

# AI AND GAMES

Fun Session

CMPUT 397

*9th Oct 2020*



# HISTORY OF AI AND GAMES

# HISTORY OF AI AND GAMES

Year	Game	AI System, <i>Creator(s)</i>
------	------	------------------------------

# HISTORY OF AI AND GAMES

Year	Game	AI System, <i>Creator(s)</i>
1959	Checkers	<i>A. Samuel</i>



# HISTORY OF AI AND GAMES

Year	Game	AI System, <i>Creator(s)</i>
1959	Checkers	<i>A. Samuel</i>
1992	Backgammon	TD-Gammon, <i>G. Tesauro</i>

# HISTORY OF AI AND GAMES

Year	Game	AI System, <i>Creator(s)</i>
1959	Checkers	<i>A. Samuel</i>
1992	Backgammon	TD-Gammon, <i>G. Tesauro</i>
1997	Chess	DeepBlue, <i>IBM</i>

# HISTORY OF AI AND GAMES

Year	Game	AI System, <i>Creator(s)</i>
1959	Checkers	<i>A. Samuel</i>
1992	Backgammon	TD-Gammon, <i>G. Tesauro</i>
1997	Chess	DeepBlue, <i>IBM</i>
2007	Checkers	Chinook, <i>J. Schaeffer et al.</i>

# HISTORY OF AI AND GAMES

Year	Game	AI System, <i>Creator(s)</i>
1959	Checkers	<i>A. Samuel</i>
1992	Backgammon	TD-Gammon, <i>G. Tesauro</i>
1997	Chess	DeepBlue, <i>IBM</i>
2007	Checkers	Chinook, <i>J. Schaeffer et al.</i>
2013	Atari 2600	DQN, <i>V. Minh et al.</i>

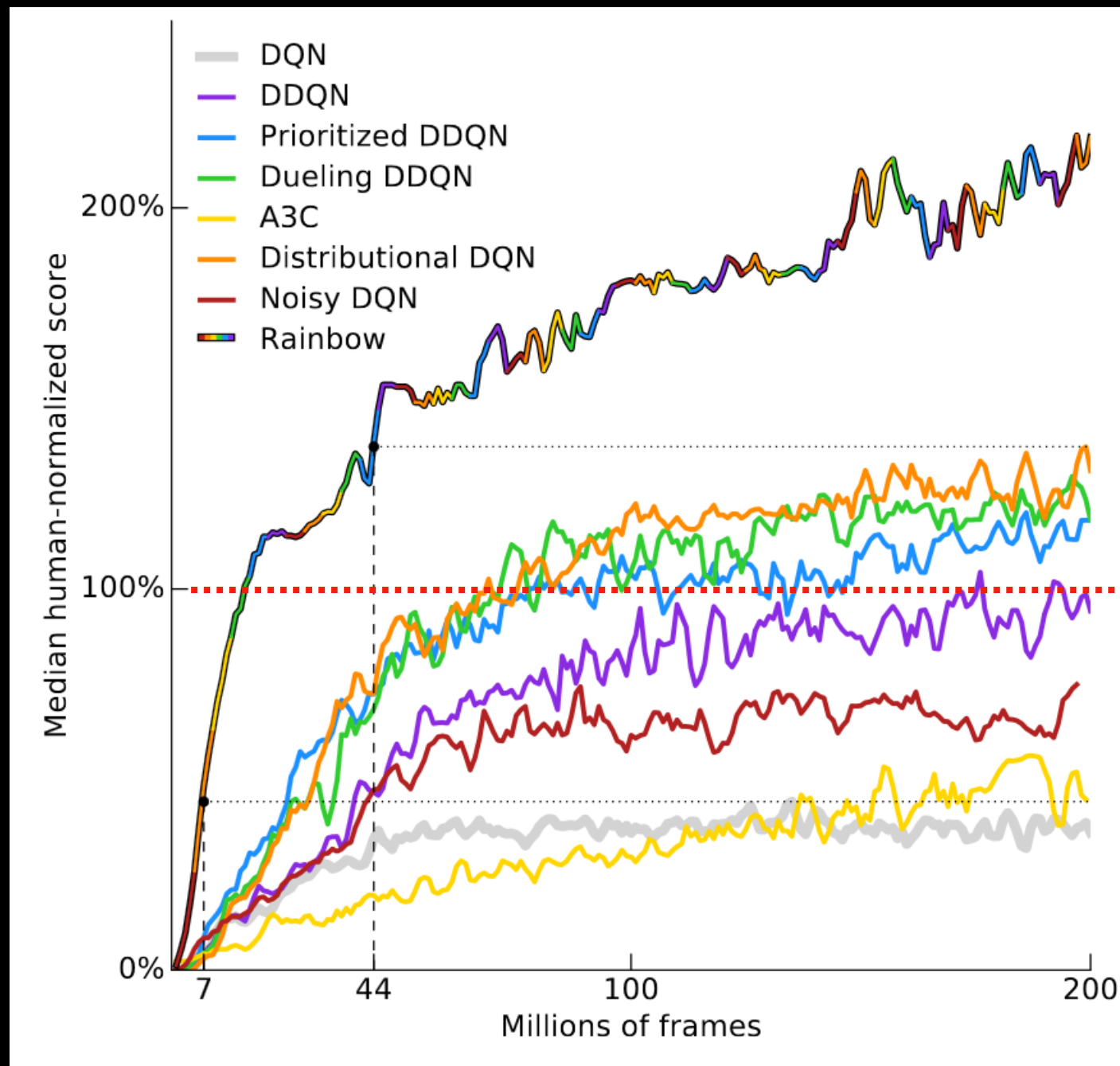
# HISTORY OF AI AND GAMES

Year	Game	AI System, <i>Creator(s)</i>
1959	Checkers	<i>A. Samuel</i>
1992	Backgammon	TD-Gammon, <i>G. Tesauro</i>
1997	Chess	DeepBlue, <i>IBM</i>
2007	Checkers	Chinook, <i>J. Schaeffer et al.</i>
2013	Atari 2600	DQN, <i>V. Minh et al.</i>
2016	Go	AlphaGo, <i>D. Silver et al.</i>

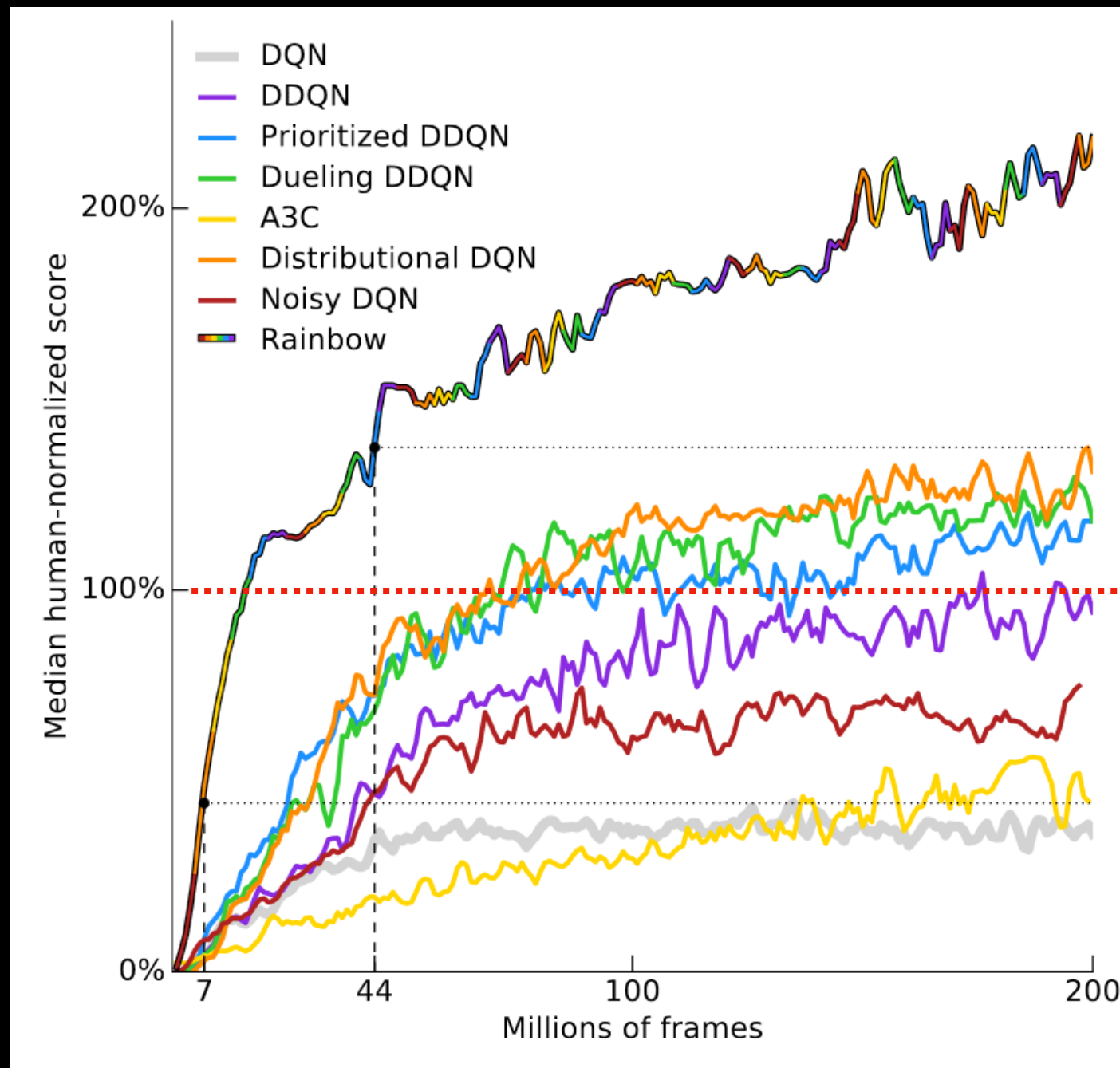
# HISTORY OF AI AND GAMES

Year	Game	AI System, <i>Creator(s)</i>
1959	Checkers	<i>A. Samuel</i>
1992	Backgammon	TD-Gammon, <i>G. Tesauro</i>
1997	Chess	DeepBlue, <i>IBM</i>
2007	Checkers	Chinook, <i>J. Schaeffer et al.</i>
2013	Atari 2600	DQN, <i>V. Minh et al.</i>
2016	Go	AlphaGo, <i>D. Silver et al.</i>
2018	Poker	DeepStack, <i>M. Bowling et al.</i>

# TRAINING TIME ON ATARI



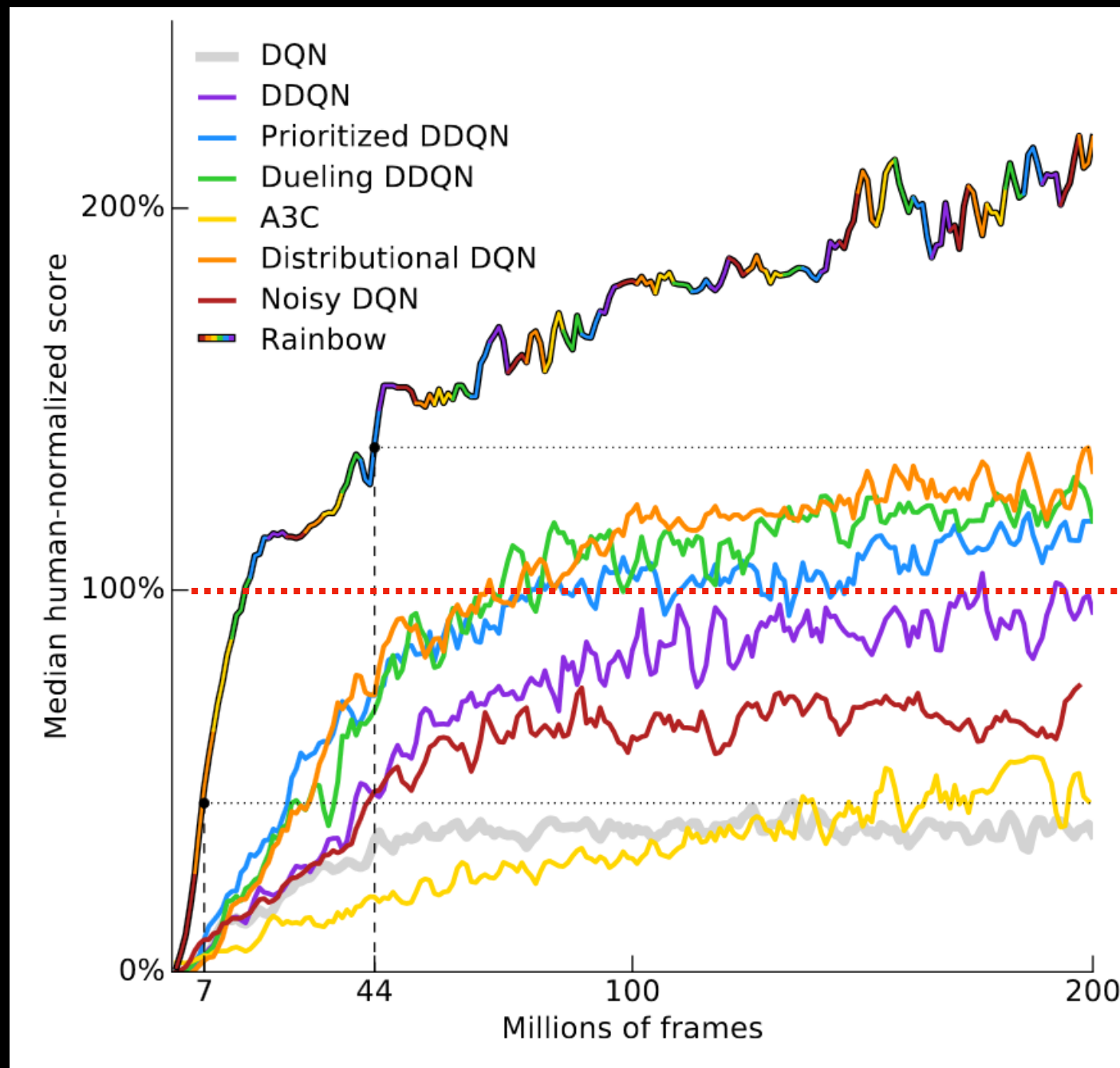
# TRAINING TIME ON ATARI



‘Rainbow’ takes around 20 *million* frames of experience to achieve the performance of an average human



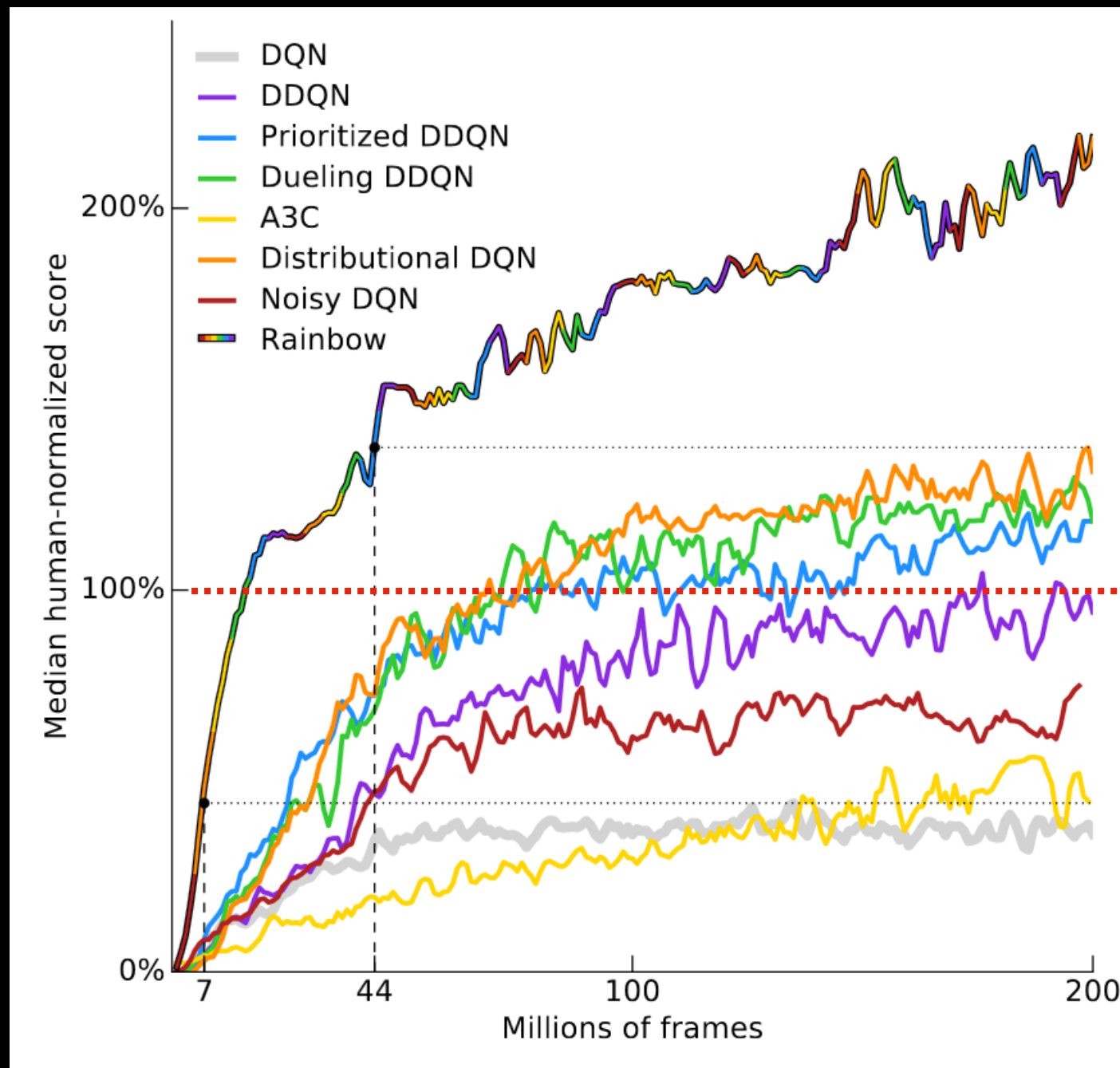
# TRAINING TIME ON ATARI



‘Rainbow’ takes around 20 *million* frames of experience to achieve the performance of an average human

Corresponds to about 83 *hours* of gameplay

# TRAINING TIME ON ATARI



‘Rainbow’ takes around 20 *million* frames of experience to achieve the performance of an average human

Corresponds to about 83 *hours* of gameplay

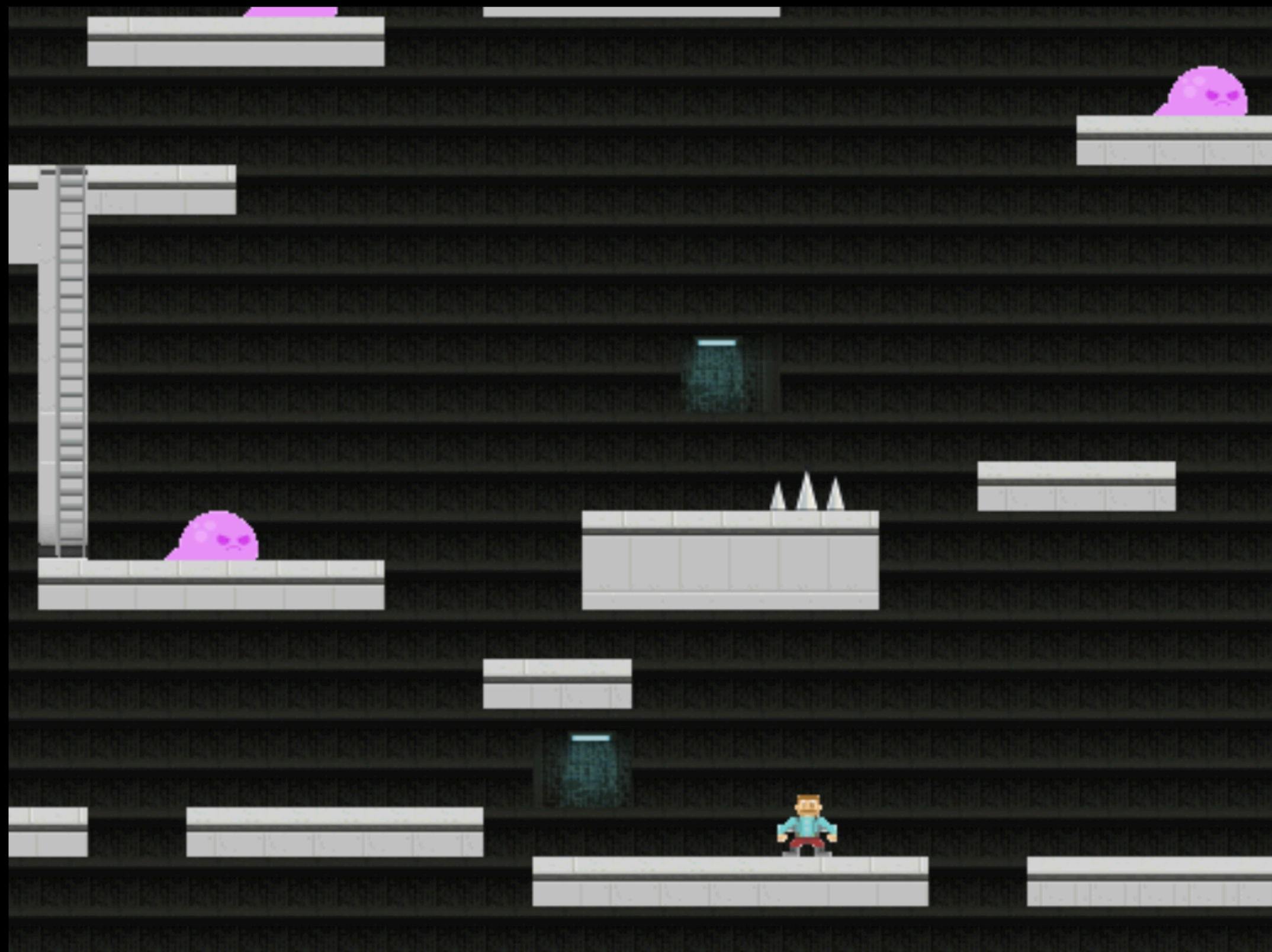
DQN couldn’t reach it even after 200 *million* frames of experience

**LET'S TRY TO UNDERSTAND  
WHY...**

**LET'S TRY TO UNDERSTAND  
WHY...**

[https://rach0012.github.io/humanRL\\_website](https://rach0012.github.io/humanRL_website)

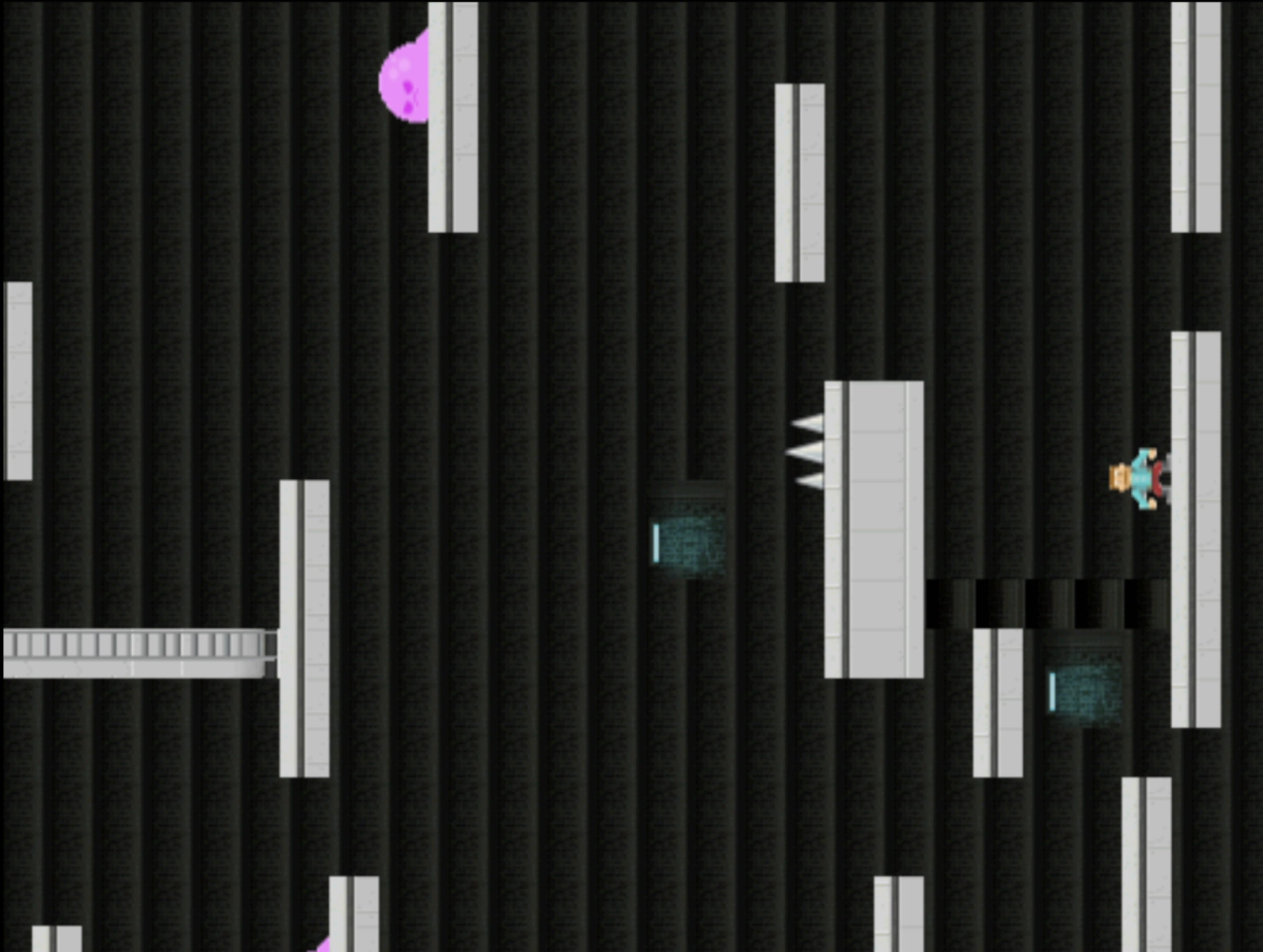
# GAME 1: (A)



[https://rach0012.github.io/humanRL\\_website](https://rach0012.github.io/humanRL_website)



# GAME 2: (H)



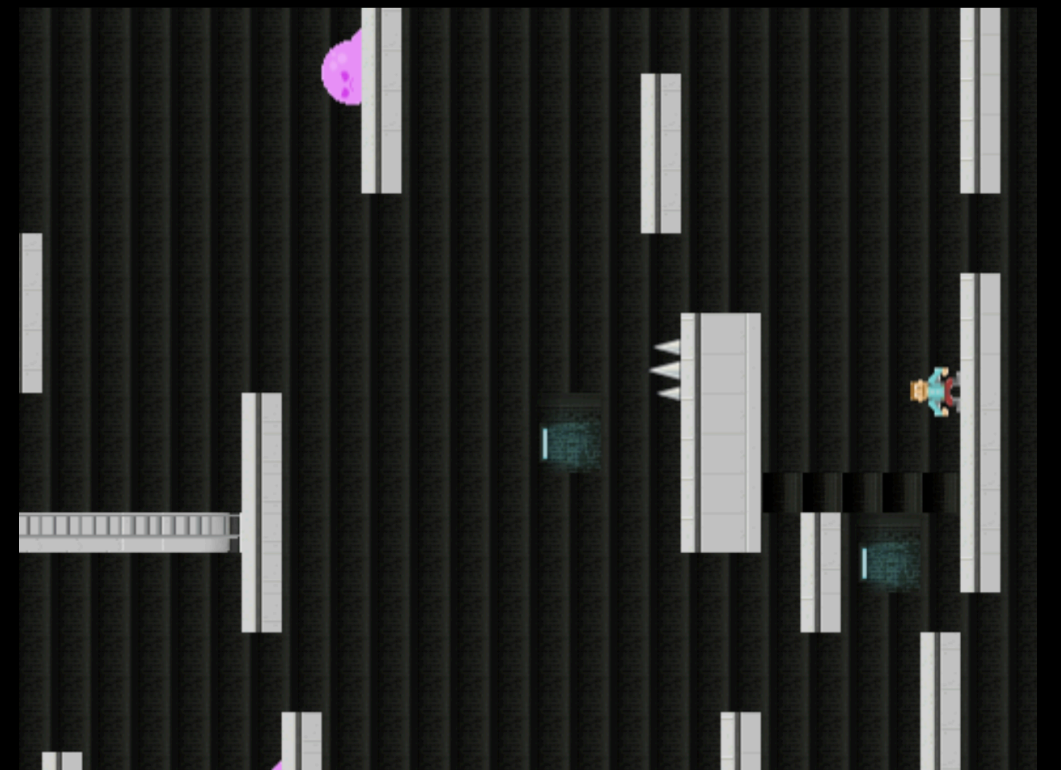
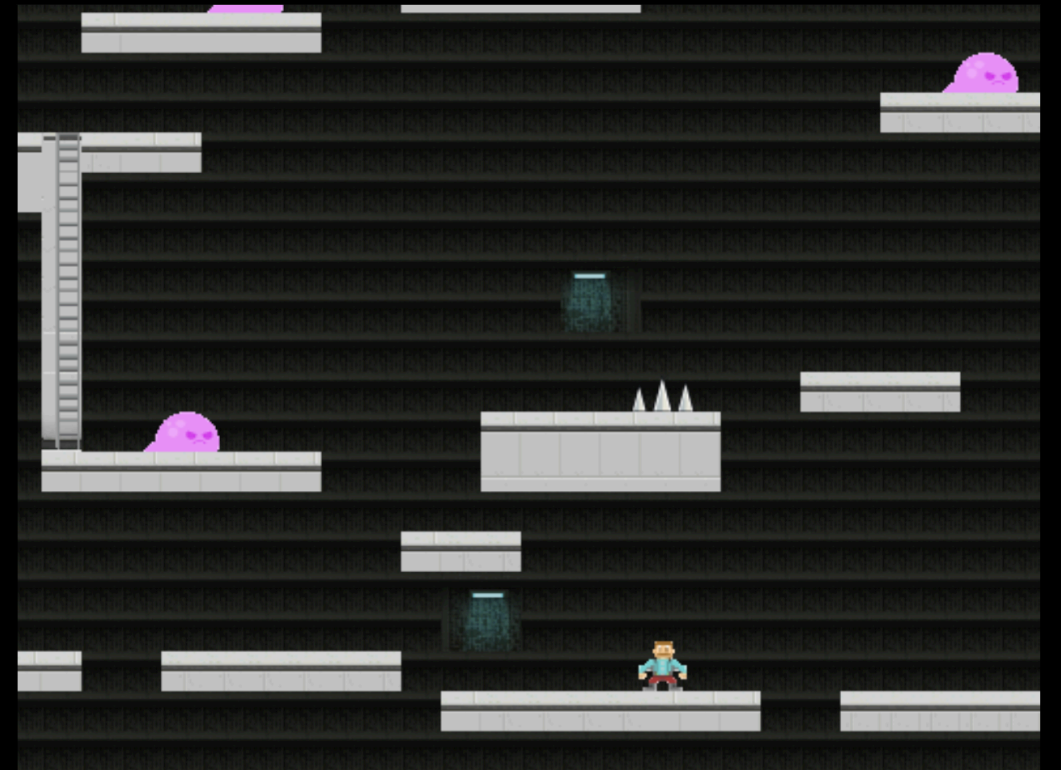
[https://rach0012.github.io/humanRL\\_website](https://rach0012.github.io/humanRL_website)

# GAME 2: (H)

What did we learn?

We rely on prior knowledge that:

- ▶ left arrow key leads left,
- ▶ right arrow key leads right,
- ▶ etc.





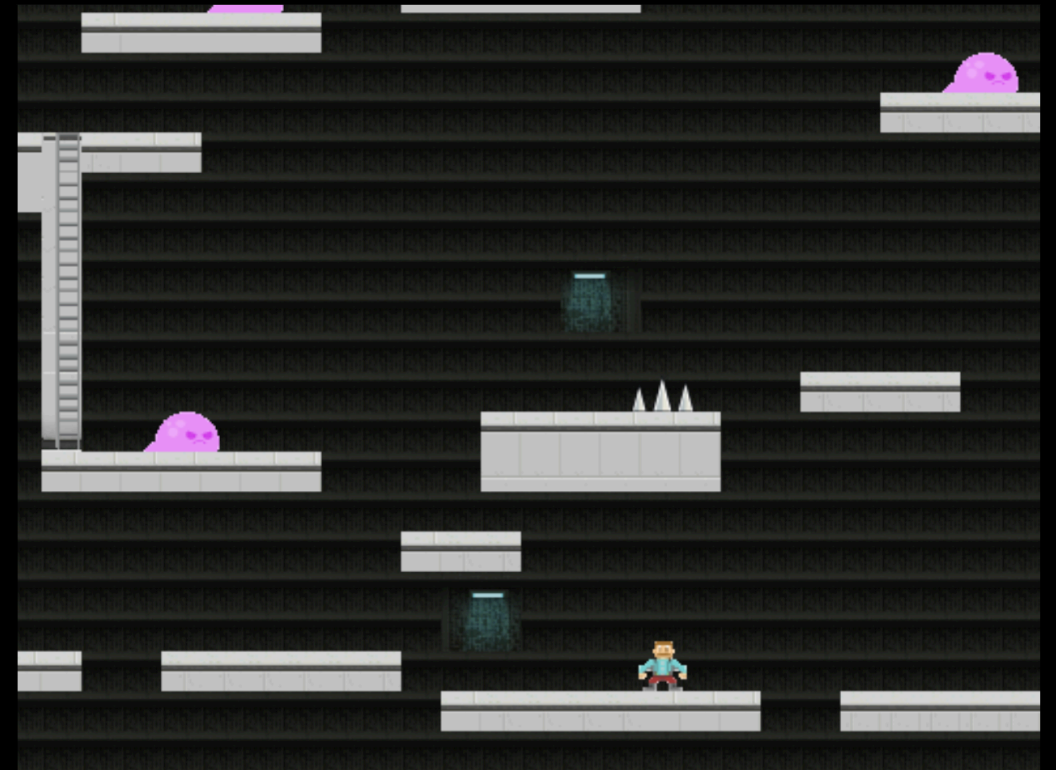
# GAME 3: (B)



[https://rach0012.github.io/humanRL\\_website](https://rach0012.github.io/humanRL_website)



# GAME 3: (B)

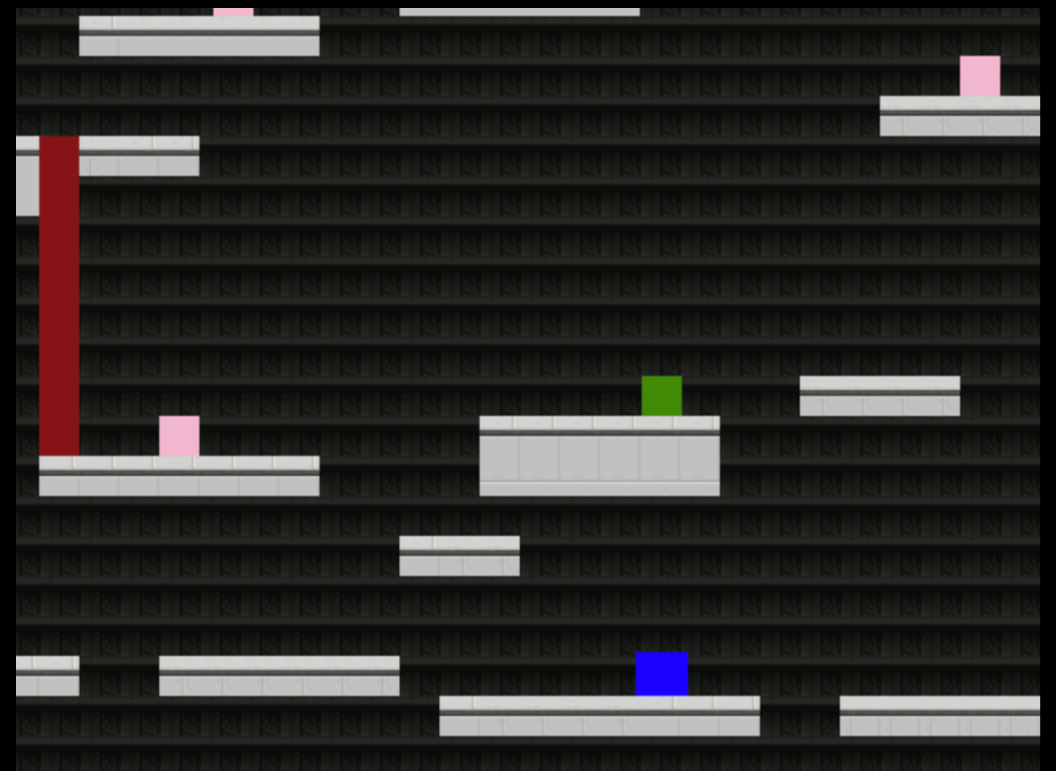
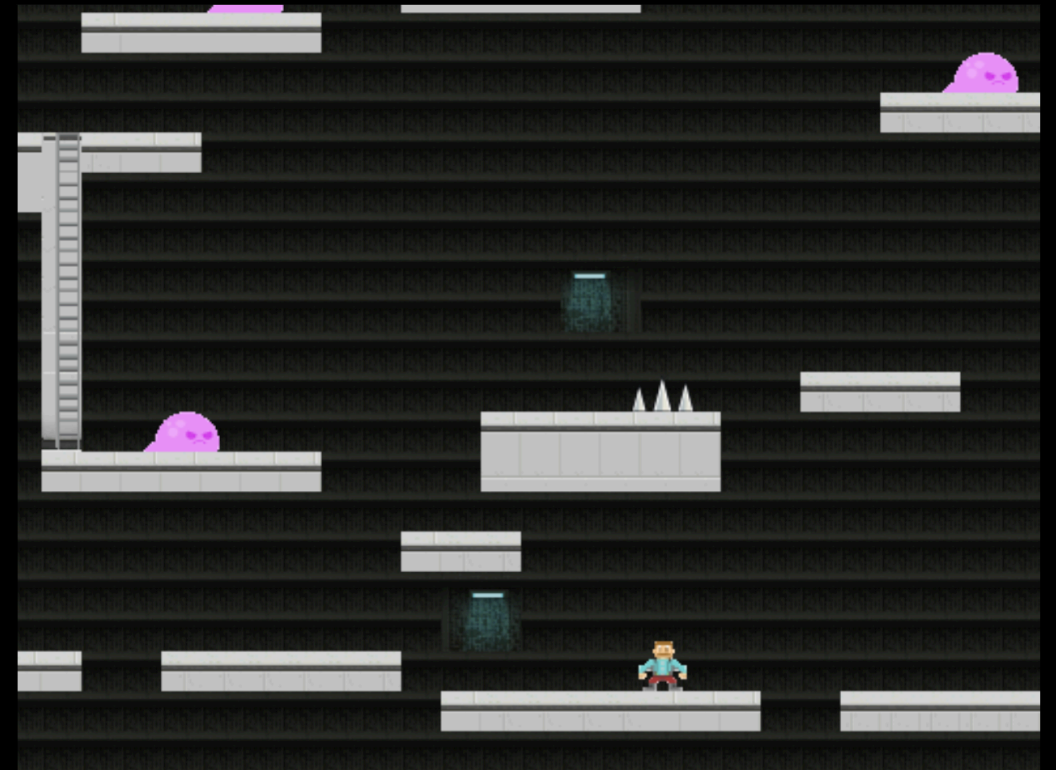


# GAME 3: (B)

What did we learn?

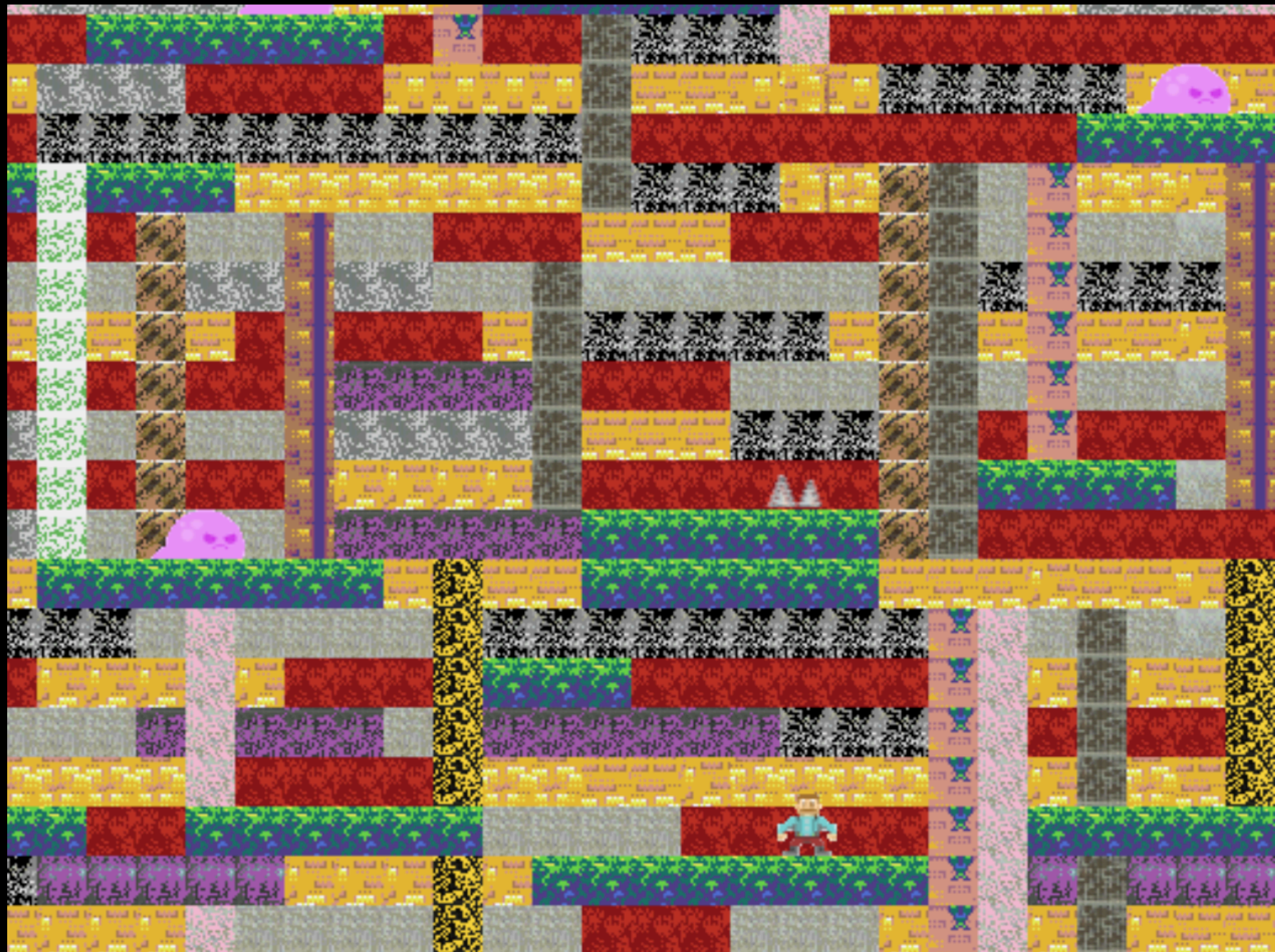
We rely on prior knowledge that:

- ▶ Ladders can be climbed
- ▶ Spikes are bad
- ▶ Monsters are bad
- ▶ Keys can open doors





# GAME 4: (E)



[https://rach0012.github.io/humanRL\\_website](https://rach0012.github.io/humanRL_website)

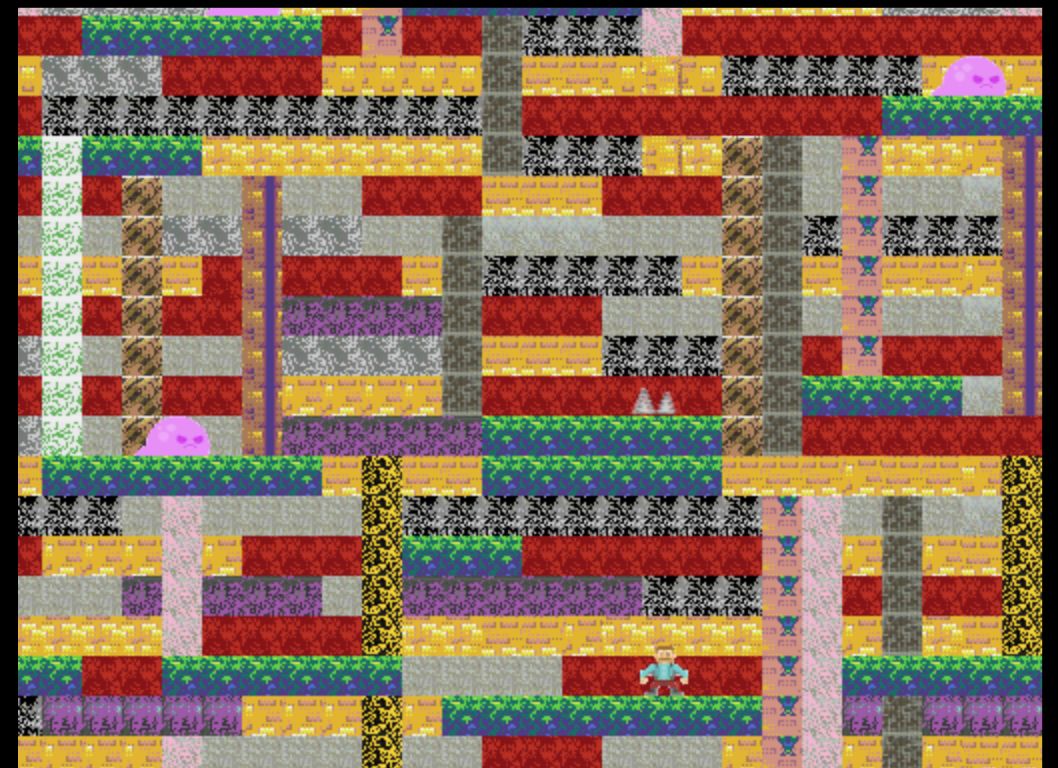
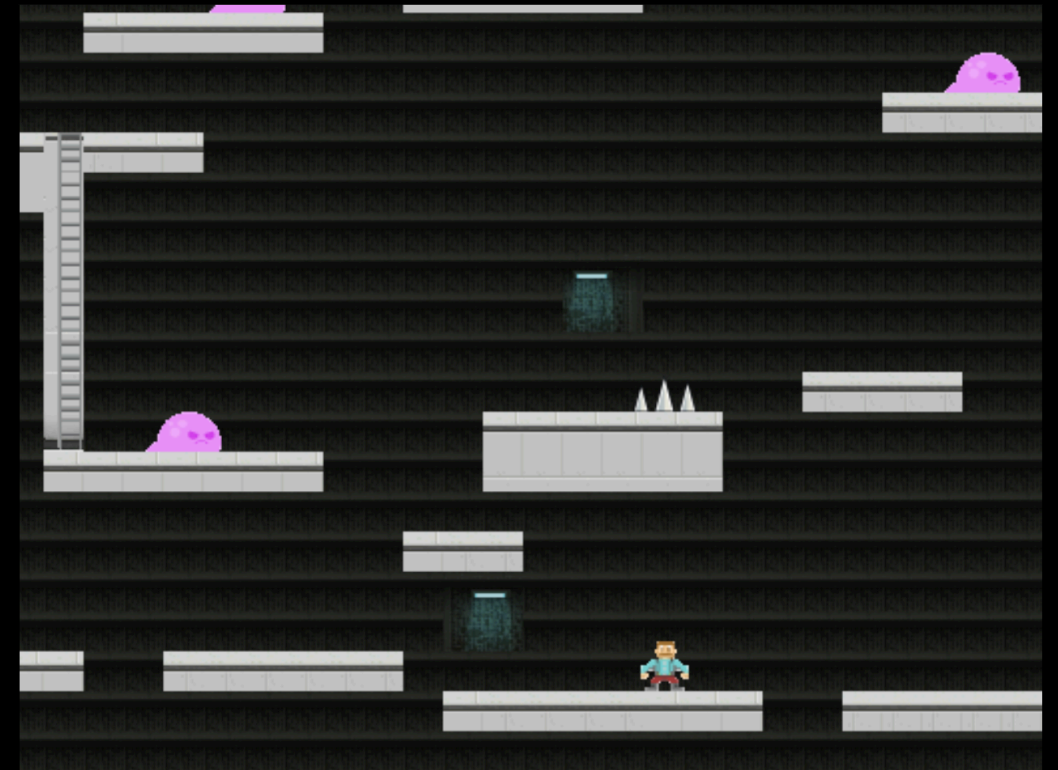


## GAME 4: (E)

# What did we learn?

# We rely on prior knowledge that:

- ▶ Black space is empty — has no objects





# GAME 5: (F)



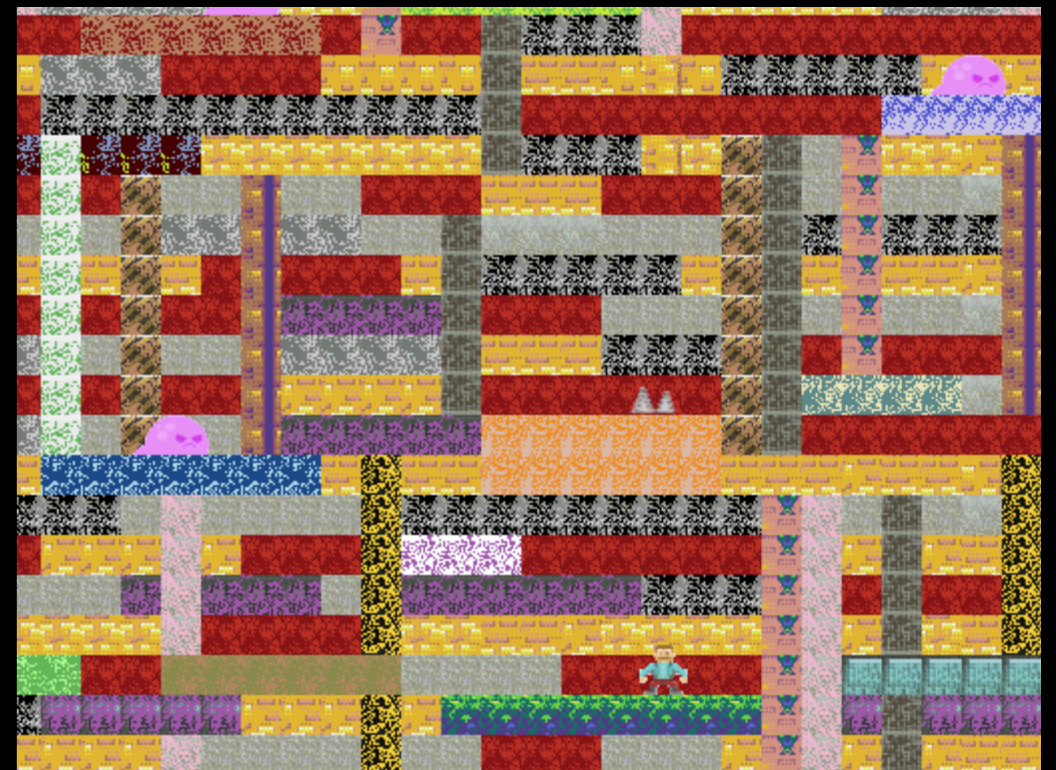
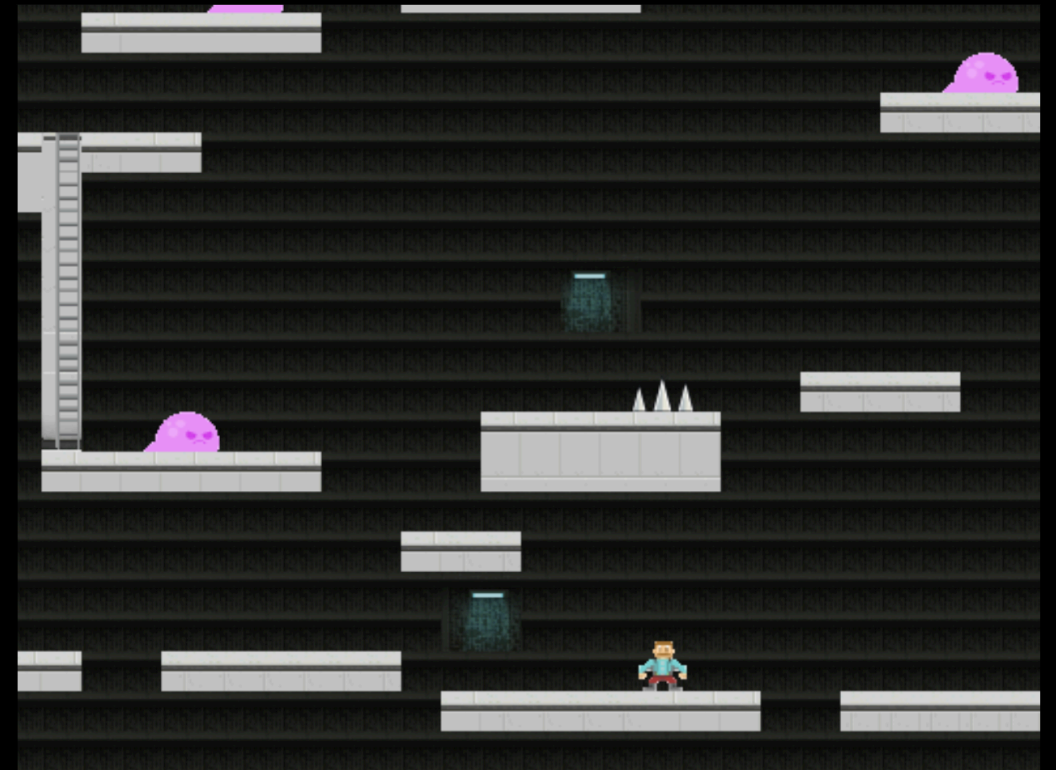


# GAME 5: (F)

What did we learn?

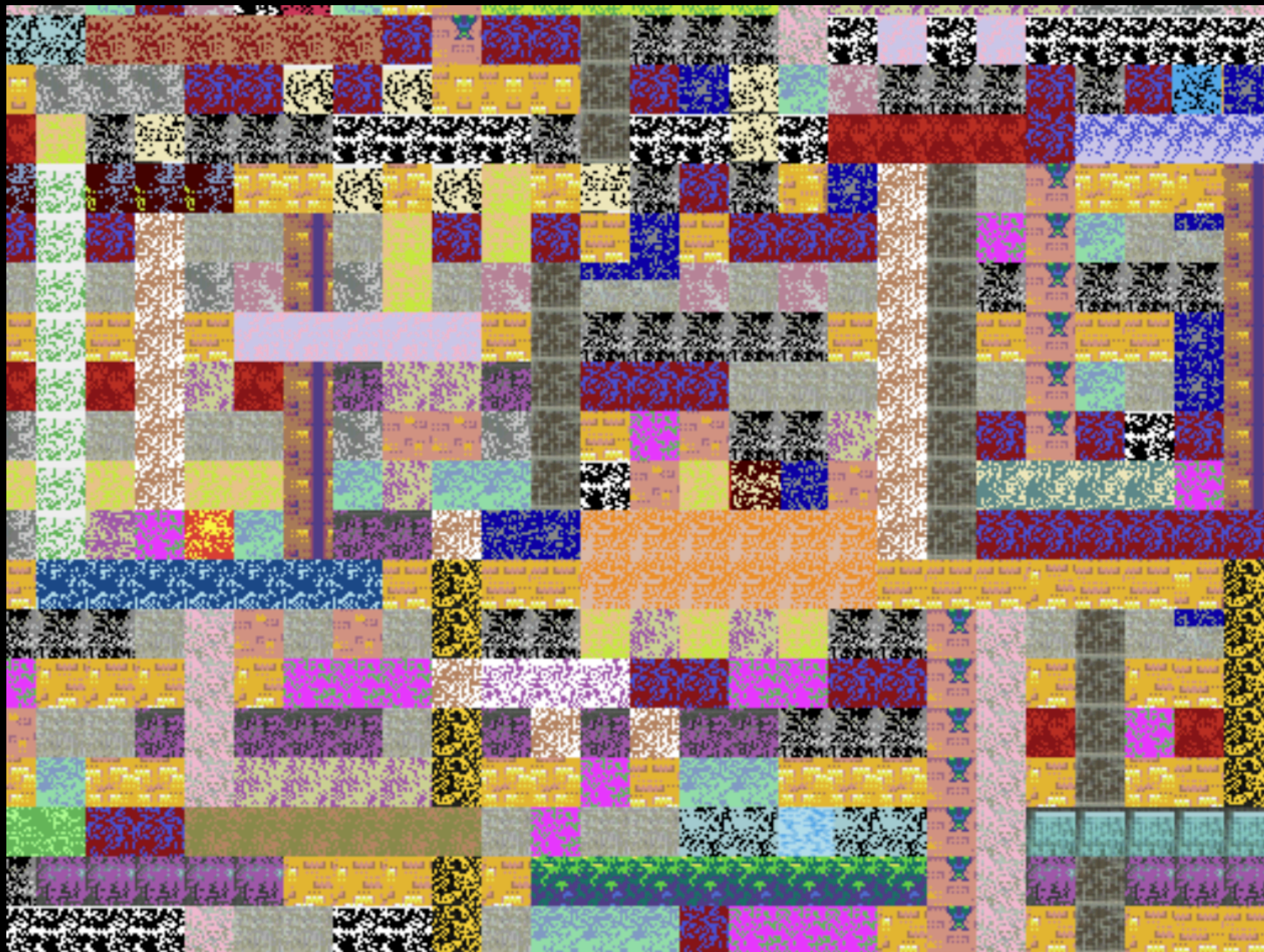
We rely on prior knowledge that:

- ▶ Things that look similar, behave similarly





# GAME 6: HARD MODE



[https://rach0012.github.io/humanRL\\_website](https://rach0012.github.io/humanRL_website)

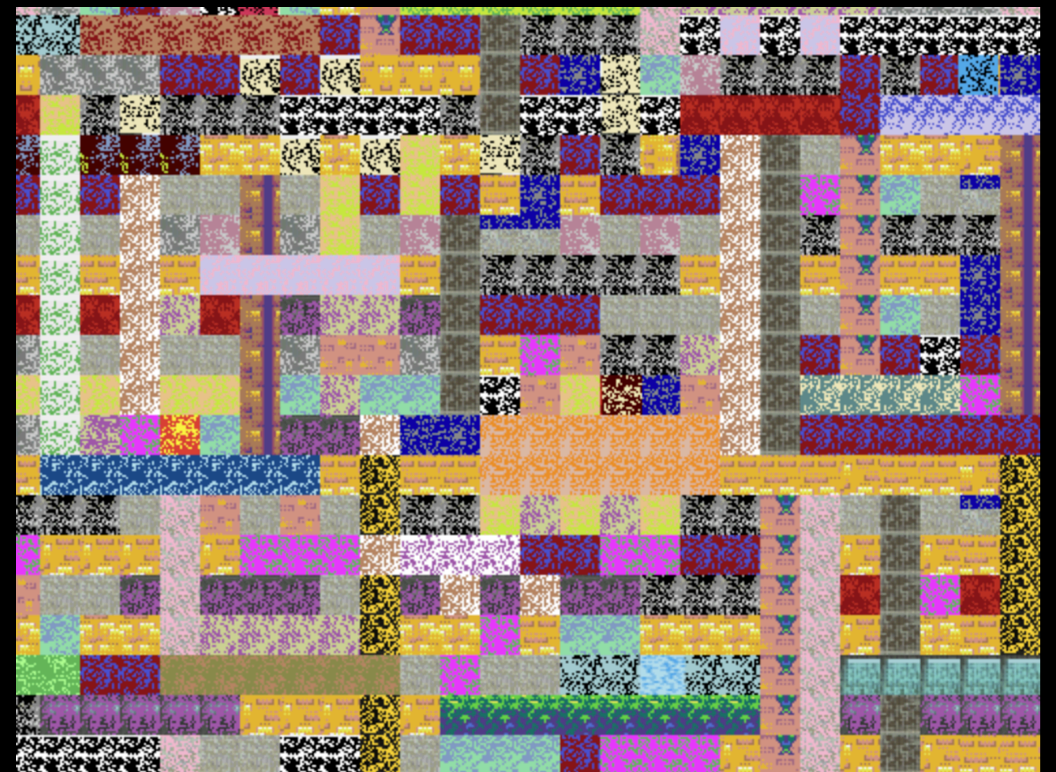
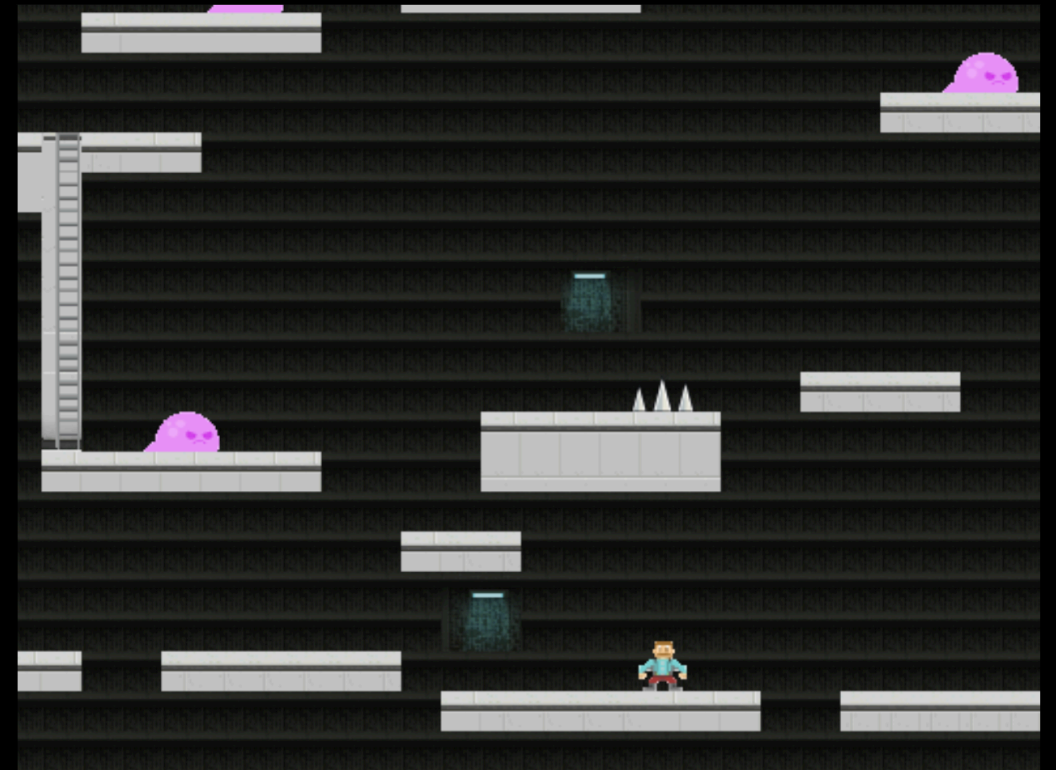


# GAME 6: HARD MODE

What did we learn?

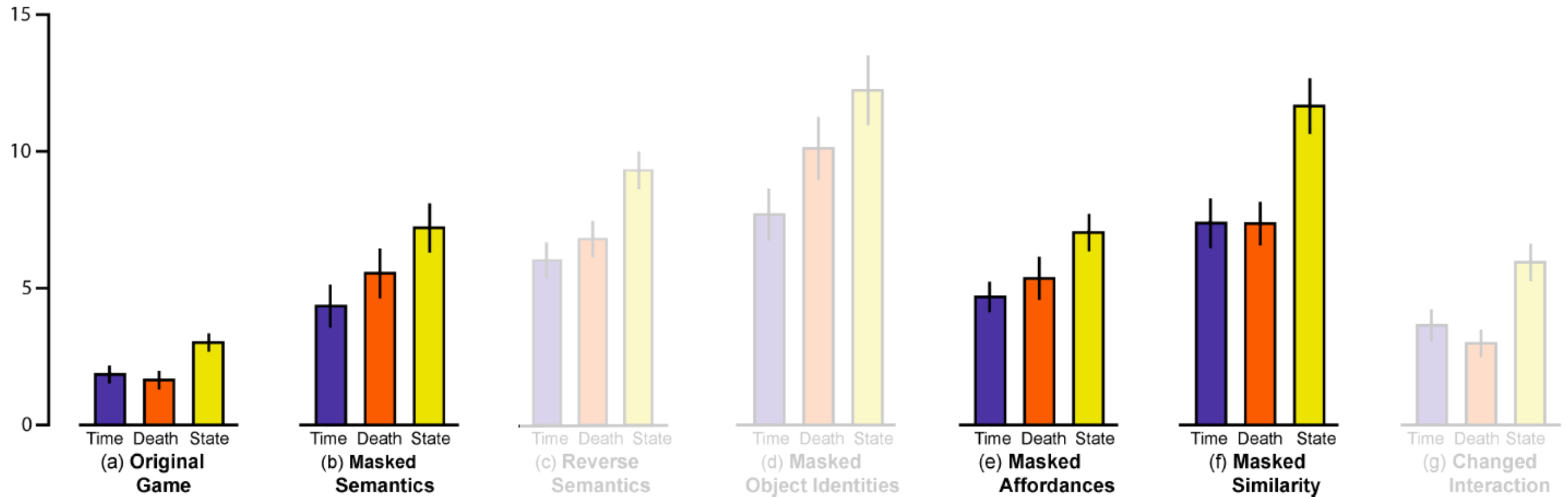
We rely on prior knowledge that:

- ▶ Left arrow key leads left, etc.
- ▶ Ladders can be climbed
- ▶ Spikes are bad
- ▶ Monsters are bad
- ▶ Keys can open doors
- ▶ Black space is empty — has no objects
- ▶ Things that look similar, behave similarly



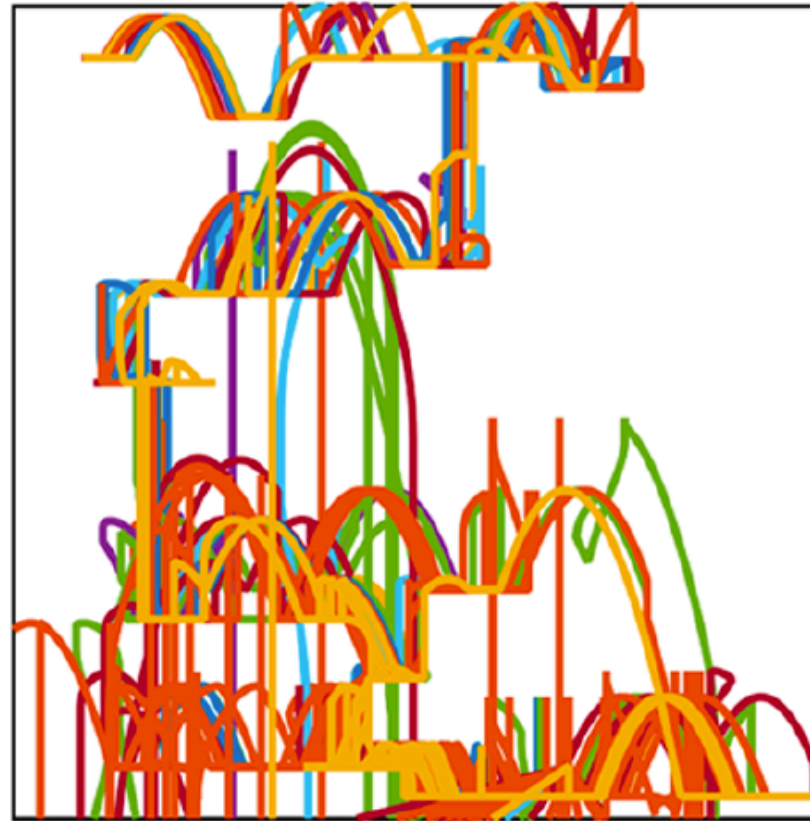
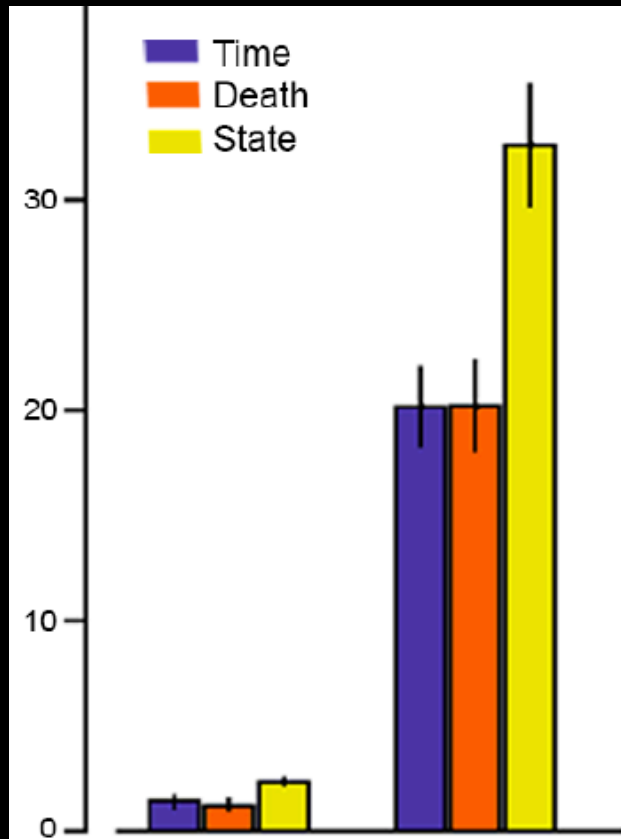


# SOME STATISTICS

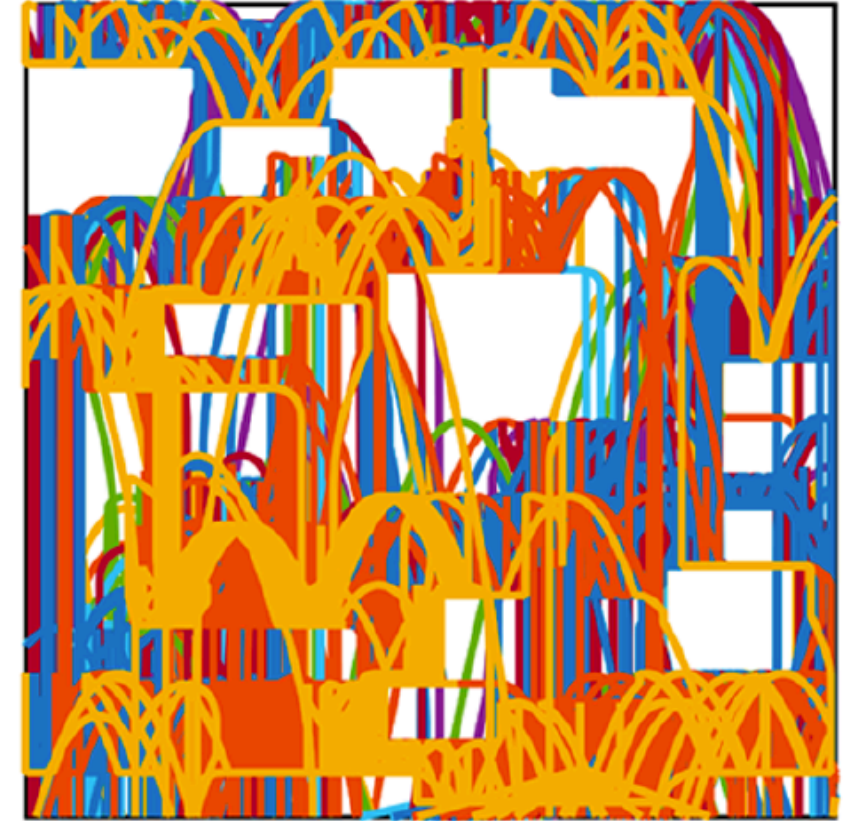


*Figure 3. Quantifying the influence of various object priors.* The blue bar shows average time taken by humans (in minutes), orange bar shows the average number of deaths, and yellow bar shows the number of unique states visited by players to solve the various games. For visualization purposes, the number of deaths is divided by 2, and the number of states is divided by 1000 respectively.

# SOME STATISTICS



Original game



Hard mode

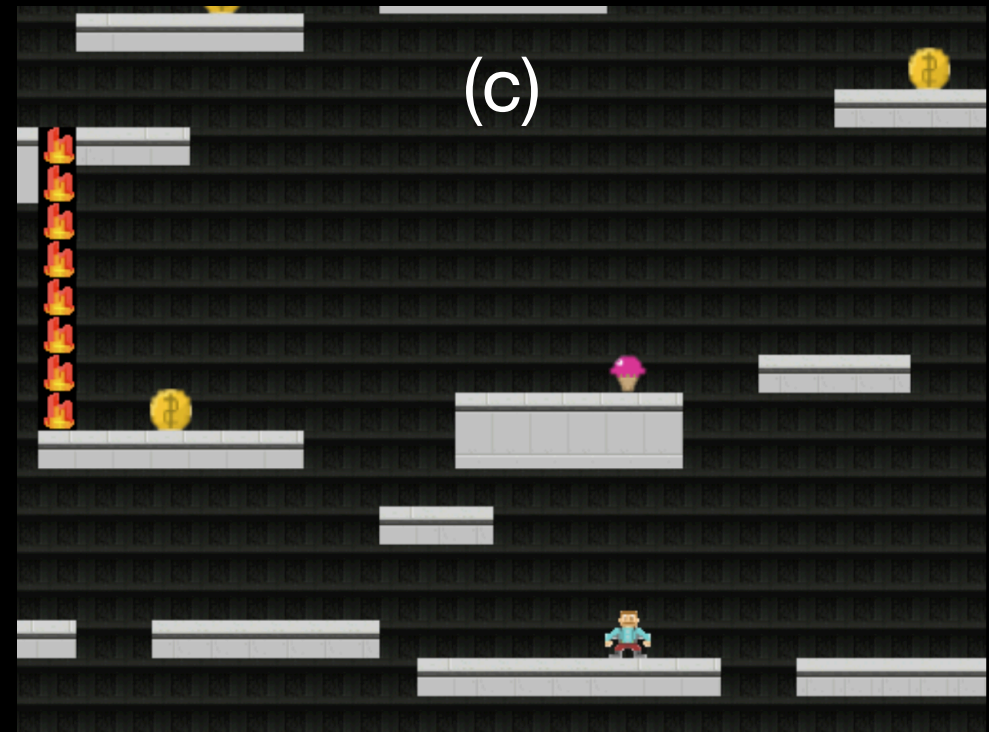


# PRIOR KNOWLEDGE — AMAZING?

- In all the previous cases, prior information was useful in solving the game/task.
- Does that mean we just try and build in as much prior knowledge as possible into our agents?

# PRIOR KNOWLEDGE — AMAZING?

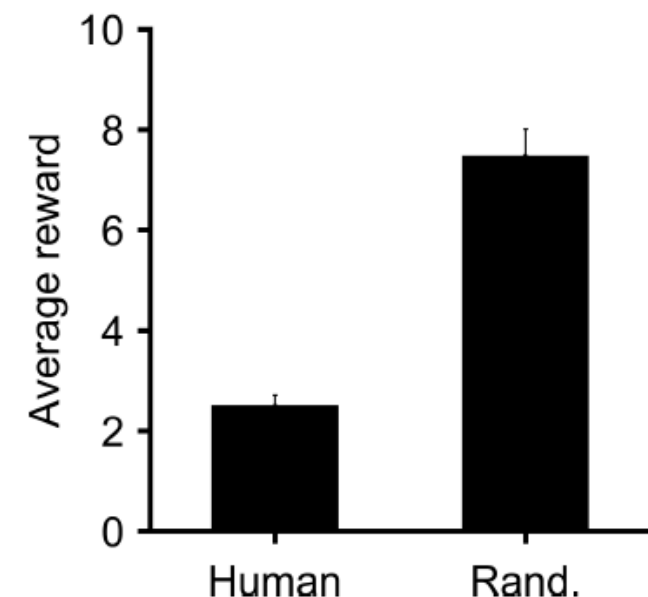
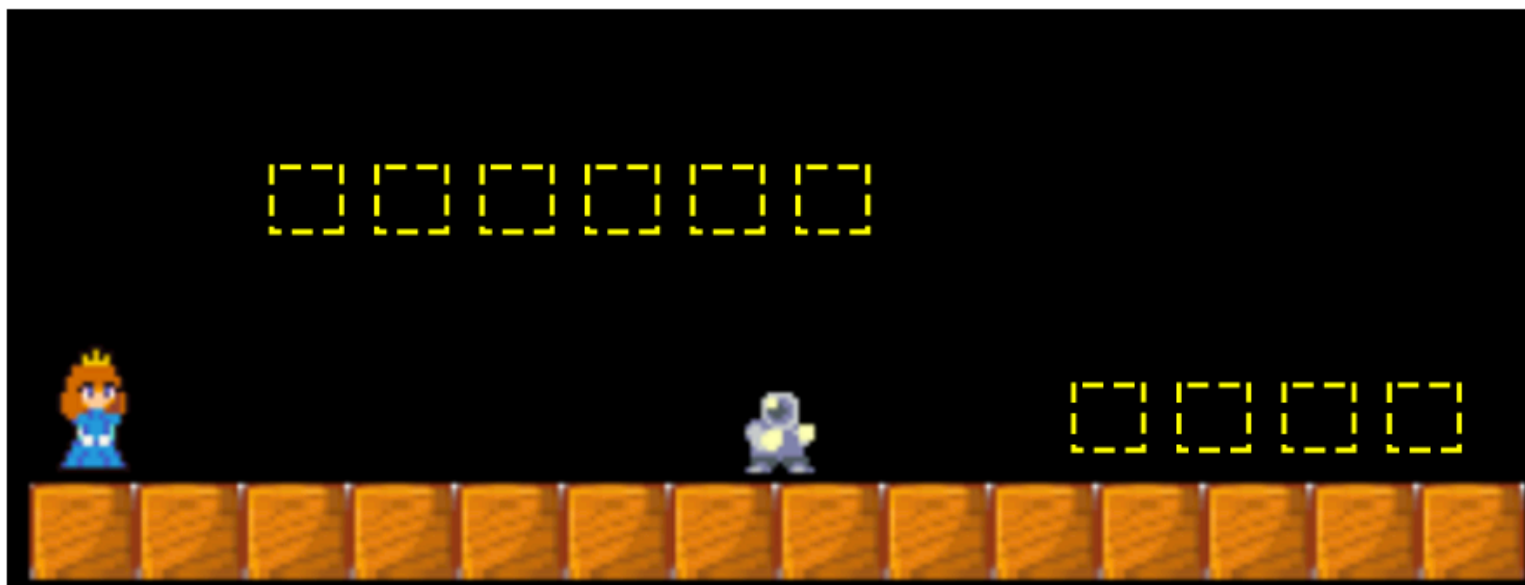
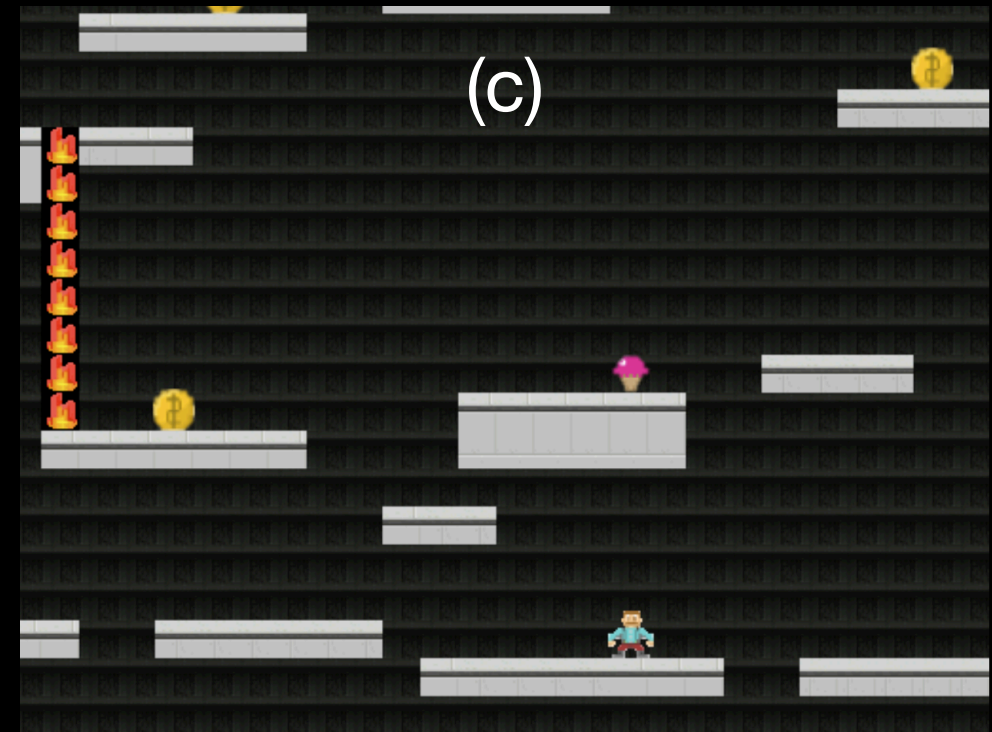
- In all the previous cases, prior information was useful in solving the game/task.
- Does that mean we just try and build in as much prior knowledge as possible into our agents?





# PRIOR KNOWLEDGE — AMAZING?

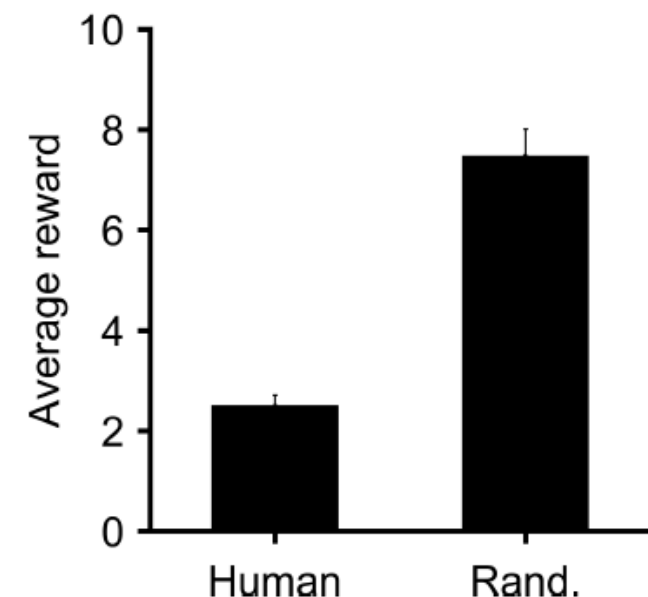
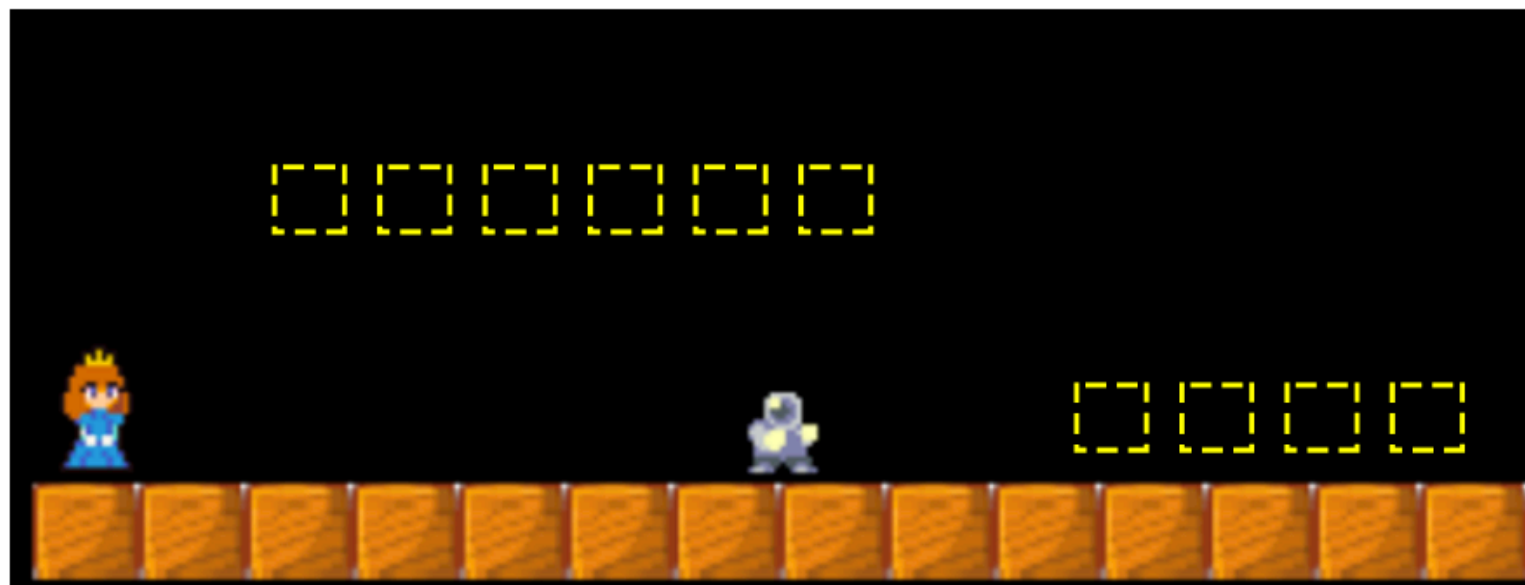
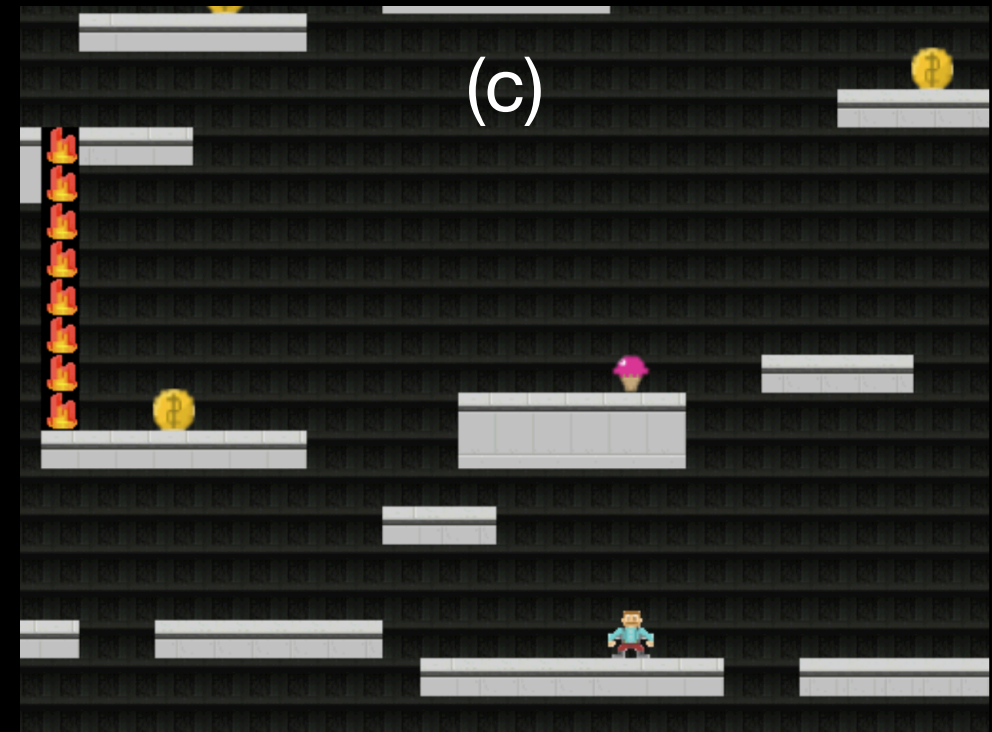
- In all the previous cases, prior information was useful in solving the game/task.
- Does that mean we just try and build in as much prior knowledge as possible into our agents?

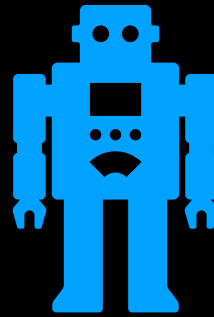
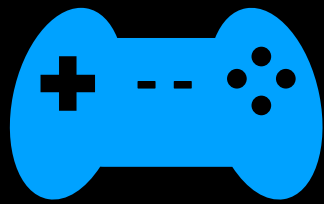


# PRIOR KNOWLEDGE – AMAZING?

- In all the previous cases, prior information was useful in solving the game/task.
- Does that mean we just try and build in as much prior knowledge as possible into our agents?

*Maybe not...*





# QUESTIONS?

Paper: <https://arxiv.org/abs/1802.10217>

Website: [https://rach0012.github.io/humanRL\\_website](https://rach0012.github.io/humanRL_website)