] ->
[ PlayerBodyH | PlayerHead1 | Crate ] sfx2

horizontal [ > Player | No Obstacle ] -> [ PlayerBodyH | PlayerHead1 ] sfx2

[ Player Apple ] [ PlayerBody ] -> [ Player Apple ] [ ]
[ Player Apple ] -> [ Player ]

[ > Player ] -> [ Player ]

```

```
DOWN [ Player | No Obstacle ] -> [ PlayerBodyV | PlayerHead1 ]
DOWN [ Crate | No Obstacle ] -> [ | Crate ]
```

```
=====
WINCONDITIONS
=====
```

```
some player on exit
```

```
=====
LEVELS
=====
```

```
message level 1 of 10
```

```
#####
#.....#
#.....#
#.....#...#
#.....#...#
#.....#...#
#..P...#...#..E.#
#####
#####
..#...#...#...#
#...#...#...#...#
#####
#####
#####
```

```
(additional levels omitted for clarity)
message congratulations!
```

Table 3: PuzzleScript games in which all levels were successfully validated in JAX (vis-a-vis solutions generated by breadth-first search in JavaScript).

Game	# Rules
Miner To Miner Empire	0
There is no-one here to help you	0
blocks	0
sokoban sanity	1
sokoban sticky	1
sumo	1
test	1
wrappingrecipe	1
Clones	1
Carnival Shooter	1
ESCAPE!	1
Futuristic Block Pushing Game	1
Carnival Shooter!	1
Bad Example	1
opposition test	1
microban	1
blank	1
Swap the block!	1
Time-reversed Microban	1
Swap Sokoban	1
Sokubunny and the colored Boxes	1
Sokoban Flipped	1
Sokoban...	1
Skeleton Assembler	1
Pulling Box Sokoban	1
Piedra	1
Opposition	1
Okosban	1
Microban I	1
Microban	1
Lis' game development adventure	1
King of walls	1
sokoban basic zoomscreen	1
sokoban basic debug	1
sokoban basic	1
soko and the band test	1
secrets test	1
perpendicular test	1
sokoban horizontal	1
sokoban eyeball	1
rigidfail1	2
sokoban match3	2
naughtysprite	2
notsnake	2
twolittlecrates1	2
actiontest	2
Crafting Table	2
Cyberpunk 2020	2
Love and Pieces	2
Puzzle-Design Homework	2
Roze	2
Silly Rabbit	2

Continued on next page

Table 3: PuzzleScript games in which all levels were successfully validated in JAX (vis-a-vis solutions generated by breadth-first search in JavaScript).

Game	# Rules
TwinPush	2
GDD301 Demo Game	2
colorban test	2
dot puzzle test	2
gdd301 test	2
lovendpieces	2
Deltarune block pushing game	2
Duplex	2
Escape the Void Full	2
Labyrinth	2
Ice Block Game	2
Come Doggo come - Alliteration	3
zenpuzzlegarden	3
Zen Puzzle Garden	3
Time-Reversed Minicosmos	3
Undertale X O Puzzle	3
Veggie Jam	3
1D Sokoban	3
Bridge	3
Maze it all	3
Snow	3
Soko-bine	3
Deadly Cat Maze	3
Footprint Puzzle	3
Harvesting Apples	3
Harvesting Apples II	3
Kill the eyes	3
scriptcross	3
rigid 11	3
The Land Of Secrets	4
Travelling salesman	4
Snake Crate	4
Remote Control Sokoban	4
RAFTURE 2 SALVATION	4
Tele	4
All These Damn Crates	4
Monster block push game	4
Multi-word Dictionary Game	4
Brush Runner	4
Bubble Boy	4
Cuddlefish	4
Handle with Care	4
Ice Maze	4
Modality	4
modality	4
wordgame	4
constellationz	4
leftrighnpcs	4
wall scout	5
TipOver	5
twolittlecrates3	5
twolittlecrates4	5
octat test	5

Continued on next page

Table 3: PuzzleScript games in which all levels were successfully validated in JAX (vis-a-vis solutions generated by breadth-first search in JavaScript).

Game	# Rules
twolittlecrates2	5
Lufia Puzzle Remake	5
Merge and Swap	5
Box Fill	5
Broken Leg Sokoban	5
Every Three Steps You Hit a Wall Out of Nowhere	5
I Only Play REAL Multiban	5
Roots	5
Poop and Salad Have Different Mindsets	5
PS Paint	5
octat	5
lunar lockout	5
Indiana Jones And His Adventure In The Lost Temple Of Owls...	6
Isle of Sokoban	6
Ocean Treasure	6
Racist Dogs	6
Power Problems	6
Pac-Puzzler	6
Impossible Sokoban	6
Bubblegoban	6
Algorithm-Generated Sokoban Levels	6
Amy	6
Bottomless puzzle-adventure	6
Tile Tiler	6
Stuffle	6
window hero	6
Eyes On You	7
Combine!	7
Sokocross	7
Son of Pacman	7
Dreaming of Strawberries	7
Shroom Party!	7
Roller Boi	7
blockfaker	7
rigid scott1	7
Escaping Limbo	8
Reap	8
Stand aside, everyone! I take large steps!	8
ZigZag Ice	8
ebony and ivory	9
The Blob	9
Ebony & Ivory	9
It Is Pitch Black	9
Mimic Translation	9
Sluzzle	9
Stickyban	9
Switcheroo	9
Absorb	9
=lost=	9
ESL Puzzle Game – CHALLENGE MODE	10
Game One	10
Kicking walls	10
Path Finder	10

Continued on next page

Table 3: PuzzleScript games in which all levels were successfully validated in JAX (vis-a-vis solutions generated by breadth-first search in JavaScript).

Game	# Rules
Skull Contraption	10
Slidings	10
slidings	10
Add Man 2 This Time It's Arithmetical	10
limerick	11
Tractor Beam Sokoban9	11
Stand II	11
Rush Hour	11
Lime Rick	11
Colorban	12
Yin-Yang	12
Jump Fox	12
Fractured Identity	13
2020	13
The Growth of H'Ragh	13
Palette	13
Variations	13
This Adventure World	13
AutoDom	13
kettle	14
Rolling-block colour-zone mazes	14
Kettle	14
Painter	14
Together Alone	14
Bichrome	15
Tempting Fate	15
Weird Bug II Weird Bugger	15
Abel's Warehouse	15
No Right Turn Sokoban	15
Mitosys	15
FS Game Dev Workshop	15
Omnipresent	15
Prototype Demo v1	16
Leo	16
Knight's Tour	16
take heart lass	17
Sticky v0.901	17
Straighten Up	17
Hysteresis mazes	17
Take Heart Lass	17
Clean Up	17
Castle Duck!	17
Darkness Sokoban	18
Solitaire	18
BUDS	18
Last day of the year	18
I Have No Mouth, And I Must Find My Hiding Spot	19
smother	19
Puzzleboi	19
Towers of Saigon	20
I Have No Mouth, And I Cannot Push Two Boxes	20
No Fault of My Own	20
Color Combination	20

Continued on next page

Table 3: PuzzleScript games in which all levels were successfully validated in JAX (vis-a-vis solutions generated by breadth-first search in JavaScript).

Game	# Rules
The Block Tower	20
Heroes of Sokoban - Ancient Japan	20
Hitori	21
Rob's first game	21
Multiply	21
Hill Buddies	21
Hanoi Winter	21
Linked	22
Lawn-Mowing Robot	22
Wrap	23
I Have No Mouth, And I Must Toggle Blocks On All Sides Of M...	23
Detroit Become Immense	23
magnetic	23
Omniban	24
candy saga 2 turbo edge	24
9x9 Go	27
Life at Home Depot	27
Meatball	27
Stick-with-it mazes	27
Stairs	27
Circulando	27
Candy's Crushes Saga	28
Light Maze	28
PackBot	30
Do You Want to Build a Snowman	30
Memory path	31
Paint	31
rigid many many	31
rigid many many2	31
Pipe Game	34
cakemonsters	34
firefighting	34
Twin Drop	35
Gears	36
Yin Yang	36
Rainbow Apples	37
Tornado Tamer	38
A port of Puzzle Wizard IQ 130+, level 18 of 72	40
Shared Bridges Game	40
Pong	41
Fabric	42
Puzzle Snake	44
RBG	51
Puzzle Stitch	58
Two-faced	67
Flippin' Ladybug	75
Upstairs Downstairs	75
At Last, A Ninja!	80

Table 4: PuzzleScript games in which one or more levels were successfully validated in JAX (vis-a-vis solutions generated by breadth-first search in JavaScript).

Game	# Rules
Minimalist	1
Password	1
the undertaking	1
zelda fan game	1
Government Final Exam	1
random test	1
Dot Puzzle 3.1	2
Mice and Cheese	2
atlas shrank test	2
the walls are closing in on me	3
Zombie Invasion	3
COIN COLLECTORS	3
Teh Interwebs	4
Greedy Explorer	4
3 brothers	4
Stand	5
Zanzlanz Adventures	5
The electrician	6
Nirvana	6
Dead Grass	6
byyourside	6
Cancel	7
Black & White	7
AVGN Dis-Ventures	7
Some lines were meant to be crossed	7
Little Girl, Big World	7
LameLightsOut	7
MazezaM	8
Nethack Sokoban	8
The Dragon's Gold	8
Sokoban Dungeon	8
mazezam	8
Dragon Quest	8
Boxes	8
The Trouble with Toasters	8
Union Move	9
Brotherhood	9
Delicious Ships!	9
Glue Factory	9
RED GREEN BLUE	9
Unity	10
Lime Richard	11
Bridge-toggle Maze	11
Circuit Breaker	11
Stand III	11
Line of Sight	11
The Dungeon of Squeamish Chickens	11
Flood	12
Hamiltwo	13
Cold Feet Sokoban	13
Rook Game	13
Rocketman	14

Continued on next page

Table 4: PuzzleScript games in which one or more levels were successfully validated in JAX (vis-a-vis solutions generated by breadth-first search in JavaScript).

Game	# Rules
Directioban	14
Electric Slide	14
NAND Circuit Simulator	14
Random Color Maze	14
Pontoluvian StraightJacket Puzzle	14
Turn-based Pac-man	14
Two-Bit's Sacraficial Adventure	14
Rewindoban	14
Pitman MZ-700	15
Sokolor EX	15
Stepping Stones	15
Midas	15
midas	15
Emerita	16
Knightoban	16
Involuntary Line Force Plus	16
The Adventures of Red and Blue!	16
Pants, Shirt, Cap	17
where did all this ice come from	17
Match Flow	18
Pushcat Jr	19
Candy Bomb	20
The Irrigation Game	20
Strong Fish	20
Grandma's Christmas Kisses	20
Nyetris DEMOv1.8	21
Hero	21
Bricks and Hammer	21
The Saga of the Candy Scroll	21
Hector	21
the saga of the candy scroll	21
Heroes of Sokoban	22
heroes of sokoban	22
Heroes of sokoban edited	22
stick candy puzzle saga	22
Bat Loves Lava Shark	23
Lights Out	23
The Nuevo Asylum	24
Coin Eater	24
Gravity Sokoban	24
Pumpkin smasher	26
Eschoban	26
Misaligned	26
Pink Matter	26
The Beginning, The End	26
Neutralize	28
dungeon janitor	28
tiny treasure hunt	28
Party Leaver 0.02	28
PuzzleScript Cell Machine Test	28
Blind Ninja	28
Dungeon Janitor	28
ledchallenge	28

Continued on next page

Table 4: PuzzleScript games in which one or more levels were successfully validated in JAX (vis-a-vis solutions generated by breadth-first search in JavaScript).

Game	# Rules
PWARE SKEG v1.1	29
push	30
PUSH	30
collapse	31
Brendan loves mondays	31
The Adventures of Pablo Diego Jose Francisco de Paula Juan N...	33
Baguettes	33
Pousser des caisses	34
Tumblebox	34
Clean your room	34
Mow Problems	34
Janitor Ted	35
Switchboard Operator	35
Heroes of Sokoban and Other Tales The Ranger	36
something	36
Boulderban	37
Sticky Cubes	38
Manic Ammo	38
Dinosaur Love Story	38
manic ammo	38
Pegs	39
Sokofun Clone	41
Crate Rotate	42
chaos wizard	42
Dinosaur Love Story Bonus Levels	43
atlas shrank	44
Atlas Shrank	44
rotate fall	45
Rabbits like Carrots	49
A Rose for Icy Heart	50
Heavy Sword	50
Frogbike Magnet Hero	51
Heroes of Sokoban II Monsters	51
12345 v1.0	52
Life is Hard	52
Bombs Away	55
Aaaah! I'm Being Attacked by a Giant Tentacle!	55
Illuminate	56
Puzzle Dash	57
IDOLS TO THE BURNT GOD	60
Miss Direction	63
Weird Bug and the Many Magic Gems	63
Let me through	64
Candy Crush Demake	65
	65
Diagonal Block Pushing Prototype	66
Two-faced	67
Elimination	69
Drop Kick	74
Smoothoban	78
Heroes of Sokoban III The Bard and The Druid	79
heroes of sokoban 3	79
Stackoban	79

Continued on next page

Table 4: PuzzleScript games in which one or more levels were successfully validated in JAX (vis-a-vis solutions generated by breadth-first search in JavaScript).

Game	# Rules
SHELL CASE	80
Feed Me Through A Straw	82
Instrumenta de Superi	84
Magnetoban	88
Snakey Jakey	93
chaos wizard ii spellcrafter	104
Broken Abacus	115
123456	117
Flower Bed	127
ParaLands	131
Infected World	133
Let's Part, E Contest Ver	138
Cooking Game	140
Wiggler demake	147
Kiri's Garden	157
Inswaption	157
Alcazar	161
STRATA-GEMS	201
Fish Friend	209
Aperture Science Sokoban Testing Initiative	226
Doors and Boxes	279

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Wizard School!	100%	1	0
Light Knight	100%	1	0
Linked	100%	7	0
Not Normal Crates	100%	18	0
A Cog's Life	100%	3	0
the gods await me	100%	1	1
IDIBAW	100%	15	3
XL-Up Maze	100%	2	4
FULL CIRCLE	100%	1	4
Elements	100%	11	5
the camel and th	100%	1	5
Rock, Paper, Scissors (v0.90 = v1.alpha)	100%	13	5
Puzzlescript Diary	100%	1	8
Find Girlfriend!	100%	15	13
Dark Maze 3	100%	1	15
Rocketman	100%	5	16
1D Sokoban	100%	2	19
Repel	100%	4	23
Wrap	100%	7	23
Hungry Kitty	100%	10	26
nekopuzzle	100%	10	26
Slidyyyyyyy	100%	5	27
Mirror Bounce	100%	4	34
Impetus	100%	5	37
Floor Painter	100%	1	38
Slippy Penguin	100%	7	39
Castle Monk	100%	2	44
Minimalist	100%	10	46
Silly Rabbit	100%	4	48
Tumblebox	100%	4	51
GDD301 Game	100%	10	51
LameLightsOut	100%	4	52
something	100%	1	60
Pilgrimage	100%	2	61
Pumpkin smasher	100%	8	62
Yin and Yang	100%	3	62
Four-room tilt mazes	100%	3	63
How Many Digits Of Pi	100%	21	65
Do You Know	100%	7	66
Rocketmen	100%	7	66
Roller Boi	100%	6	76
HyperMaze	100%	1	83
Fruit Ghost	100%	1	88
Forever Hallway	100%	2	103
Rolling-block colour- zone mazes	100%	9	117
STUCK IN A HOLE	100%	17	121
Hysteresis mazes	100%	5	168
Inchworm	100%	1	174
Tumblin'	100%	6	192

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
CC2	100%	1	204
Random Color Maze	100%	8	217
Stick-with-it mazes	100%	5	220
Absorb	100%	2	302
Game 101	100%	7	367
Puzzle Game Supreme	100%	4	384
With Cheese	100%		
Memories Of Castle-	100%	10	396
mouse			
castlemouse	100%	10	396
Animal Cascade	100%	10	396
Eyes On You	100%	2	442
COIN COLLECTORS	100%	10	452
Ouroboros	100%	6	478
Saute Mouton	100%	1	484
Roots	100%	5	509
Boats Cars & Trains	100%	9	510
Impasse	100%	23	529
Drop Maze	100%	5	570
Space Shoot-em-puzzle	100%	10	592
Mimic Translation	100%	7	629
LA CASE DE CANIMI	100%	10	686
YE PAPEL			
Rewindoban	100%	5	812
Darkness Sokoban	100%	2	877
sokoban basic	100%	2	900
sokoban basic zoom-	100%	2	900
screen			
sokoban sanity	100%	3	900
sokoban basic debug	100%	2	900
Stuffle	100%	2	900
RBG	100%	1	930
Shield Game	100%	3	973
Kill the eyes	100%	1	974
Loup-Mouton-Chou	100%	1	1127
PWARE SKEG v1.1	100%	6	1256
Fishman	100%	4	1278
It Is Pitch Black	100%	5	1331
Brain	100%	1	1353
Racist Dogs	100%	3	1355
Black & White	100%	3	1381
Blind Ninja	100%	9	1481
Piedra	100%	4	1489
Towers of Hanoi	100%	4	1553
Metamorph	100%	4	1562
Crate Rotate	100%	11	1583
Path Finder	100%	1	1589
Detroit Become Im-	100%	1	1606
mense			
Clean Up	100%	8	1667
Line of Sight	100%	1	1741
m c eschers armageddon	100%	4	1921

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
MC Escher's Equestrian	100%	4	1921
Armageddon	100%	5	2164
Deserted Island	100%	1	2272
Ebony & Ivory	100%	1	2272
ebony and ivory	100%	26	2351
Slide Rule	100%	3	2472
Veggie Jam	100%	10	2779
byyourside	100%	1	2909
Five Pulloban Puzzles	100%	1	3246
Hill Buddies	100%	1	3407
Tricky Tower	100%	6	3863
Game One	100%	11	3882
Mad Queens	100%	1	4417
Vines	100%	1	4508
Shroom Party!	100%	11	4606
TipOver	100%	1	4771
Deltarune block pushing game	100%	6	4831
Magical Shoe Adventure	100%	6	4988
icecrates	100%	6	4988
IceCrates	100%	4	5013
CrateBlob	100%	6	5063
Duplex	100%	5	5296
Crate Game Five	100%	9	5297
Sokoban...	100%	4	5371
Modality	100%	4	5371
modality	100%	5	5587
Little Girl, Big World	100%	10	5712
naborciM	100%	10	5712
Time-reversed Microban	100%	3	5786
SSR Demake	100%	4	6426
Rob's first game	100%	2	6569
Connect Four	100%	6	6710
Abel's Warehouse	100%	31	7101
Bubble Butler CMD RE-ORGANIZE	100%	7	7258
silver lungs	100%	4	7571
magnetic	100%	1	7671
Savior	100%	8	7878
Switcheroo	100%	5	7946
Cyberpunk 2020	100%	4	8770
Laser	100%	1	9063
Sponge Game	100%	1	9382
A port of Puzzle Wizard	100%	1	10372
IQ 130+, level 18 of 72	100%	14	10670
Pac-Puzzler	100%	2	11620
Strange Warehouse	100%	11	12189
sokoban match3	100%	11	12189
slidings	100%	23	12375
Slidings	100%		
Sokubunny and the colored Boxes	100%		

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Pulling Box Sokoban	100%	10	12813
Stickyban	100%	7	14038
The Observable Uni-verse	100%	7	14706
Multi-word Dictionary Game	100%	1	15875
wordgame	100%	1	15875
octat	100%	8	16009
Marble Shoot	100%	11	16243
Michael Bay's Legend of Zelda	100%	3	16377
Stairways	100%	3	16942
Line Force	100%	1	17338
candy saga 2 turbo edge	100%	8	17544
Tele	100%	7	17585
Coin Eater	100%	7	17672
Spring	100%	3	18160
RAFTURE 2 SALVATION	100%	9	21102
Shooting Pool	100%	6	21536
Escape the Void Full	100%	17	21803
Not in the mood	100%	1	24884
Nom	100%	1	28142
Back home	100%	13	29166
Towerfall	100%	1	29462
Dangerous Dungeon	100%	10	29769
lunar lockout	100%	4	30351
Game Loading	100%	3	30552
The McKelvey-Schofield Theorem	100%	2	32827
Ocean Treasure	100%	12	35663
Party Demon	100%	3	36069
Kettle	100%	11	36298
kettle	100%	11	36298
Possibility Space	100%	3	38034
Spacekoban	100%	20	38599
Sokoban Dungeon	100%	3	41174
Spider's Hollow	100%	11	44385
Glue Factory	100%	4	45418
Sokobot	100%	2	45673
Backstabber	100%	6	48770
Bat Loves Lava Shark	100%	3	51127
Opposition	100%	14	53575
No Right Turn Sokoban	100%	4	57140
ESCAPE!	100%	5	58846
Graded Sir	100%	5	59089
dreizack	100%	2	63708
Bottomless puzzle-adventure	100%	2	64133
Time-Reversed Minicosmos	100%	40	65958
Parity Match	100%	12	67572
Swap Sokoban	100%	10	74488

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
The Dungeon of Squeamish Chickens	100%	12	76510
Son of Pacman	100%	4	82138
Logger Isles	100%	1	86503
Nirvana	100%	4	91325
Zombie Pitfall	100%	3	94736
Zombie Invasion	100%	4	96570
Philosophy	98%	113	100000
Blind Maze a1	96%	25	100000
Pig Friends	95%	20	8
Pivotal	95%	20	100000
Forest of Tizenhat	94%	18	100000
smother	94%	16	100000
Smother	94%	16	100000
The World Beneath the Surface	94%	16	100000
Easy Enigma	94%	16	100000
AVGN Dis-Ventures	93%	30	100000
Power Block	93%	15	100000
Zanzlanz Adventures	93%	15	100000
Fast Forward	93%	15	100000
sokobond demake	93%	15	100000
Velocity Castle	93%	14	27266
Snow	92%	13	100000
Involuntary Line Force Plus	92%	13	100000
sum three horizontally to 8!	92%	13	100000
Travelling salesman	92%	12	100000
pikkuruinen	92%	12	100000
Colour Chained	92%	12	100000
color chained	92%	12	100000
Unity	92%	12	100000
Take Heart Lass	92%	12	100000
take heart lass	92%	12	100000
Colorban	91%	23	100000
Drop Kick	91%	34	100000
coincounter	91%	11	11042
Coin Counter	91%	11	11042
Ball smashes blocks	91%	11	100000
Labyrinth	91%	11	100000
rotate fall	90%	20	184
Lime Richard	90%	10	100000
Sokoban Flipped	90%	10	100000
Heavy Sword	90%	10	100000
12345 v1.0	90%	30	100000
MazezaM	90%	30	100000
mazezam	90%	30	100000
Algorithm-Generated Sokoban Levels	90%	20	100000
the undertaking	89%	38	70
Escaping Limbo	89%	9	10493

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
lovendpieces	89%	9	36029
Love and Pieces	89%	9	36029
Pants, Shirt, Cap	89%	9	100000
This Adventure World	89%	9	100000
Pac-Slide V0.3	89%	9	100000
Candy's Crushes Saga	88%	26	100000
Beams and Flowers	88%	17	100000
The Growth of H'Ragh	88%	8	2552
Coin Dropper	88%	24	100000
Miner To Miner Empire	88%	8	100000
Isle of Sokoban	88%	8	100000
Soko-bine	88%	8	100000
Stealth Agent Gemini	87%	15	50846
Mice and Cheese	87%	15	100000
Ice Cubes	86%	29	100000
Sweet Hints	86%	7	100000
Mars Attacks	86%	7	100000
push	86%	21	100000
PUSH	86%	21	100000
3 brothers	86%	14	100000
EpicJamGame	86%	7	100000
Stepping Stones	86%	7	100000
Stairs	86%	7	100000
Miss Direction	85%	20	100000
Angize	85%	13	100000
Dot Puzzle 3.1	84%	19	100000
easyenigma	84%	19	100000
Schleimban	83%	6	13478
Eschoban	83%	6	24838
Party Leaver 0.02	83%	6	100000
PackBot	83%	12	100000
Insey Winsey Spider	83%	6	100000
Combine!	83%	12	100000
The electrician	83%	6	100000
sokodig	83%	18	100000
Boupha's Candle Quest	83%	12	100000
Prototype Demo v1	83%	6	100000
Magnet Jack	83%	6	100000
Manni the Mountaineer	83%	6	100000
Closure Demake	83%	6	100000
Jailbreak	83%	6	100000
Soliquid	82%	11	100000
Stand II	82%	11	100000
Star Burger Inc	81%	16	100000
Doktor Lezer	81%	16	100000
cakemonsters	81%	36	100000
Lord of darkness	80%	10	1649
GDD301 Demo Game	80%	5	2971
Subterranean Maze of Caves	80%	5	16476
Transparency	80%	10	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Xtreme Gardening	80%	10	100000
Palette	80%	15	100000
Misaligned	80%	10	100000
Gravity Sokoban	80%	10	100000
Routing Puzzle	80%	10	100000
Aaaah! I'm Being At- tacked by a Giant Tenta- cle!	80%	10	100000
Hanoi Winter	80%	5	100000
Stand III	80%	15	100000
Stand	80%	5	100000
Color Combination	80%	5	100000
The Blob	80%	5	100000
Castle Duck!	80%	5	100000
Single Map Symbol Mi- croban	80%	10	100000
ZigZag Ice	80%	5	100000
Microban	80%	10	100000
microban	80%	10	100000
VEXT EDIT	80%	5	100000
Rippler's Tower	80%	25	100000
Union Move	79%	14	99590
High Noon RPG game	79%	14	100000
Enqueue	79%	14	100000
Boxes	78%	9	100000
Unfinished Arrows	78%	9	100000
Game Flood	78%	9	100000
Lis' game development	78%	9	100000
adventure	77%	13	100000
Fabric	77%	13	100000
Summer	76%	17	100000
Puzzle Script Tutorial	76%	21	100000
pusH	75%	4	38
wall scout	75%	4	1060
PuzzleScript Cell Ma- chine Test	75%	4	100000
Flying Kick	75%	16	100000
Instrumenta de Superi	75%	8	100000
L.A.S.E.R	75%	20	100000
it dies in the light	75%	4	100000
Memory path	75%	12	100000
Janitor Ted	75%	12	100000
Crafting Table	75%	4	100000
Merge and Swap	75%	8	100000
whaleworld	75%	8	100000
TimeStopp	75%	8	100000
Together Alone	75%	4	100000
Box Fill	75%	8	100000
It Dies In The Light	75%	4	100000
Jesus Respawns	75%	12	100000
You're Toast!	75%	4	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
2D Whale World	75%	8	100000
Skull Contraption	75%	4	100000
diesinthelight	75%	4	100000
Tile Tiler	75%	4	100000
Hey you! Stop blocking the laser!	74%	23	100000
Unconventional Guns	73%	15	100000
Manic Ammo	73%	15	100000
manic ammo	73%	15	100000
Color Fill Land	73%	11	7702
Dinosaur Love Story	73%	22	100000
Don't Play On the Ice	73%	11	100000
constellationz	73%	11	100000
Heroes of Sokoban and Other Tales The Rogue	72%	18	100000
Warping Through the Halls	71%	7	10486
Two-Bit's Sacraficial Adventure	71%	14	100000
where did all this ice come from	71%	14	100000
Sokofun Clone	71%	7	100000
Train	71%	7	100000
Let me through	71%	7	100000
Yin Yang	71%	14	100000
Alley	71%	7	100000
Rook Game	71%	7	100000
BulletDodger	71%	24	100000
Stackoban	70%	30	100000
Rainbow Apples	70%	10	100000
zelda fan game	70%	10	100000
Neutralize	70%	10	100000
Touchdown Heroes (Prototype)	70%	10	100000
I'm Sick Today	69%	13	100000
One Way Street	67%	3	100000
Goblin Hooblob	67%	6	100000
Mitosys	67%	15	100000
Clean your room	67%	12	100000
Reap	67%	6	100000
Knightoban	67%	6	100000
Cold Feet Sokoban	67%	3	100000
Infected World	67%	21	100000
Cave Story Demake v2.3	67%	3	100000
Delicious Ships!	67%	15	100000
Don't Die	67%	6	100000
Heroes of Sokoban - Ancient Japan	67%	6	100000
Remote Control	67%	3	100000
Sokoban	67%	9	100000
Pink Matter	67%	9	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Indiana Jones And His Adventure In The Lost Temple...	67%	3	100000
Add Man 2 This Time	67%	15	100000
It's Arithmetical	67%	12	100000
Dotsnake	67%	3	100000
Escape The Haunted Mansion	67%	3	100000
Collapsable Sokoban	67%	12	100000
Lieutenant Americium	67%	3	100000
Oxygen Not Required	67%	6	100000
Teh Interwebs	67%	6	100000
Gingham	67%	3	100000
Electric Slide	67%	3	100000
Weird Dave	67%	9	100000
Broken Leg Sokoban	67%	101	100000
Catrap	66%	20	100000
Winter	65%	17	100000
Bricks and Hammer	65%	14	100000
Candy Bomb	64%	14	100000
Wiggler demake	64%	14	100000
Sticky v0.901	64%	11	100000
Fractured Identity	64%	22	100000
heroes of sokoban	64%	22	100000
Heroes of Sokoban	64%	19	100000
Hi, Vmind!	63%	8	100000
Microbe Destroyers	62%	32	100000
Snakey Jakey	62%	8	100000
Dead Grass	62%	8	100000
VRPS	62%	49	100000
Sluzzle	61%	5	16651
shoving	60%	10	100000
GDD301 Lab 010	60%	5	100000
The Trouble with Toasters	60%	10	100000
Al's Egg Roll	60%	20	100000
Autumn	60%	5	100000
blockfaker	60%	5	100000
Poop and Salad Have Different Mindsets	60%	5	100000
The Nuevo Asylum	60%	10	100000
Tractor Beam Sokoban9	60%	10	100000
Ultimate Sokoban	60%	5	100000
Supreme	60%	5	100000
Kicking walls	60%	5	100000
BUDS	60%	5	100000
2020	60%	10	100000
Doors and Boxes	60%	5	100000
Dating Queens	60%	5	100000
Mini Nomerads 2 Constructaballs	60%	10	100000
Candy Crush Demake	60%		

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Clones	60%	10	100000
Heroes of Sokoban and Other Tales The Ranger	59%	22	100000
Magnetron	59%	22	100000
Cyberbox	59%	17	100000
Ice and Tar	58%	12	100000
Atlas Shrank	58%	12	100000
The Explorer 3.0	58%	12	100000
atlas shrank	58%	12	100000
	58%	19	100000
Omniban	58%	19	100000
Bubblegoban	57%	7	768
Bridge-toggle Maze	57%	7	100000
Yin-Yang	57%	21	100000
Push - A - Crate	57%	14	100000
Hazard Golf	57%	7	100000
Mini Nomerads	56%	9	100000
Santa's Great Escape	56%	9	100000
Two-faced	56%	9	100000
Snek	56%	9	100000
heroes of sokoban 2	55%	22	100000
Everything Antimatters	55%	11	100000
Sokocross	55%	22	100000
A Knight's Tour	55%	11	100000
Heroes of Sokoban II	55%	22	100000
Monsters			
There is no-one here to help you	55%	11	100000
A Rose for Icy Heart	55%	11	100000
123456	54%	28	100000
Brotherhood	53%	15	100000
Inswaption	52%	21	100000
Pitman MZ-700	52%	50	100000
sokoban horizontal	50%	2	315
gdd301 test	50%	2	613
sokoban sticky	50%	2	13082
lava flow	50%	2	44619
Undertale X O Puzzle	50%	2	100000
niggle	50%	4	100000
Lights Out	50%	10	100000
Shared Bridges Game	50%	8	100000
Hack It!	50%	2	100000
1D Rubik's Cube	50%	2	100000
Beam Islands	50%	8	100000
Pushller	50%	4	100000
Collapse	50%	2	100000
collapse	50%	2	100000
Flipush An Exciting Square Game	50%	6	100000
Every Three Steps			
You Hit a Wall Out of Nowhere	50%	2	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Snakeoban	50%	10	100000
A Knight's Path	50%	2	100000
Antimatter Labs	50%	12	100000
Ball Bros	50%	10	100000
Twinkle, Twinkle Little Square	50%	8	100000
Out There	50%	8	100000
Directioban	50%	6	100000
Match Pub	50%	10	100000
Cancel	50%	8	100000
Footprint Puzzle	50%	4	100000
Suplexer	50%	4	100000
Maze it all	50%	4	100000
SHELL CASE	50%	6	100000
Switchboard Operator	50%	6	100000
Strong Fish	50%	18	100000
FROWN INVERSION SQUAD	50%	20	100000
Juxtaposition	50%	16	100000
SHY SNAKE	50%	2	100000
Festive Lights	50%	10	100000
Stand Off	50%	10	100000
Bridge	50%	4	100000
Multiply	50%	6	100000
Kick the Can	50%	4	100000
Frogbike Magnet Hero	50%	2	100000
Hamiltwo	50%	2	100000
The Circulatory System	50%	24	100000
Global Office - CSGO Emerita	50%	16	100000
Brendan loves mondays	50%	10	100000
Flower Bed	47%	17	100000
Sokobaiogenesis	46%	13	100000
Illuminate	45%	31	100000
Microban I	45%	155	100000
Baguettes	44%	9	100000
Filler	44%	18	100000
Pontoluvian Straight-Jacket Puzzle Variations	44%	9	100000
Hero	44%	25	100000
Bit Blocks	43%	58	100000
Cops N Robbers	43%	7	100000
Match Flow	43%	14	100000
Dreaming of Strawberries	43%	7	100000
Omnipresent	43%	7	100000
Crates and Portals	41%	29	100000
cratopia	41%	17	100000
Cratopia	41%	17	100000
Zombie Rescue	40%	15	10905
Sokolor EX	40%	5	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Simple Block Plagarism	40%	5	100000
Game			
Circuit Breaker	40%	5	100000
Dragon Quest	40%	10	100000
I Only Play REAL Multi-	40%	5	100000
ban			
I Have No Mouth, And			
I Must Find My Hiding	40%	20	100000
Spot			
Heroes of sokoban	40%	25	100000
edited			
Grandma's Christmas	40%	10	100000
Kisses			
Monster block push	40%	10	100000
game			
Swap the block!	40%	5	100000
Drop Swap	40%	5	100000
Keys and Doors 0.1.0	40%	5	100000
Explod	40%	10	100000
midas	40%	15	100000
Midas	40%	15	100000
the walls are closing in	40%	10	100000
on me			
Alcazar	40%	10	100000
I Have No Mouth, And I			
Must Toggle Blocks On	40%	10	100000
All ...			
Puzzleboi	40%	5	100000
The Adventures of Pablo			
Diego Jose Francisco de	39%	18	100000
Pa...			
Godzilla	38%	8	100000
Atomix demake	38%	8	100000
Broken Abacus	38%	16	100000
-ooo-	38%	8	100000
The Great Emu War	38%	24	100000
NAND Circuit Simulator	38%	8	100000
Number Friends -work	37%	27	100000
in progress-			
the saga of the candy	37%	30	100000
scroll			
The Saga of the Candy	37%	30	100000
Scroll			
Reset	36%	11	245
The Workshop	36%	11	100000
PUZZLETALE	36%	11	100000
heroes of sokoban 3	36%	22	100000
Heroes of Sokoban III	36%	22	100000
The Bard and The Druid			
Newton's Crates	36%	14	100000
Medusa's Wrath v0.2.1	36%	14	100000
Miner Guy	33%	3	100000
Pipe Game	33%	3	100000
Four Colour Theorem	33%	3	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Feed Me Through A Straw	33%	27	100000
Ouzo Pixel Party	33%	9	100000
Bad Example	33%	3	100000
GASP	33%	3	100000
Knight's Tour	33%	6	100000
sokochrome	33%	9	100000
Pousser des caisses	33%	3	100000
Memory Push	33%	9	100000
Circulando	33%	3	100000
Dinosaur Love Story	33%	6	100000
Bonus Levels			
Life at Home Depot	33%	9	100000
Diagonal Block Pushing	33%	3	100000
Prototype			
Straighten Up	33%	6	100000
Season Finale	33%	9	100000
Tidy the Cafe!	33%	6	100000
At Last, A Ninja!	33%	3	100000
Tornado Tamer	33%	6	100000
Power Problems	31%	16	100000
limerick	30%	10	100000
Let's Part, E Contest Ver	30%	10	100000
The Adventures of Red			
and Blue!	30%	10	100000
Aerobatics	30%	10	100000
Gobble Rush!	29%	21	464
gobble rush	29%	21	464
ParaLands	29%	7	100000
Muddy Sokoban Level			
Set I	29%	7	100000
KidsCoding	29%	7	100000
Diamond Mine	29%	7	100000
Sokoslam	28%	18	100000
Boulder Puzzle	28%	18	100000
Sticky Cubes	27%	11	100000
Lime Rick	27%	11	100000
IDOLS TO THE	27%	26	100000
BURNT GOD			
EntrepotPhage Demake	27%	30	100000
STRATA-GEMS	27%	30	100000
Mow Problems	27%	15	100000
Nethack Sokoban	26%	19	100000
Stained Glass	25%	4	100000
dungeon janitor	25%	4	100000
Bichrome	25%	12	100000
ScriptCross	25%	4	100000
scriptcross	25%	4	100000
i am two	25%	8	100000
TwinPush	25%	4	100000
Akari 2	25%	4	100000
N Queens Puzzle	25%	8	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Dungeon Janitor	25%	4	100000
Towers of Saigon	25%	48	100000
Elementals	25%	4	100000
Space Expedition	25%	8	100000
Ice Maze	25%	4	100000
Darn gifs for my game	25%	4	100000
Magnetoban	25%	12	100000
Elimination	25%	20	100000
The Land Of Secrets	25%	4	100000
XL-Plan	25%	4	100000
The Dragon's Gold	25%	12	100000
Aperture Science			
Sokoban Testing Initia-	24%	17	100000
tive			
Tour de Four	23%	13	100000
FACED	23%	26	100000
Fox Cave v0.32 (this ver-	22%	9	100000
sion added Demon)			
Smoothoban	22%	45	100000
Collect Gnocchi	21%	19	100000
The Adventure of Puz-	20%	5	100000
zler			
Pornography for Begin-	20%	5	100000
ners			
Stage Floor Assembler	20%	15	100000
Pegs	20%	15	100000
Surround Yourself With	20%	5	100000
Dogs			
The Irrigation Game	20%	10	100000
Robot Repairs 1.2	20%	20	100000
ESL Puzzle Game -	20%	10	100000
CHALLENGE MODE			
Slime Squisher	19%	16	100000
Boxes & Balloons	18%	22	100000
Okosban	18%	11	100000
atlas shrank test	17%	12	211
Flippin' Ladybug	17%	6	100000
Crocodiles Love Cook-	17%	12	100000
ies			
Sheep	17%	12	100000
The Beginning, The End	17%	12	100000
SwapBot	17%	12	100000
-=lost=-	17%	6	100000
Nyetris DEMOv1.8	17%	12	100000
Fish Friend	16%	25	98211
Leo	15%	13	100000
Some lines were meant	14%	7	100000
to be crossed			
IMS445 Puzzlescript	14%	7	100000
Game Buoy Deploy	14%	7	100000
Cooking Game	14%	7	100000
Fall Leaves	14%	7	100000
Fallow Land	14%	14	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Repeater	14%	7	100000
Turn-based Pac-man	12%	8	100000
Telefrag	12%	8	100000
Solitaire	12%	8	100000
PrograMaze	12%	8	100000
Hector	11%	9	100000
Color Totems	11%	18	100000
Boolean Bloom 0.37	11%	9	100000
stick candy puzzle saga	10%	20	100000
JAM3 Game	10%	10	100000
Upstairs Downstairs	10%	20	100000
Greedy Explorer	10%	10	100000
Gloves and Boots	9%	11	100000
Snake Crate	9%	11	100000
Do You Want to Build a Snowman	8%	12	100000
Hitori	7%	15	100000
Weird Bug and the Many	6%	17	100000
Magic Gems			
Rabbits like Carrots	5%	20	100000
chaos wizard	5%	22	100000
Puzzle Snake	0%	1	1
Pushing It	0%	1	1
Impostor crated!	0%	1	1
againexample	0%	1	2
Metal and Grass	0%	8	4
Runner 2	0%	1	6
Crypt of the Necro-dancer DEMAKE	0%	1	6
Gamlet v16.09.04 Alpha	0%	5	6
Channel	0%	1	11
Scale the Tower	0%	1	12
sokoban empty	0%	2	16
test	0%	1	24
wrappingrecipe	0%	1	25
opposition test	0%	1	32
secrets test	0%	1	32
soko and the band test	0%	1	48
colorban test	0%	1	51
Helping	0%	4	56
sumo	0%	1	60
Picky Leaks	0%	15	72
blank	0%	2	81
Dungeon Walk II	0%	1	120
Stephen's Delight			
Tic-Fa-No	0%	1	121
rigid 11	0%	1	127
Pong	0%	1	169
Switch It Up	0%	1	246
Jump Fox	0%	1	303
Carnival Shooter!	0%	1	392

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Carnival Shooter	0%	1	392
sokoban eyeball	0%	1	506
eyeball	0%	1	506
rigidfail1	0%	1	670
Sleeping Bear	0%	1	770
leftrighnps	0%	1	888
Depth-First Maze	0%	1	996
Soko and the Band	0%	1	1096
naughtysprite	0%	1	1190
Zaap	0%	1	1280
Getting ready for bed.	0%	1	1688
perpendicular test	0%	1	3906
Color Slider	0%	5	5723
mazetest	0%	1	6001
octat test	0%	1	10040
limerick test	0%	1	14268
Uphill Pathfinder	0%	1	15578
Stand aside, everyone! I take large steps!	0%	1	38377
Wormy	0%	1	59979
Puzzle Dash	0%	5	100000
Dungeon Generator V - Key Hunt edition	0%	1	100000
Paul Bunyan's Demade Island Adventure (v2)	0%	1	100000
Bombs Away	0%	1	100000
Quantum Childminding - Expanded Levelset	0%	8	100000
tiny treasure hunt 1.1	0%	7	100000
Trains Love Snowmen	0%	1	100000
Life is Hard	0%	1	100000
threes	0%	1	100000
Blow the ship down!	0%	1	100000
Dharma Dojo demake	0%	5	100000
Painter	0%	1	100000
RED GREEN BLUE	0%	2	100000
Noleap	0%	1	100000
Cuddlefish	0%	1	100000
Amy	0%	8	100000
Brush Runner	0%	1	100000
randomrobots	0%	1	100000
Treasure Hunt	0%	2	100000
Turing Machine	0%	1	100000
spooki	0%	1	100000
Freeform Bal Ru's Curse	0%	1	100000
Sokoboros	0%	1	100000
9x9 Go	0%	1	100000
Knighted chtp 2	0%	1	100000
tunnel rat	0%	2	100000
Colors	0%	1	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
A desecration of Jack	0%	1	100000
Lance's I'm too far gone	0%	1	100000
Sixty-Five	0%	1	100000
Pushcat Jr	0%	8	100000
Ogres Are Like Onions	0%	1	100000
ironic game	0%	1	100000
mijn puzzelspel	0%	3	100000
3 in a row	0%	1	100000
Boulderban	0%	2	100000
rigid many many2	0%	1	100000
rigid many broken	0%	2	100000
Bal Ru's Curse cruelty-free demake	0%	1	100000
Roze	0%	1	100000
pretender to the crown	0%	16	100000
tiny treasure hunt	0%	1	100000
rigid scott1	0%	1	100000
firefighting	0%	1	100000
riverpuzzle	0%	1	100000
Pipe Navigation	0%	1	100000
Last day of the year	0%	1	100000
Quantum Childminding	0%	4	100000
Club Penguin	0%	1	100000
Deadly Cat Maze	0%	1	100000
Princess of Isometria v0.10	0%	8	100000
Overreaction	0%	1	100000
Minecraft	0%	1	100000
Zoop Demake	0%	1	100000
castlecloset	0%	1	100000
Harvesting Apples II	0%	1	100000
Touhouban - Beta	0%	6	100000
Snowman Decorator	0%	1	100000
Little Glube	0%	1	100000
Gears	0%	1	100000
Bug Exterminator	0%	5	100000
Fi & BOXES!!!	0%	1	100000
zokoban	0%	1	100000
I'm too far gone	0%	1	100000
Impossible Sokoban	0%	1	100000
Lufia Puzzle Remake	0%	1	100000
Tiny Farm 0.1	0%	3	100000
Paint	0%	6	100000
Little Dungeons	0%	6	100000
Tempting Fate	0%	1	100000
Shoop	0%	1	100000
Skeleton Assembler	0%	1	100000
Cave Explorer	0%	1	100000
Cannon Fodder	0%	1	100000
Snake Scout	0%	1	100000
King of walls	0%	1	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
The Dragon's Cave	0%	2	100000
Grunt work of the revolution puzzles	0%	9	100000
Zen Puzzle Garden	0%	1	100000
Come Doggo come - Al-literation	0%	5	100000
MicroChip	0%	1	100000
PS Paint	0%	3	100000
Government Final Exam	0%	1	100000
Twin Drop	0%	5	100000
dungeons	0%	1	100000
gunslinger	0%	2	100000
Simple Block Pushing Game	0%	5	100000
PENTAGO	0%	1	100000
actiontest	0%	1	100000
i herd u liek water templs	0%	1	100000
Lawn-Mowing Robot	0%	7	100000
Drainage	0%	9	100000
Laserverse	0%	1	100000
zenpuzzlegarden	0%	11	100000
Pipes	0%	5	100000
notsnake	0%	3	100000
window hero	0%	1	100000
Light Maze	0%	9	100000
dot puzzle test	0%	3	100000
MATCH 3 PUZZLE	0%	1	100000
Net	0%	1	100000
legend of zokoban	0%	1	100000
Meatball	0%	1	100000
Puzzle Stitch	0%	2	100000
rigid parallel many	0%	1	100000
W3RDS	0%	1	100000
Dash Boulder	0%	1	100000
FS Game Dev Workshop	0%	5	100000
Luxbend II Marbles	0%	1	100000
Hack It! - Tilt Game	0%	1	100000
ColorLift	0%	1	100000
Rush Hour	0%	1	100000
The Lucid Dreamer	0%	1	100000
Give me ur key	0%	4	100000
bouncers	0%	1	100000
Bouncers	0%	5	100000
Bubble Boy	0%	5	100000
Puzzle Defender	0%	3	100000
Colorblend	0%	1	100000
chaos wizard ii spell-crafter	0%	1	100000
MatchLess 0.2	0%	4	100000
Weird Bug II Weird Bugger	0%	1	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
Count Mover	0%	1	100000
Ice Block Game	0%	1	100000
rigid one unlimited	0%	1	100000
A Unexamined Life Is Not Worth Meowing	0%	1	100000
Un-Lock	0%	1	100000
All These Damn Crates	0%	4	100000
PuzzleTech 0.73	0%	1	100000
Minsweeper Demake	0%	1	100000
Flash Point Fire Rescue	0%	1	100000
Two level puzzle	0%	3	100000
Attack of the Martian Giants	0%	7	100000
Snake	0%	1	100000
SNAQUIA	0%	1	100000
Colour Chained Endless	0%	1	100000
Pixel Master	0%	8	100000
Flower Dance	0%	1	100000
The Block Tower	0%	1	100000
Password	0%	3	100000
AutoTrom	0%	1	100000
The Riddle Machine	0%	5	100000
Paint2	0%	1	100000
Puzzle-Design Home-work	0%	5	100000
Snakeskin	0%	8	100000
AutoDom	0%	1	100000
twolittlecrates4	0%	1	100000
Ancient Puzzles of Colossal Caves indev4	0%	8	100000
Bloxyd	0%	24	100000
twolittlecrates3	0%	1	100000
Futuristic Block Pushing Game	0%	1	100000
ledchallenge	0%	4	100000
N Step Punt	0%	2	100000
Thirty-Two Nanobots	0%	1	100000
Tunnel Rat	0%	2	100000
Fireproof bomber	0%	5	100000
Polyomino Puzzles	0%	3	100000
Good Example	0%	1	100000
blocks	0%	1	100000
Kiri's Garden	0%	7	100000
Dull Pacman Clone	0%	1	100000
Spider Solitaire	0%	1	100000
Battle-Mage	0%	1	100000
random test	0%	6	100000
dropswap	0%	5	100000
No Fault of My Own	0%	3	100000
Luxbend	0%	1	100000
rigid many many	0%	2	100000
Jormungandr	0%	1	100000

Continued on next page

Table 5: Results of BFS on full dataset of games, with max 100,000 max search iterations and a timeout of 1 minute.

Game	Solved Levels %	# Total Levels	Max Search Iterations
twolittlecrates1	0%	1	100000
2048	0%	1	100000
watch your step	0%	1	100000
Harvesting Apples	0%	1	100000
Cell Division	0%	1	100000
rigid parallel unlimited	0%	1	100000
I Have No Mouth, And I	0%	1	100000
Cannot Push Two Boxes	0%	1	100000
Forest Road	0%	10	100000
randomspawner	0%	1	100000
Sok7	0%	1	100000
sok7	0%	1	100000
Ice Breaker	0%	1	100000
Big Dave	0%	1	100000
1D M3	0%	1	100000
Handle with Care	0%	3	100000
Merger Tiles - 2048	0%	1	100000
Remix	0%	1	100000
twolittlecrates2	0%	1	100000
Go	0%	5	100000
Bunshin master 0.75	0%	1	100000

1128 Additional references

- 1129 [35] Philip Bontrager, Wending Lin, Julian Togelius, and Sebastian Risi. Deep interactive evo-
1130 lution. In *Computational Intelligence in Music, Sound, Art and Design: 7th International*
1131 *Conference, EvoMUSART 2018, Parma, Italy, April 4-6, 2018, Proceedings*, pages 267–282.
1132 Springer, 2018.
- 1133 [36] Sam Earle, Samyak Parajuli, and Andrzej Banburski-Fahey. Dreamgarden: A designer assis-
1134 tant for growing games from a single prompt. In *Proceedings of the 2025 CHI Conference on*
1135 *Human Factors in Computing Systems*, pages 1–19, 2025.
- 1136 [37] Timothy Merino, Megan Charity, and Julian Togelius. Interactive latent variable evolution for
1137 the generation of minecraft structures. In *Proceedings of the 18th International Conference on*
1138 *the Foundations of Digital Games*, pages 1–8, 2023.
- 1139 [38] Jimmy Secretan, Nicholas Beato, David B D Ambrosio, Adelein Rodriguez, Adam Campbell,
1140 and Kenneth O Stanley. Picbreeder: evolving pictures collaboratively online. In *Proceedings*
1141 *of the SIGCHI conference on human factors in computing systems*, pages 1759–1768, 2008.
- 1142 [39] Graham Todd, Alexander Padula, Matthew Stephenson, Éric Piette, Dennis JNJ Soemers, and
1143 Julian Togelius. Gavel: Generating games via evolution and language models. *arXiv preprint*
1144 *arXiv:2407.09388*, 2024.