July 19th - July 28th

Progress Presentation

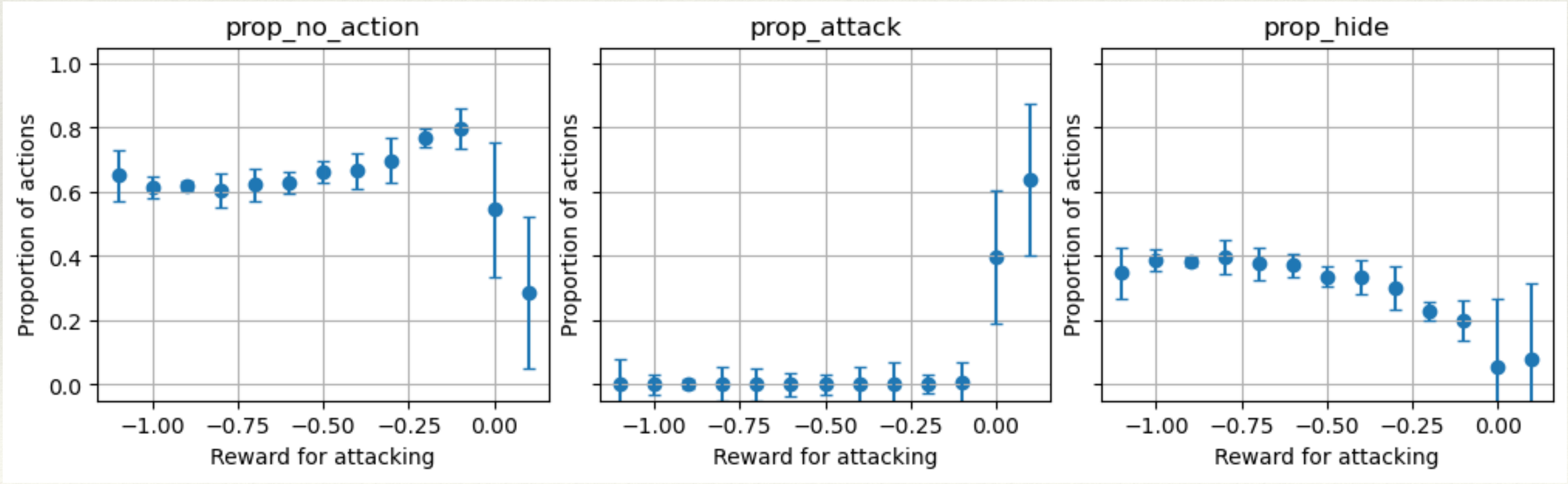
- results from experiment 1
- further experiments
- is the lowest forest level 0 or 1?

Experiment 1

- Independent variable: the reward for attacking another civilisation
  - values tested: -1.1, -1.0, …, -0.1, 0, 0.1 (13 values)
- Dependent variables:
  - proportions of different actions taken
  - attack streak length distribution
  - average reward per time step
  - action utility prediction error

Experiment 1

Proportions of Different Actions



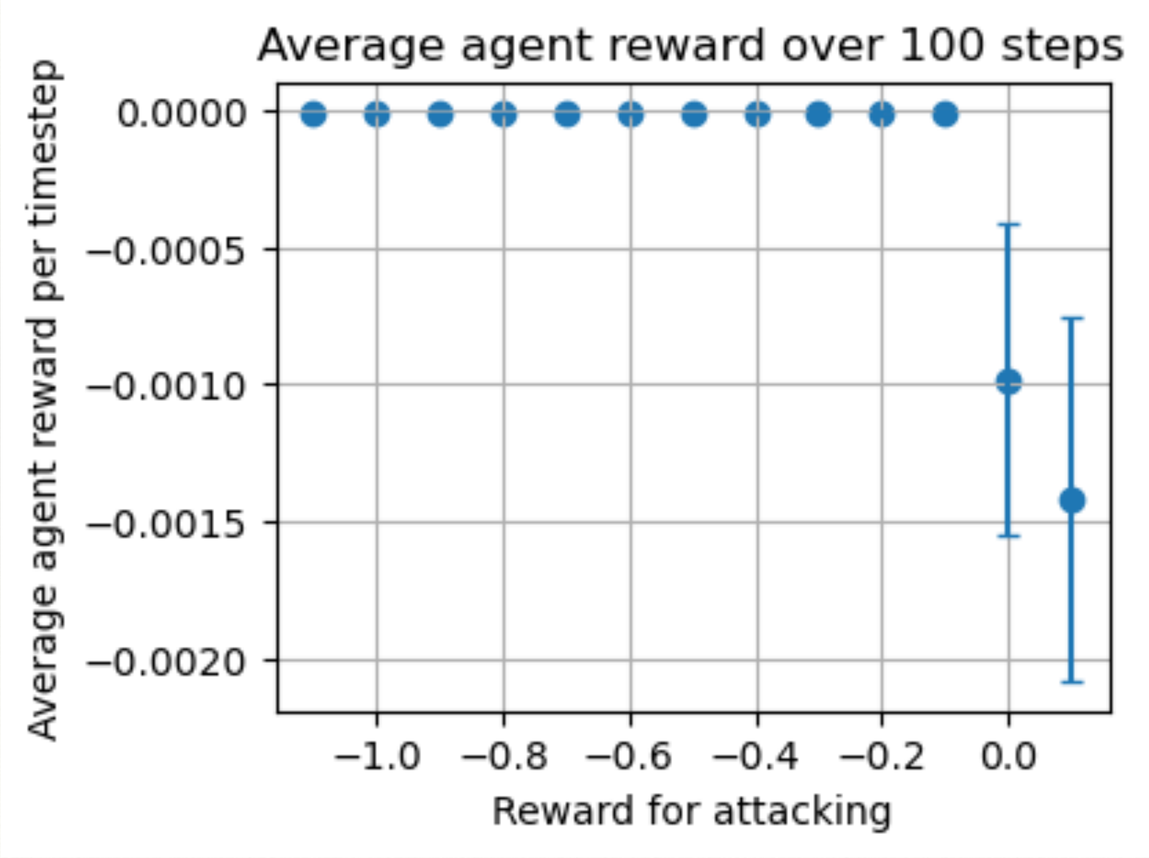
Experiment 1

Attack Streak Length Distribution

(Should do a few longer runs)

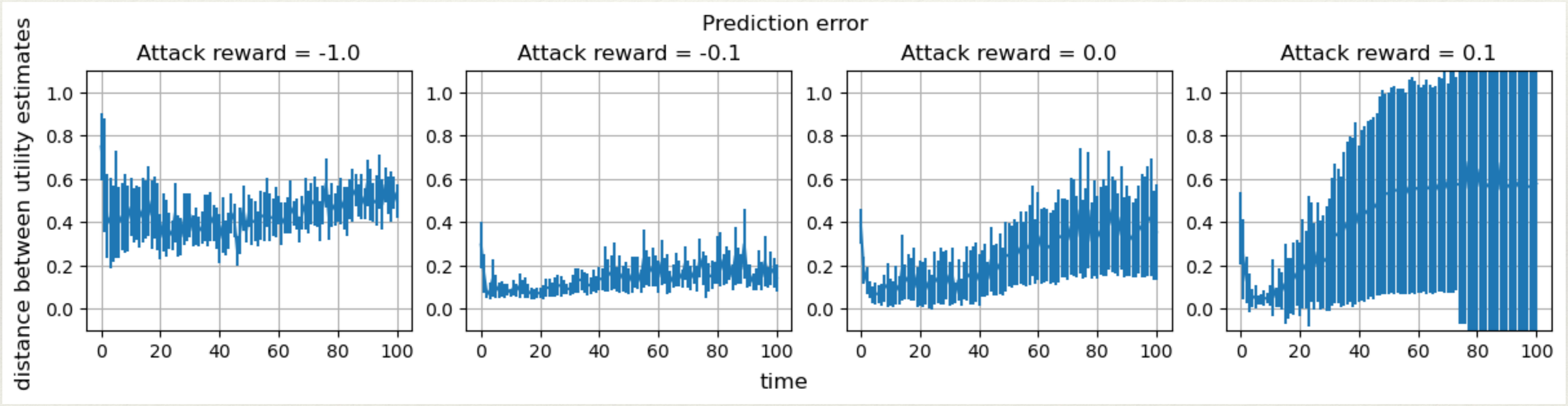
Experiment 1

Average Reward per Time Step



Experiment 1

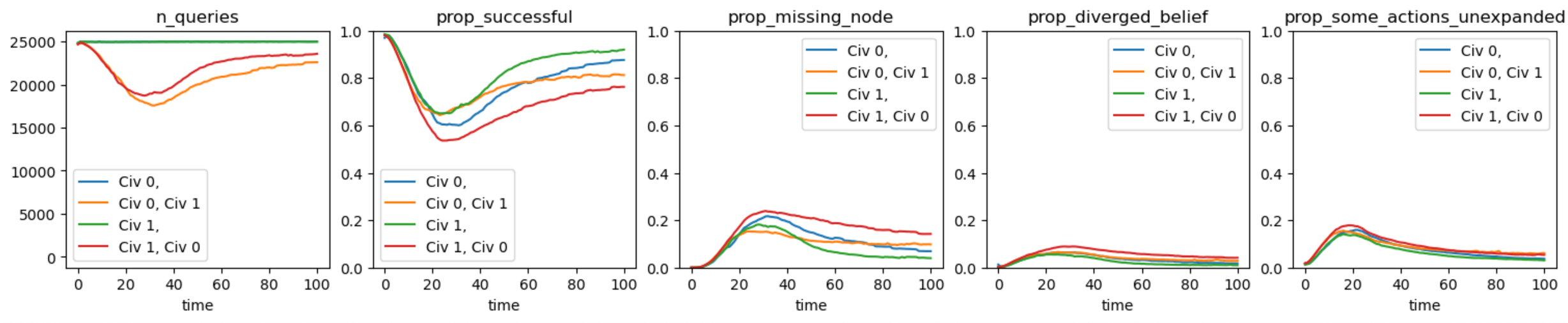
Prediction Error



prediction error of j predicting $i = \left\| \hat{U}(b_{i,L}, \cdot) - \sum_{is_{j,L}} b_{j,L}(is_{j,L}) \hat{U}(b_{(j,i),L-1}, \cdot) \right\|$ (Note: $b_{i,L}$ and $b_{j,L}$ are from separate I-POMDPs)

Experiment 1

Lower Forest Query Success Rate



Further Experiments

Which one to focus on?

1. investigating the **algorithm** (convergence of utility estimates, beliefs, …)

2. investigating the **optimal policy** computed by the algorithm (beliefs under which it is optimal to attack, …)

3. investigating the **resulting system** of agents acting according to the optimal policy (attack streak length distribution, …)

Is the lowest level forest level 0 or 1?

subintentional models. Second level beliefs are defined in terms of first level models and so on.
Formally, define spaces:

$$
\begin{array}{lllllllll}
IS_{i,0} & = & S, & \Theta_{j,0} & = & \{\langle b_{j,0}, \widehat{\theta}_j \rangle : b_{j,0} \in \Delta(IS_{j,0})\}, & M_{j,0} & = & \Theta_{j,0} \cup SM_j \\
IS_{i,1} & = & S \times M_{j,0}, & \Theta_{j,1} & = & \{\langle b_{j,1}, \widehat{\theta}_j \rangle : b_{j,1} \in \Delta(IS_{j,1})\}, & M_{j,1} & = & \Theta_{j,1} \cup M_{j,0} \\
 & & . & & & . & & & \\
 & & . & & & . & & & \\
 & & . & & & . & & & \\
IS_{i,l} & = & S \times M_{j,l-1}, & \Theta_{j,l} & = & \{\langle b_{j,l}, \widehat{\theta}_j \rangle : b_{j,l} \in \Delta(IS_{j,l})\}, & M_{j,l} & = & \Theta_{j,l} \cup M_{j,l-1}
\end{array}
$$

Definition 4 (Finitely Nested I-POMDP) *A finitely nested I-POMDP of agent i, I-POMDP$_{i,...}$ is:*

"Subintentional models", e.g. uniform random

Level 0: the interactive states hold a state. The other agents taking random actions creates a POMDP for the level 0 agent (where only the agent itself acts) in which the transition and observation functions are different. But rewards also become random

Level 1: the level 1 interactive states hold a state and implicitly the subintentional model for other agents (that they act randomly)

Notes from Meeting

- Sudden jump in number of attacks when attack_reward = 0: could be evidence of a “phase transition” / criticality
- slope of attack streak length distribution is of interest
- behaviour of prediction error for attack_reward=0.1 seems odd — check calculations
- likewise, behaviour of lower forest query success rate seems quite unusual
- to the question of whether lowest level is 0 or 1: 0 is intuitively preferable, but use your judgment
- **Experiment 2**: more agents, don’t vary attack reward (keep at 0). Do longer runs: run until the attack streak length distribution doesn’t change significantly
- **Experiment 3**: used to confirm the results of experiment 1 on a more global level.
  - Vary reasoning level (0 or 1), reward (but in a smaller, more interesting range), softargmax coefficient (determine range) and discount coefficient.
  - Do a Saltelli sample (no 2nd order terms), maybe 128 samples at first (can do more later if needed)
  - Plot same as experiment 1, but show average over all runs (letting the other parameters vary)