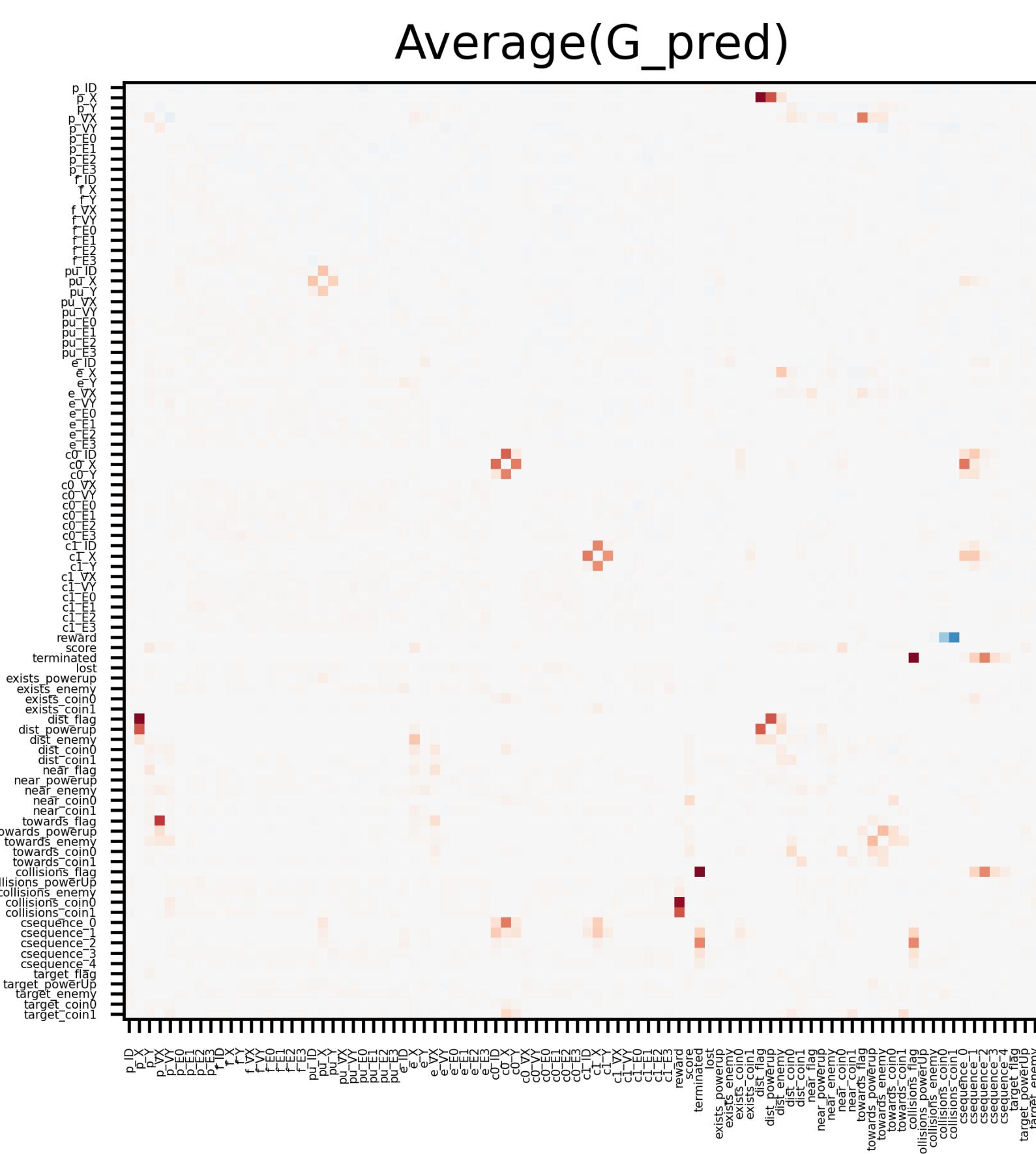
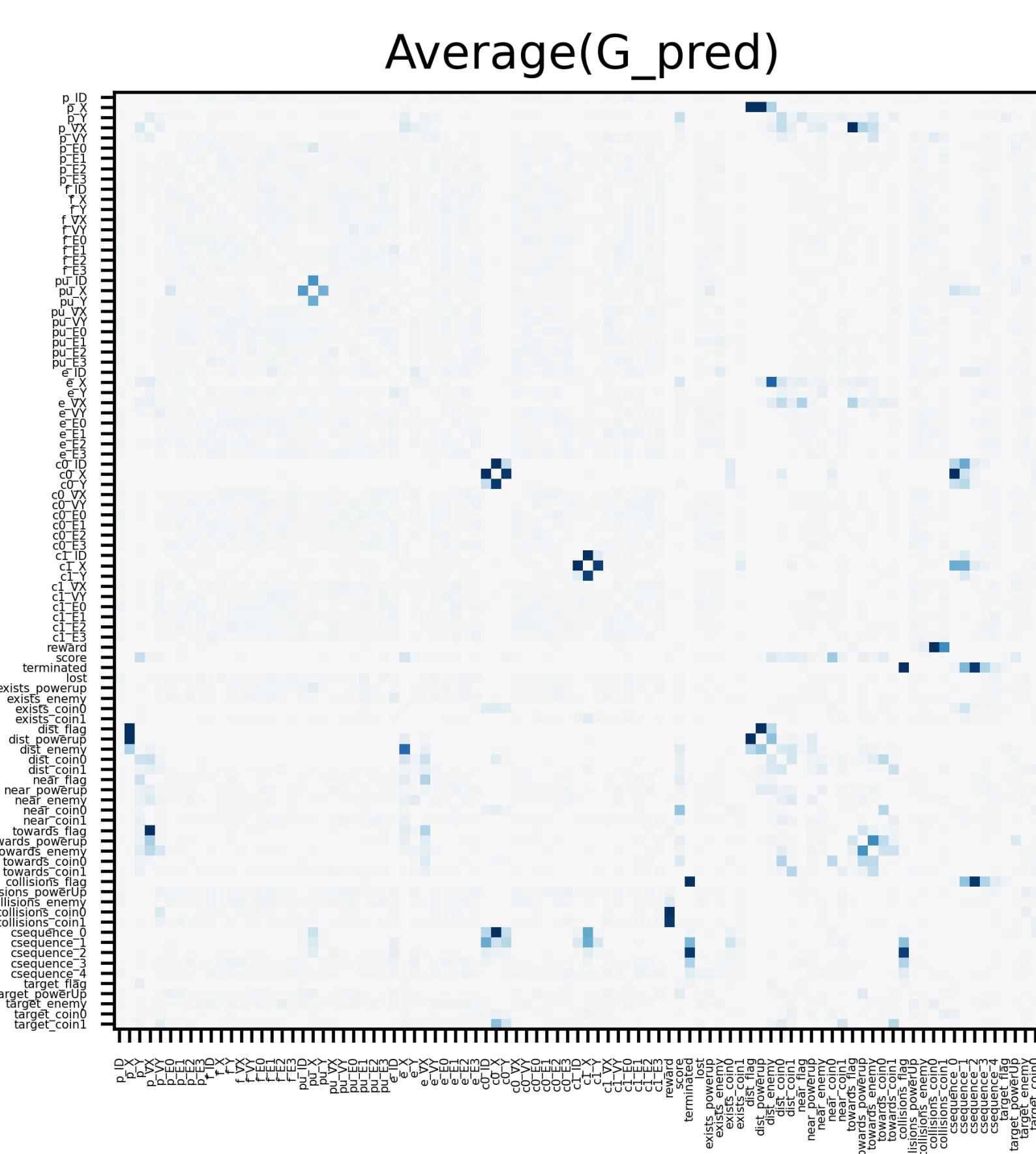


Key questions:

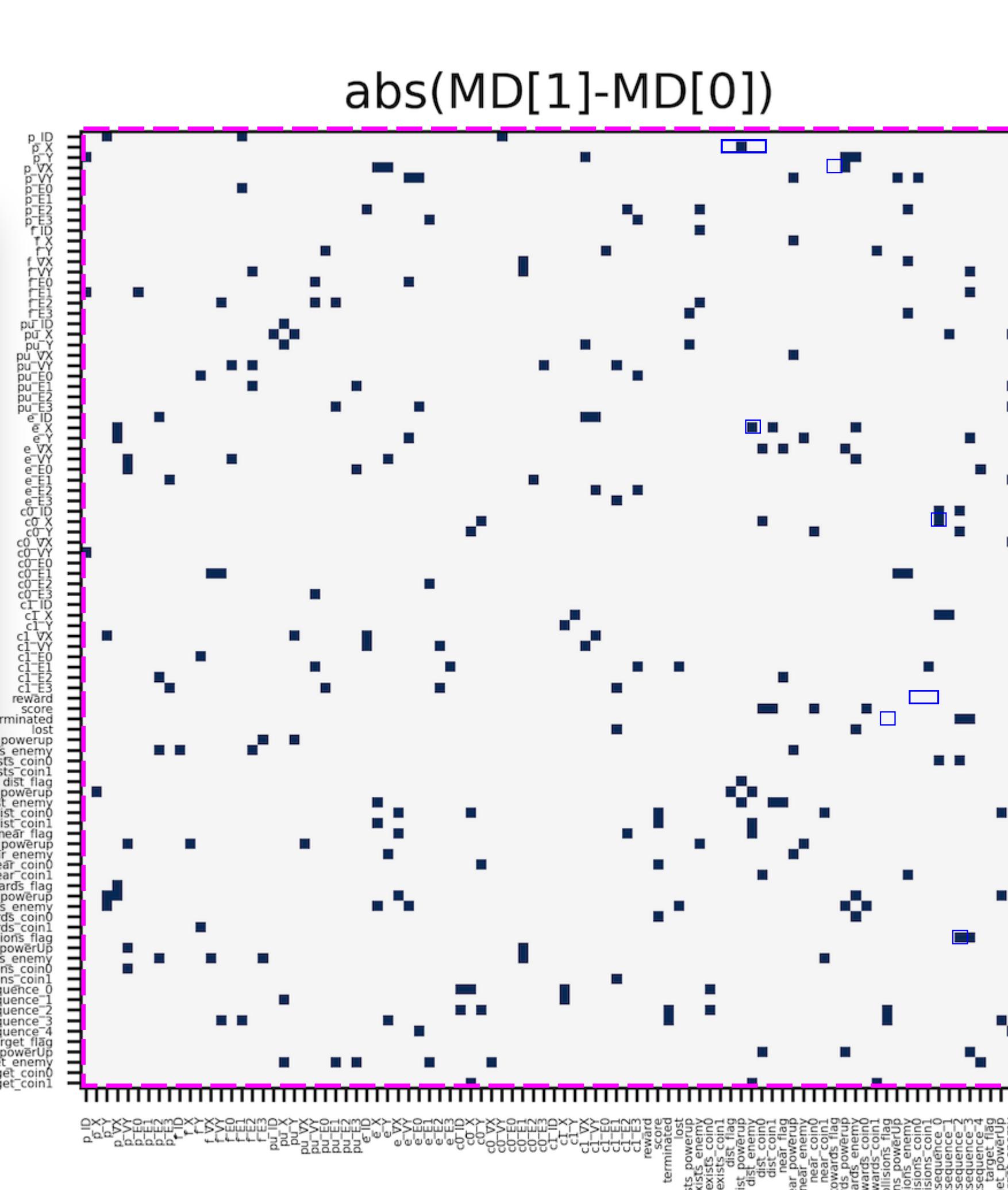
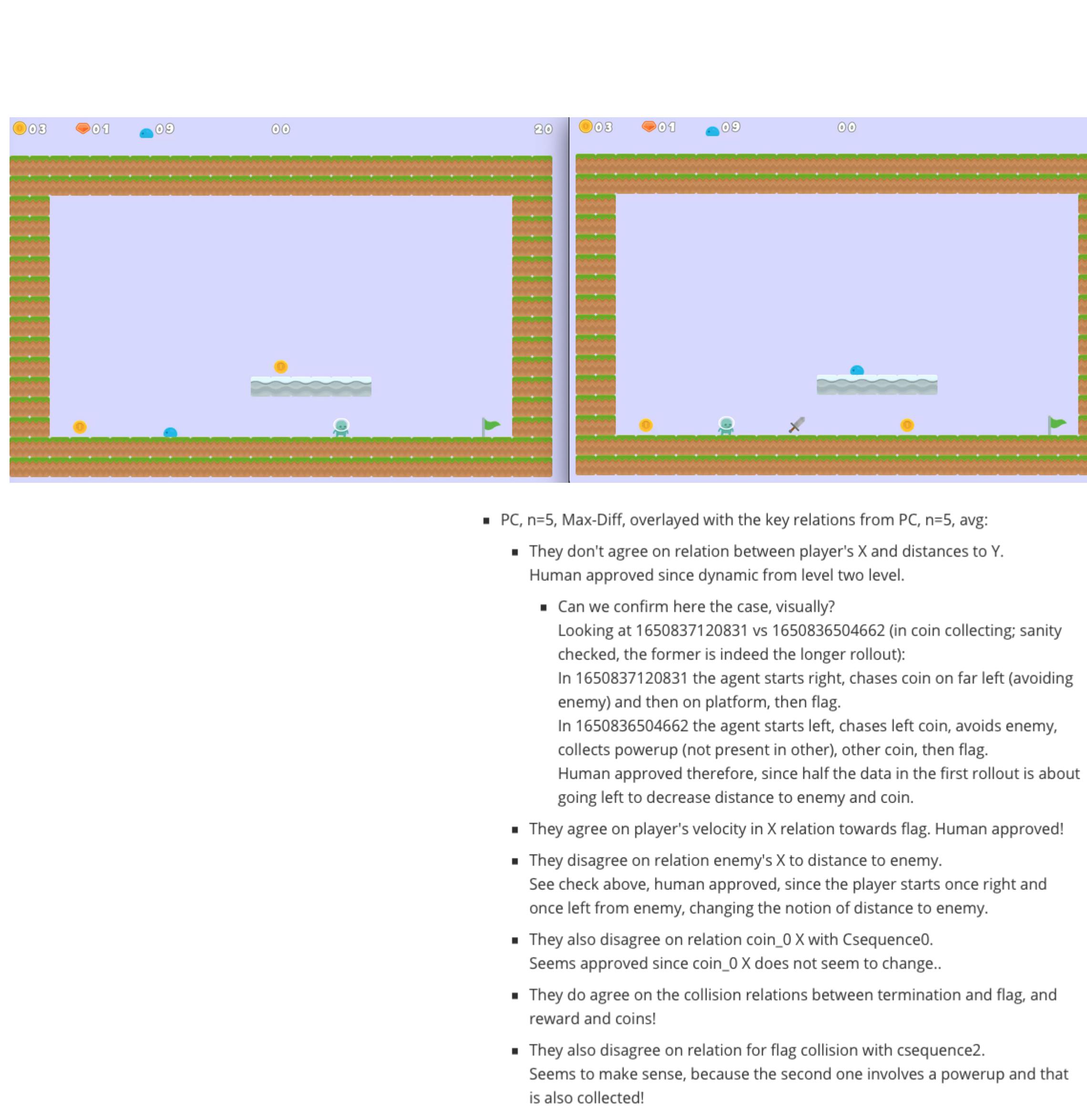
- * how important is n?
- * how important the algo?
- * what different is told by the two types?
- * do the state descriptions agree?
- * what can and cannot be applied to SCIT?
- * does the graph reveal the behavior deployed (e.g. coin collecting)?



- PC, n=5, avg, not checked the 5 recordings, looking only at stronger activations:
 - Player's X negative relation to Distance of Flag, Powerup, and Enemy. Human approved, since increasing X value (pointing to the right), decreases the distance to Flag (which always stands right) and since Enemy can vary, here the activation is less.
 - But why no relation to distance to coin? Especially when its coin collecting behavior?
 - Player's Velocity in X negative relation to Towards Flag. Human approved, since initialising movement in X towards the flag.
 - But why negative?
 - Also, not towards enemy or other, because the relation cancels out? Since in the case of flag, position is fixed?
 - Enemy's X negative relation to Distance to Enemy. Human approved if indeed the enemy started from the left of the player, then it closes the distance.
 - Coin X has negative relation to Csequence0. It is whether the Coin (in this case id 0) was collided with first, right? Probably! Since coin collecting behavior and if more right, then maybe more often when the player moves on platform.
 - Reward positive relation to Collisions with Coins. Human approved since increasing the reward increases "the reasons for getting the reward" such as coins in coin collecting behavior.
 - Coin collisions negative relation reward. Makes sense after the collision occurred (values goes back down) the reward increases right?
 - Terminated negative relation with collision flag. Human approved in the sense that they certainly occur together.
 - Collision Flag negative relation with Csequence2. Interesting, so usually the third thing achieved? Makes sense right? Since there are always 2 coins!



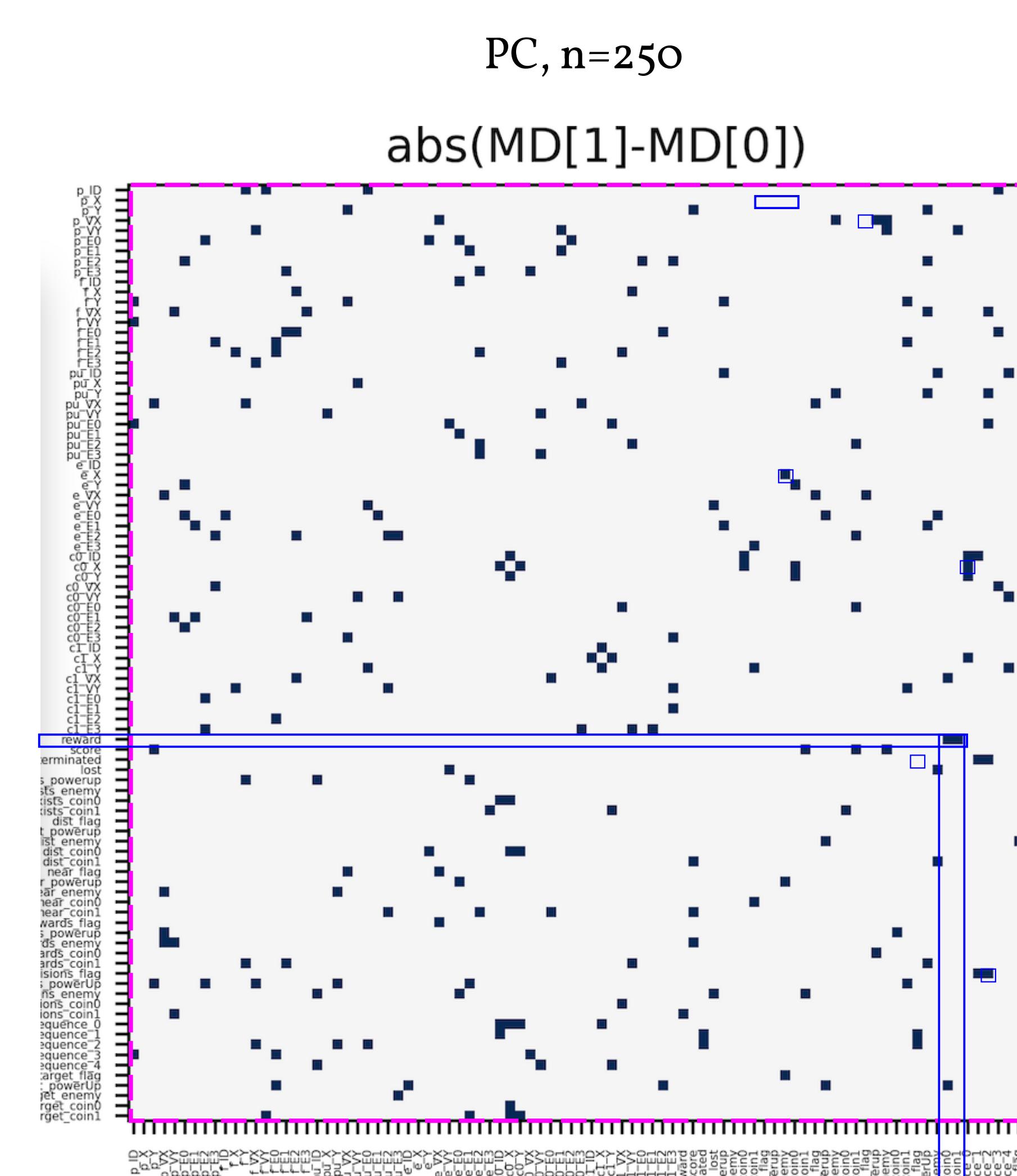
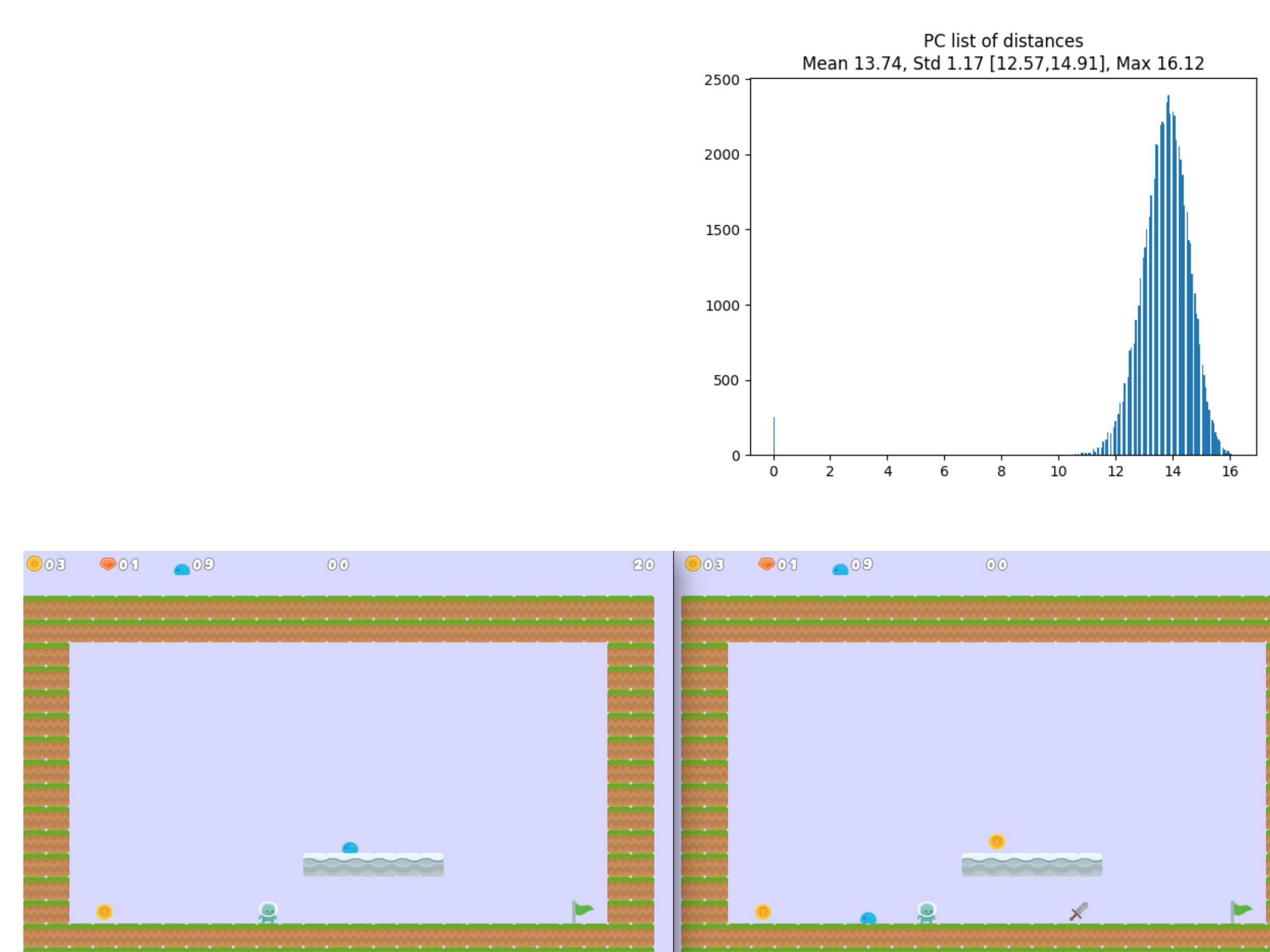
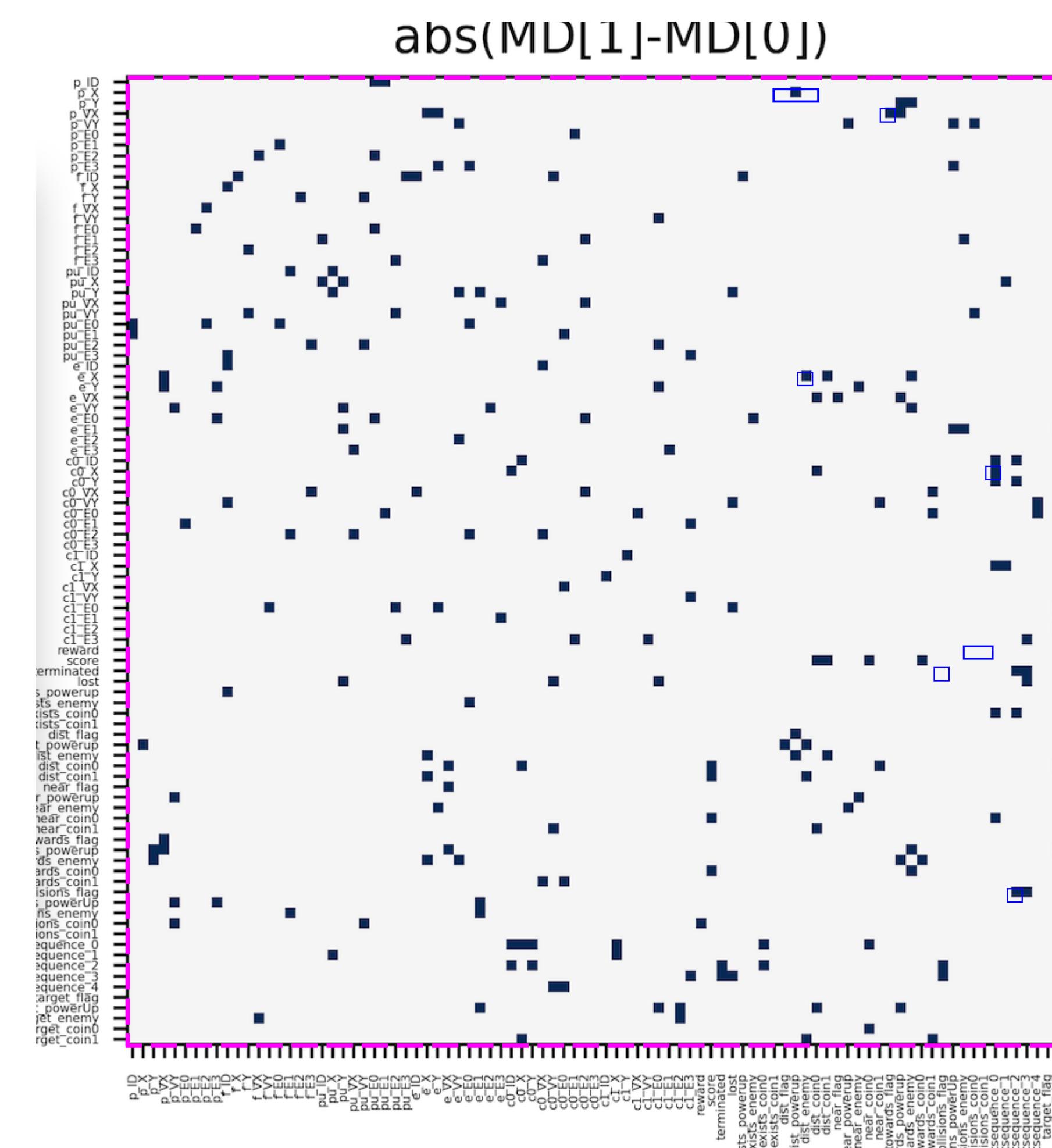
- FCI, n=5, avg, not checked the 5 recordings, looking only at stronger activations, comparing to PC n=5:
 - Completely blue (positive values).
 - More symmetric than PC (but not perfectly symmetric still).
 - Target powerup negative relation to towards powerup. Human approved since if powerup is not target anymore, then generally you don't go towards it!
- FCI, n=250, avg, not checked the 5 recordings, looking only at stronger activations, comparing to PC n=250:
 - All of the expected noise got kicked out, only the "core pattern" remain and they are in agreement with n=5.
 - Csequence relations vary, which makes sense. A stronger activation for flag as csequence2 makes sense since we collect coins, and we have two coins, so flag is 3rd - again, this confirms coin collecting behavior!
 - Distance powerup negative relation to distance flag. So on average the powerup is placed left from the player since going towards flag would be increasing powerup distance?
 - Like before.
 - CoinX associated with target Coin1. Human approved since the positioning relative to player will determine which one first to pick.



PC, n=5, Max-Diff, overlaid with the key relations from PC, n=5, avg:

- They don't agree on relation between player's X and distances to Y. Human approved since dynamics from level two level.
- Can we confirm here the case visually? Looking at 1650837120831 vs 1650836504662 (in coin collecting, sanity checked, the former is indeed the longer rollout; avoiding enemy) and then on platform, then flag.
- In 1650837120831, agent starts right, chase coin on far left (avoiding enemy), collects powerup (not present in others), other coin, avoids enemy. Human approved therefore, since half the data in the first rollout is about going left to decrease distance to enemy and coin.
- They agree on player's velocity in X relation towards flag. Human approved!
- They disagree on relation enemy's X to distance to enemy. See check above, human approved, since the player starts once right and once left from enemy, changing the notion of distance to enemy.
- They also disagree on relation coin_0 X with Csequence0. Seems approved since coin_0 X does not seem to change.
- They do agree on the collision relations between termination and flag, reward and coins.
- They also disagree on relation for flag collision with csequence2. Seems to make sense, because the second one involves a powerup and that is also collected!

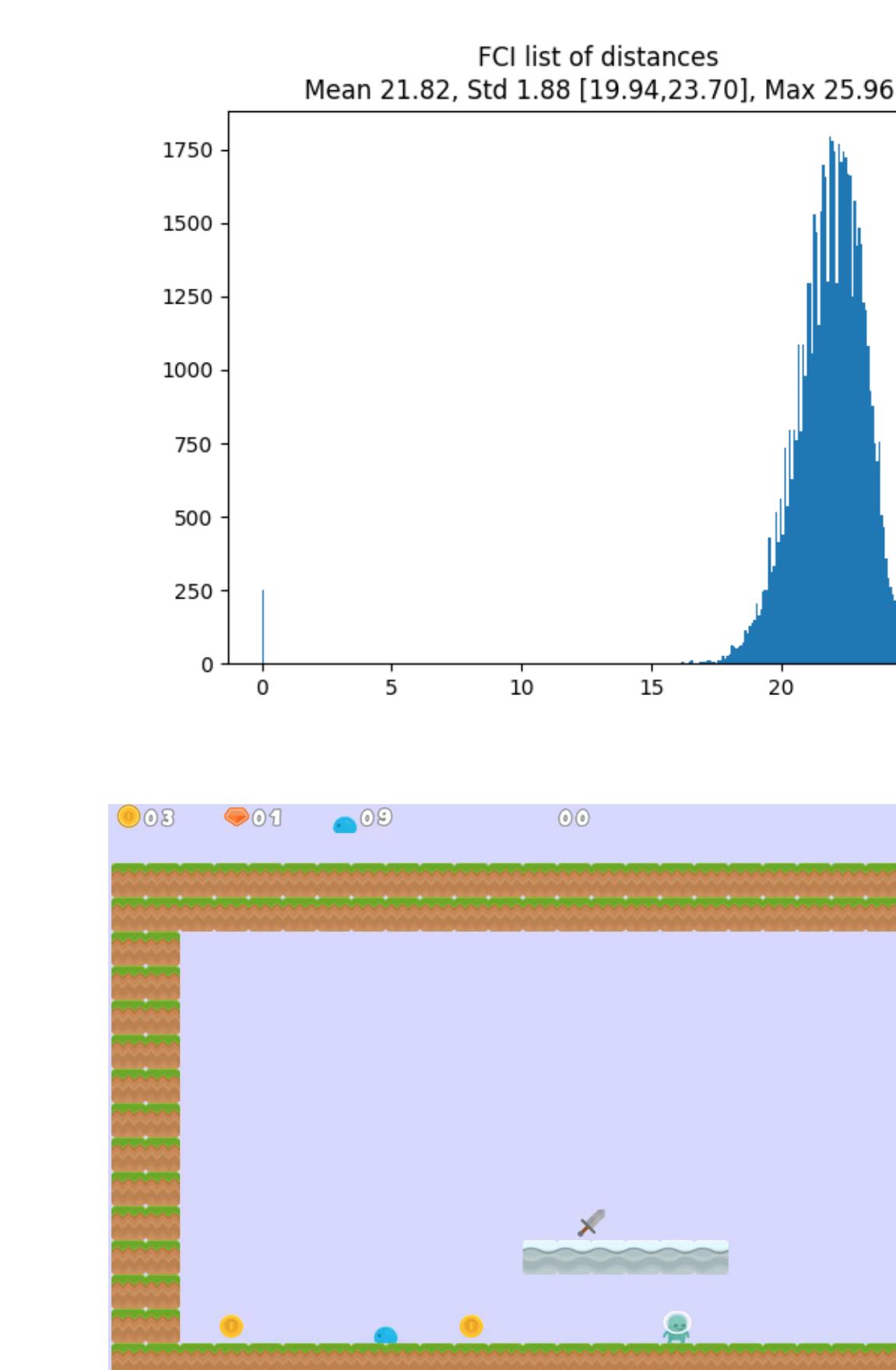
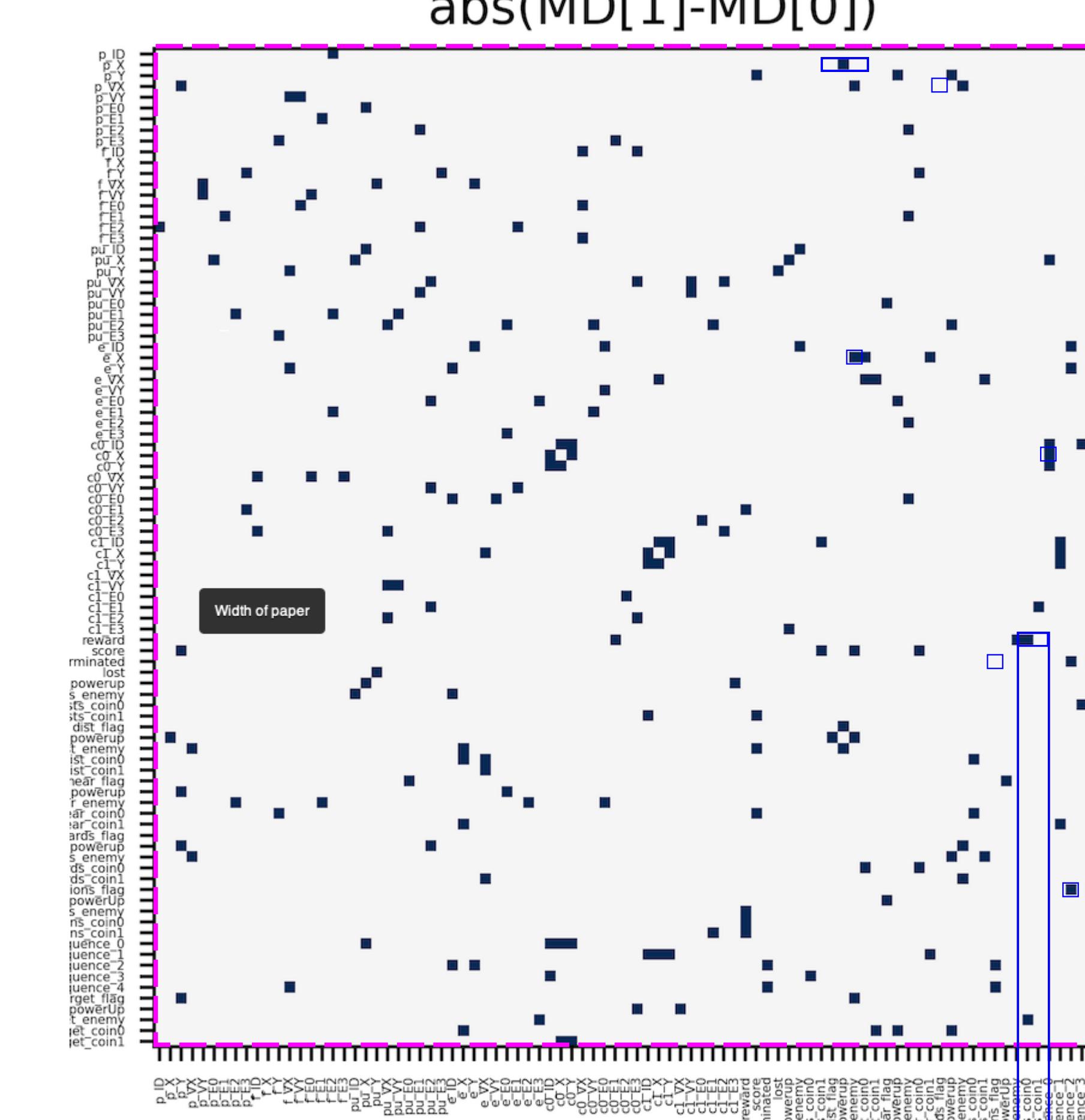
FCI, n=5



- PC, n=250, Max-Diff, overlaid with the key relations from PC, n=5, avg (which agreed with n=250):
 - ids 65 (165083604134), 223 (1650837586215) In 165083604134, agent starts left to platform, enemy on top, only one coin, no powerup, gets coin, then flag.
 - In 1650837586215, agent starts same but left is enemy, stealths behind enemy (killed enemy and no powerup collected), then gets second coin on platform, then flag.

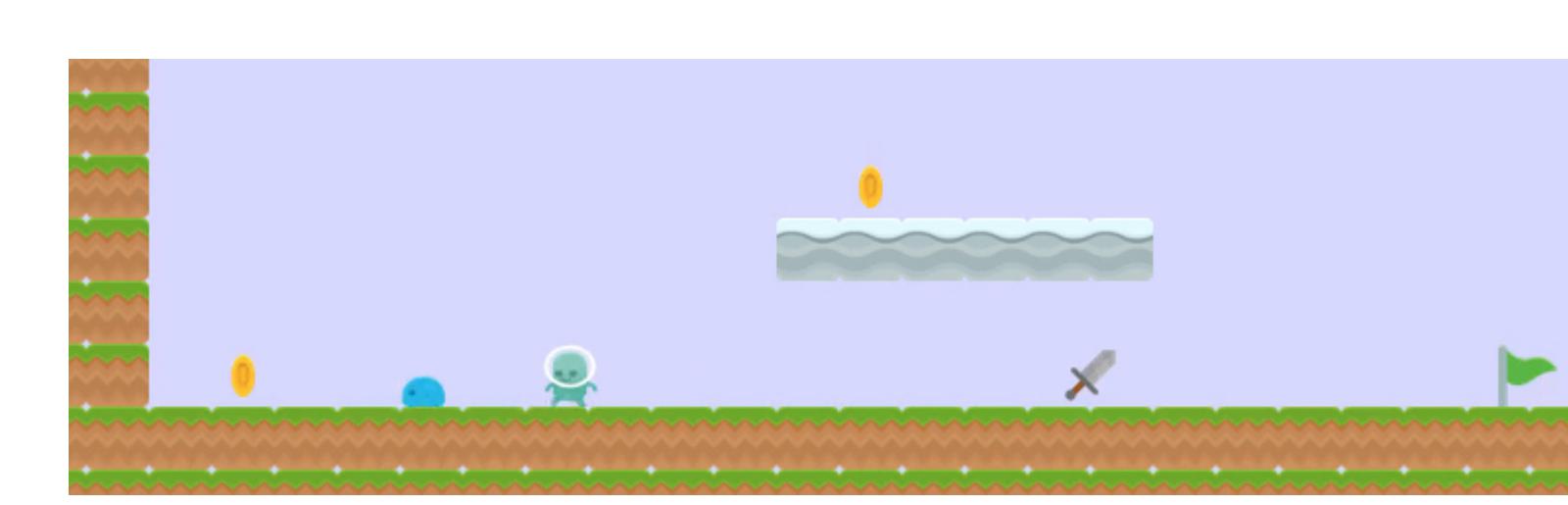
- Notable change: disagree on the coin collisions related to reward aspect. Human approved since difference in coins (maybe only the union coin, why it was not concerned), and also agreement on enemy, since neither kills enemy.

FCI, n=250



- Same, disagree on collisions and reward, but here FCI disagrees on coin0 with reward and enemy with reward whereas PC had no enemy but both coins disagreeing.

Human approved, in one enemy is killed, in other not.



Very interesting for explanation, the stealth behavior!

The player knows that the coin is guarded and cannot kill the enemy, that's why it waits until it has a brief moment of time to collect the coin while unguarded. Can we achieve this example if we maybe have a "guarded" property?

From n=5 only:

PC and FCI same. Avg. tells really something about relations. Max-Diff really tells something about the variance in the environment generation. 5 samples already enough to get a good grasp of existence of systemic relations. The found relations actually imply coin collecting behavior (maybe not uniquely though).

* On last point: player's X and object distances reveals active behavior, enemy X and distance reveals active objects, reward and coin reveals that coins were collected, csequence covering coins first shows that coins were prioritized, all coins appearing before suggests the focus on coins.

From n=250:

More samples, as expected, stabilizes results. More evidence for variability in environment captured by max-diff. Also, max-diff captures "the fundamental problem of causal inference" by having a notion of "never experienced" (see above, e.g., never killed enemy, so cannot know there is positive reward - *this is the fundamental problem of causal inference!*).