Progress Presentation

27.2. - 13.3.

Agenda

First working version of the model (kind of)

Problems

Signalling game

Overview

- model runs in steps
- at the beginning of each round, all agents receive a noisy technology signal which they use for:
  - updating technology beliefs
  - detecting if someone has been destroyed. If so, they update hostility beliefs
- after belief updates, one agents is randomly chosen and gets to choose an action from $\{$attack neighbour 1, $\ldots$, attack neighbour m, hide, do nothing$\}$
  - attacks are successful if attacker is stronger, or if equal, with 50% probability
  - hiding decreases agent visibility by 30% (but can be repeated)

Tech belief updates

- this is far from final!
- technology beliefs are represented with binomial distributions (≈ “discrete normal distribution”)
- observation = actual * visibility factor + Gaussian noise
- new belief distribution is Bin(n=10, p=observation)

Detecting a destroyed civilisation

- for each neighbour j, estimate (using MC) the probability $\mathbb{P}(\hat{t}_j^{(i)} < \min(t_j^{(i)}, 0.1))$
- the one with highest such probability (exceeding 90%) is deemed destroyed, and hostility belief update is initiated

Hostility belief updating
1/2

- H_j^{past} : j has been hostile in the past
- H_j^{now} : j was hostile this round
- X : x was destroyed this round
- $h_j^{(i)}$ is i's current hostility belief regarding j

- $c_j^{(i)} = \mathbb{P}[t_j^{(i)} > \max(t_x^{(j)}, r^{-1}(d(x,j)))]$ (currently estimated with MC)

Hostility Belief Updating

2/2

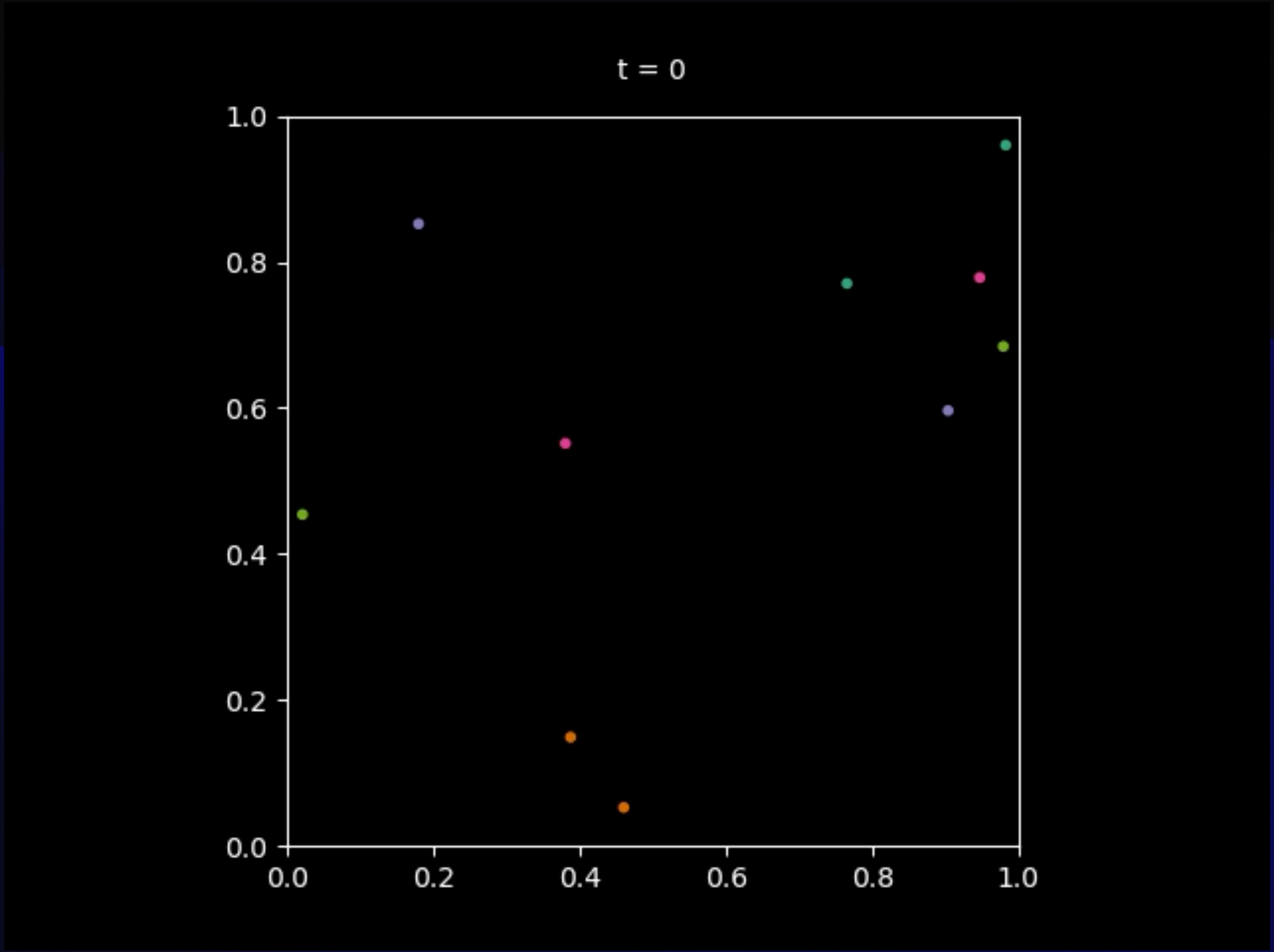
- $\mathbb{P}(H_j^{\text{now}} \mid X) = \dfrac{h_j^{(i)} c_j^{(i)}}{\sum_k h_k^{(i)} c_k^{(i)}}$
  (assuming hostility and capability are independent)

- update:
  $$\hat{h}_j^{(i)} = \mathbb{P}(H_j^{\text{past}} \cup H_j^{\text{now}} \mid X)$$
  $$= \mathbb{P}(H_j^{\text{past}} \mid X) + \mathbb{P}(H_j^{\text{now}} \mid X) - \mathbb{P}(H_j^{\text{past}} \cap H_j^{\text{now}} \mid X)$$
  $$\approx \mathbb{P}(H_j^{\text{past}}) + \mathbb{P}(H_j^{\text{now}} \mid X) - \mathbb{P}(H_j^{\text{past}})\mathbb{P}(H_j^{\text{now}} \mid X)$$
  $$= h_j^{(i)} + \frac{h_j^{(i)} c_j^{(i)}}{\sum_k h_k^{(i)} c_k^{(i)}} - h_j^{(i)} \frac{h_j^{(i)} c_j^{(i)}}{\sum_k h_k^{(i)} c_k^{(i)}}$$
  - this assumes hostility now and in the past are independent (given X)

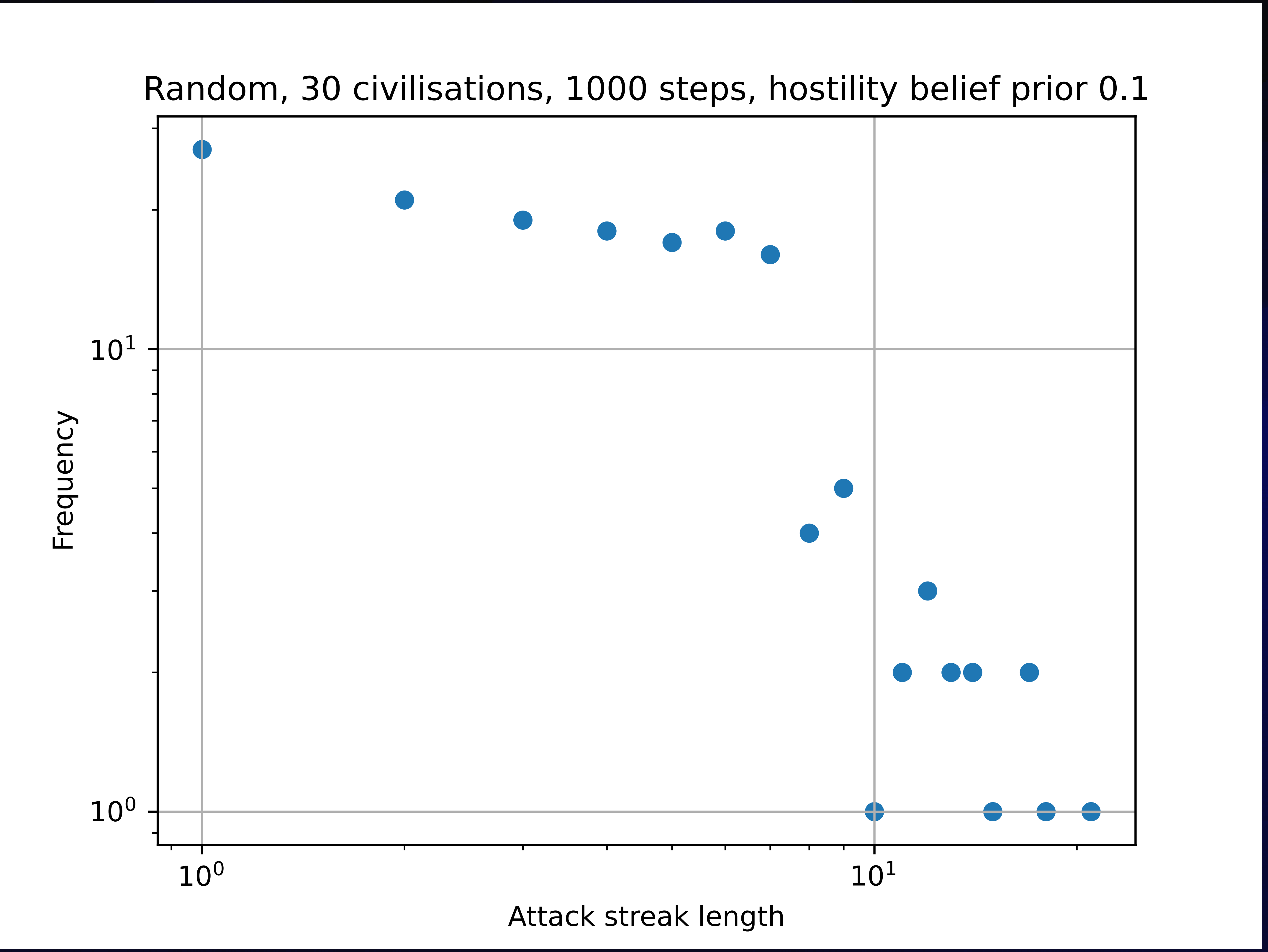
Actions

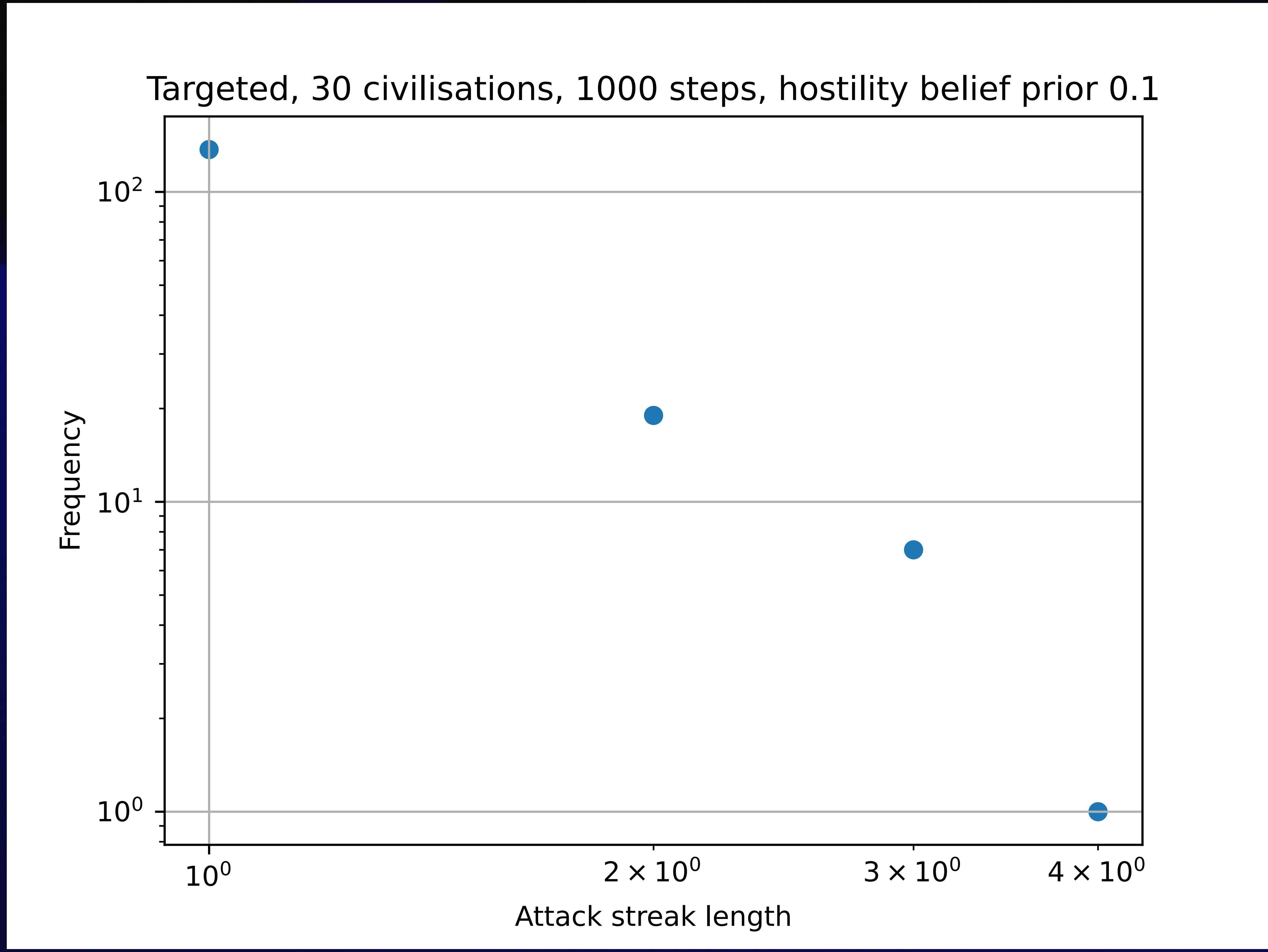
- goal: agents generate MAIDs and calculate equilibria
- so far: random action or ”targeted”
  - targeted: if agent is sure that one or more neighbours are hostile, attack the one where chance of success (estimate that we are stronger than that neighbour) is the highest. Otherwise attack random neighbour with 10% probability (“slip”) or randomly choose either hide or skip

Initial results



Initial results





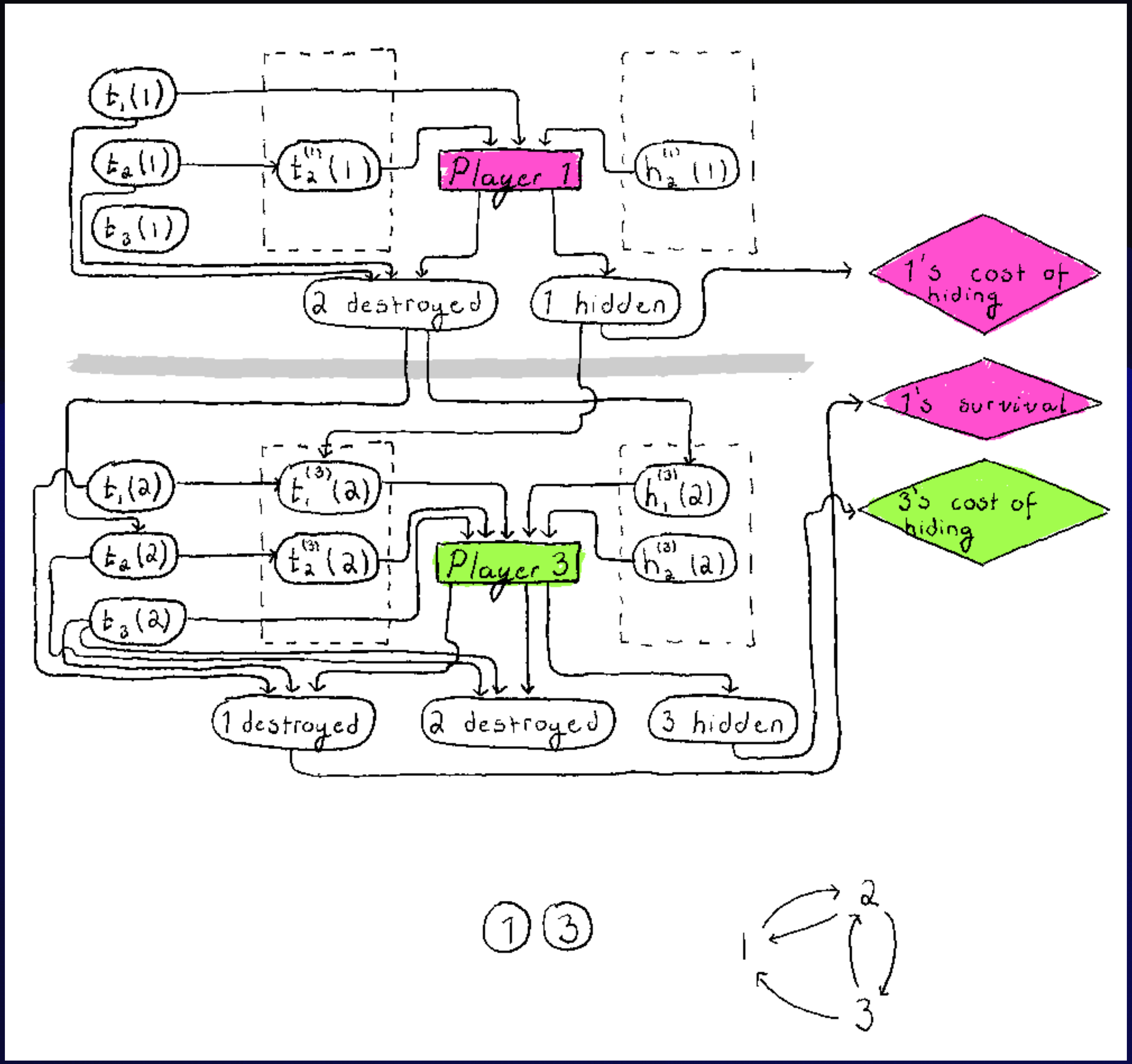
Problems

MAIDs and beliefs of others

- When agent constructs the MAID, it needs to also have beliefs about others' beliefs

- Previous actions can influence the structure of the MAID down the line

    - for example, if 1 destroyed 3 in its first round (if it could), then 3 would not be able to observe anyone else



Journal of Artificial Intelligence Research 24 (2005) 49-79 Submitted 09/04; published 07/05

A Framework for Sequential Planning in Multi-Agent Settings

Piotr J. Gmytrasiewicz PIOTR@CS.UIC.EDU
Prashant Doshi PDOSHI@CS.UIC.EDU
Department of Computer Science
University of Illinois at Chicago
851 S. Morgan St
Chicago, IL 60607

Abstract

This paper extends the framework of partially observable Markov decision processes (POMDPs) to multi-agent settings by incorporating the notion of agent models into the state space. Agents maintain beliefs over physical states of the environment and over models of other agents, and they use Bayesian updates to maintain their beliefs over time. The solutions map belief states to actions. Models of other agents may include their belief states and are related to agent types considered in games of incomplete information. We express the agents' autonomy by postulating that their models are not directly manipulable or observable by other agents. We show that important properties of POMDPs, such as convergence of value iteration, the rate of convergence, and piece-wise linearity and convexity of the value functions carry over to our framework. Our approach complements a more traditional approach to interactive settings which uses Nash equilibria as a solution paradigm. We seek to avoid some of the drawbacks of equilibria which may be non-unique and do not capture off-equilibrium behaviors. We do so at the cost of having to represent, process and continuously revise models of other agents. Since the agent's beliefs may be arbitrarily nested, the optimal solutions to decision making problems are only asymptotically computable. However, approximate belief updates and approximately optimal plans are computable. We illustrate our framework using a simple application domain, and we show examples of belief updates and value functions.

Signalling game?

- there is typically a sender and a receiver, and the sender has private information (the sender is either "good" or "bad"). The sender is trying to convince the receiver that the sender is of type "good". Cost of sending different signals depends on the type of the sender
- problem with applying this to my model
  - there are no civilisations of different types
  - the signals that civilisations would send have no information content, so they cannot differentiate between them.
    - But: later we could assume the civilisations can send different signals (but what are the costs?)