# Solving Graph-based Public Good Games with Tree Search and Imitation Learning

Victor-Alexandru Darvariu[1,2], Stephen Hailes[1], and Mirco Musolesi[1,2,3]

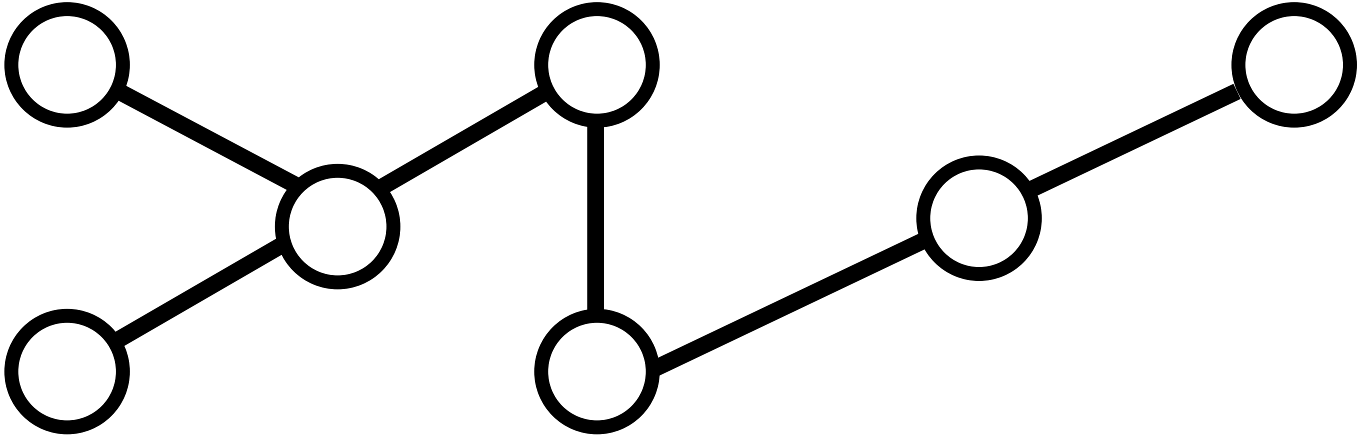
[1]University College London [2]The Alan Turing Institute [3]University of Bologna

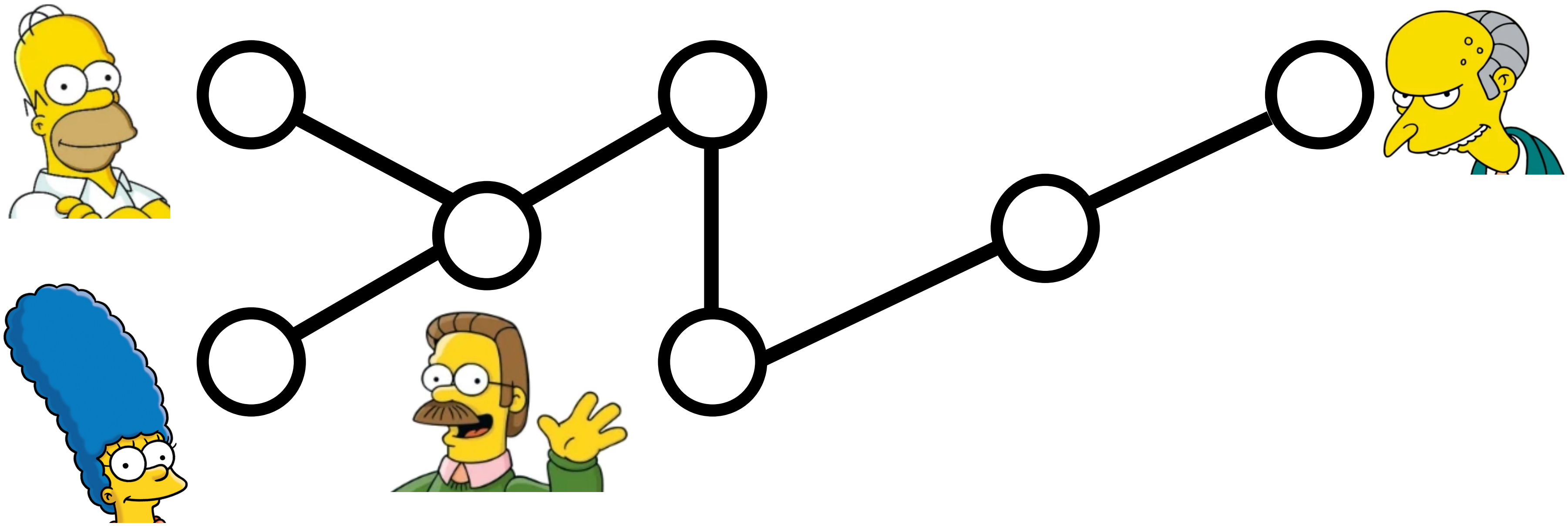
{v.darvariu, s.hailes, m.musolesi}@cs.ucl.ac.uk

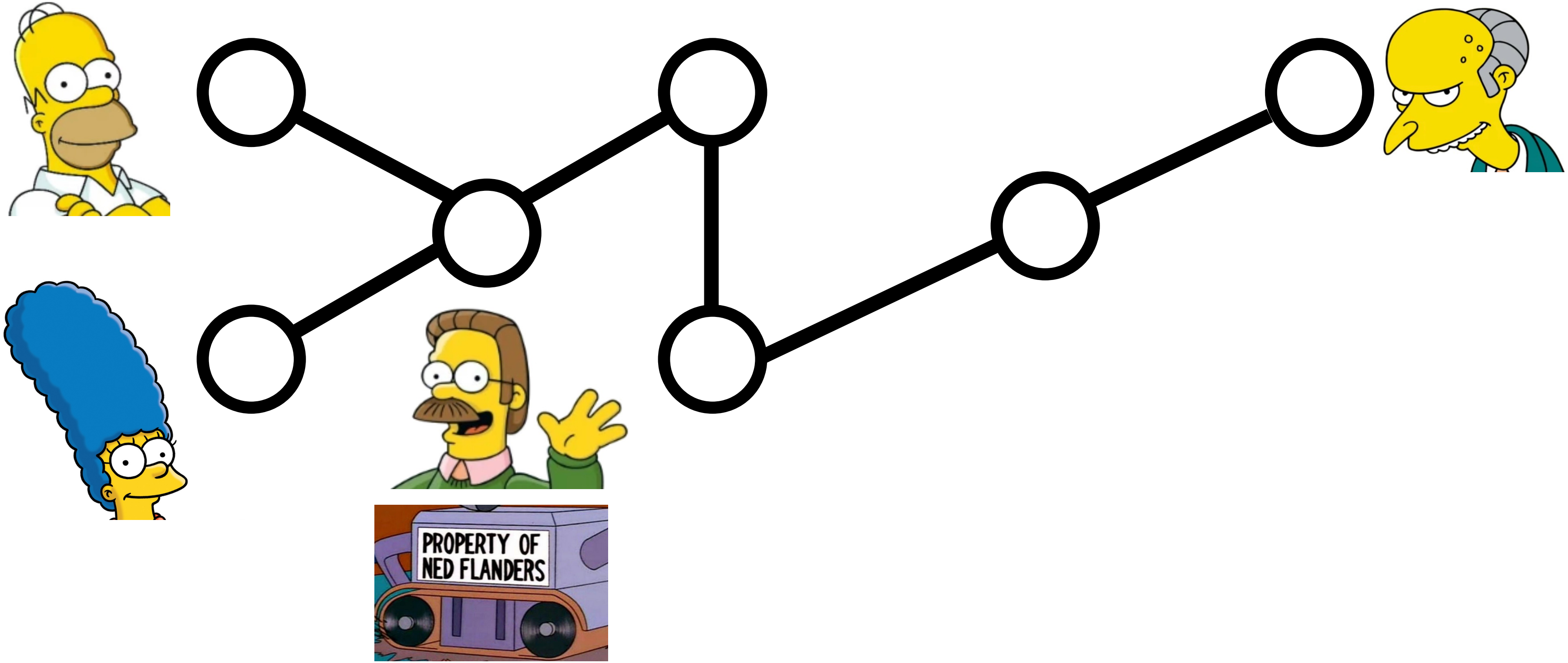
# Outline

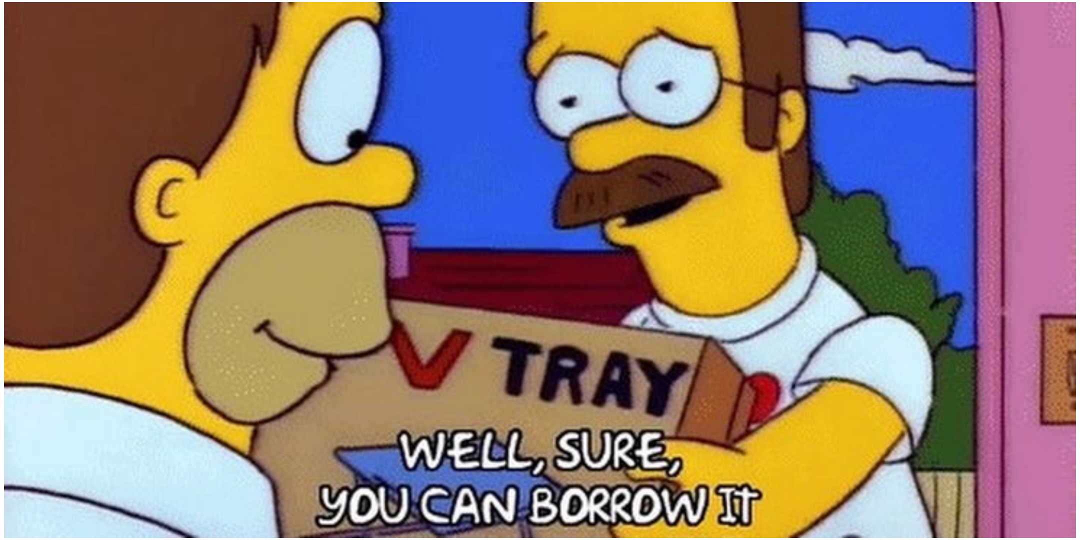
1. The Networked Best-Shot Public Goods Game
2. Our Approach for Finding Equilibria
3. Results & Discussion

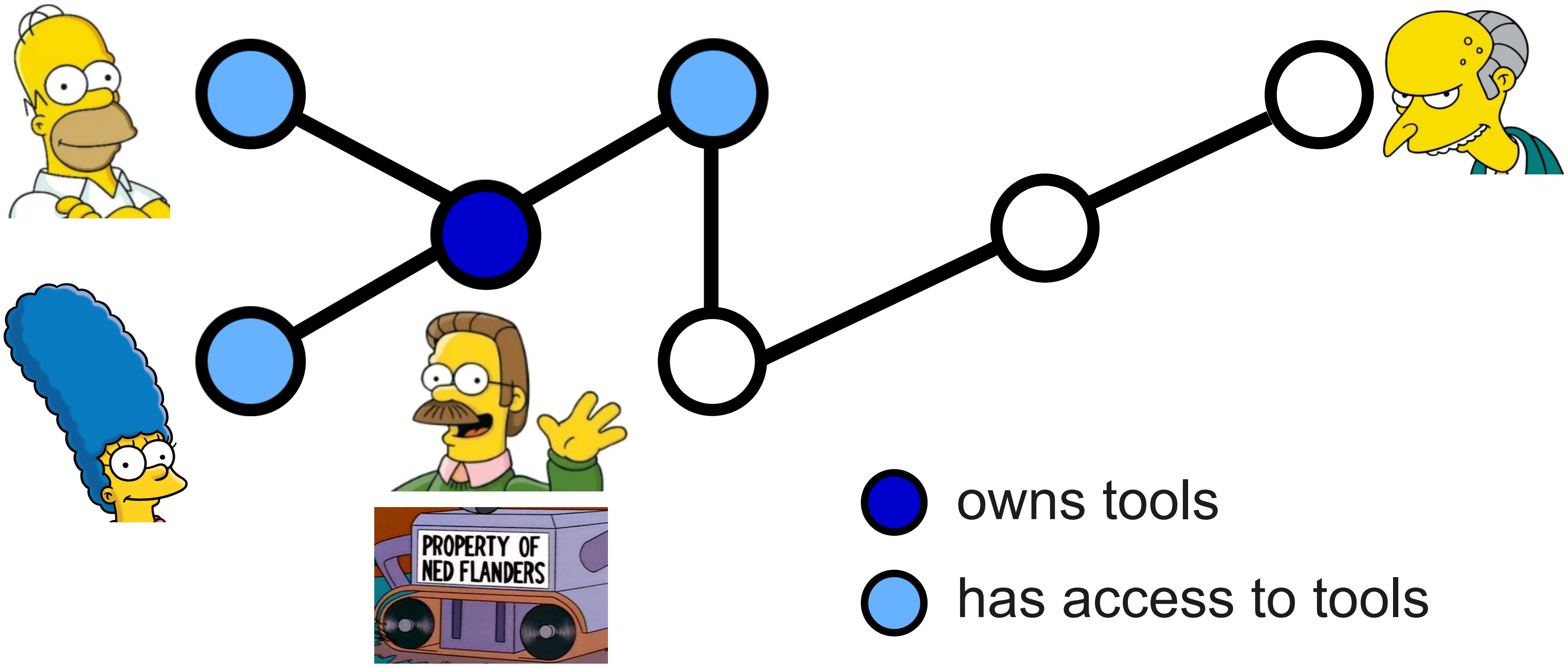
# The Networked Best-Shot Public Goods Game






PROPERTY OF
NED FLANDERS


TRAY
WELL, SURE,
YOU CAN BORROW IT


PROPERTY OF
NED FLANDERS
owns tools
has access to tools

# Public goods games (PGG)

- Form of n-party *social dilemma*
- Means of studying tensions between decisions that benefit only the individual vs. wider society
- Example applications:
    - Provisioning of public infrastructure & services
    - Dynamics of research & innovation
    - Meeting climate change targets

# Networked, best-shot PGGs

- *Networked*: impact of contributions limited along connections of a network
- *Best-shot*: utilities are binary and utility saturated if player or neighbour owns good

# Formally…

- Undirected, unweighted graph $G = (N, E)$
- Vertices $N = \{N_1, N_2, \dots N_n\}$ represent players
- Neighbourhood $\mathcal{N}_i = \{i\} \cup \{N_j \in N | (i, j) \in E\}$
- Action profile $\mathbf{a} = (a_1, \dots, a_n)$
- Acquiring good costs $c_i \in (0, 1)$, may differ between players

# Utilities and equilibria

- Utilities defined as:

$$u_i(\mathbf{a}) = \begin{cases} 1 - c_i, \text{ if } a_i = 1 \\ 1, \text{ if } a_i = 0 \ \wedge \ \exists j \in \mathcal{N}_i \ . \ a_j = 1 \\ 0, \text{ if } a_i = 0 \ \wedge \ \forall j \in \mathcal{N}_i \ . \ a_j = 0 \end{cases}$$

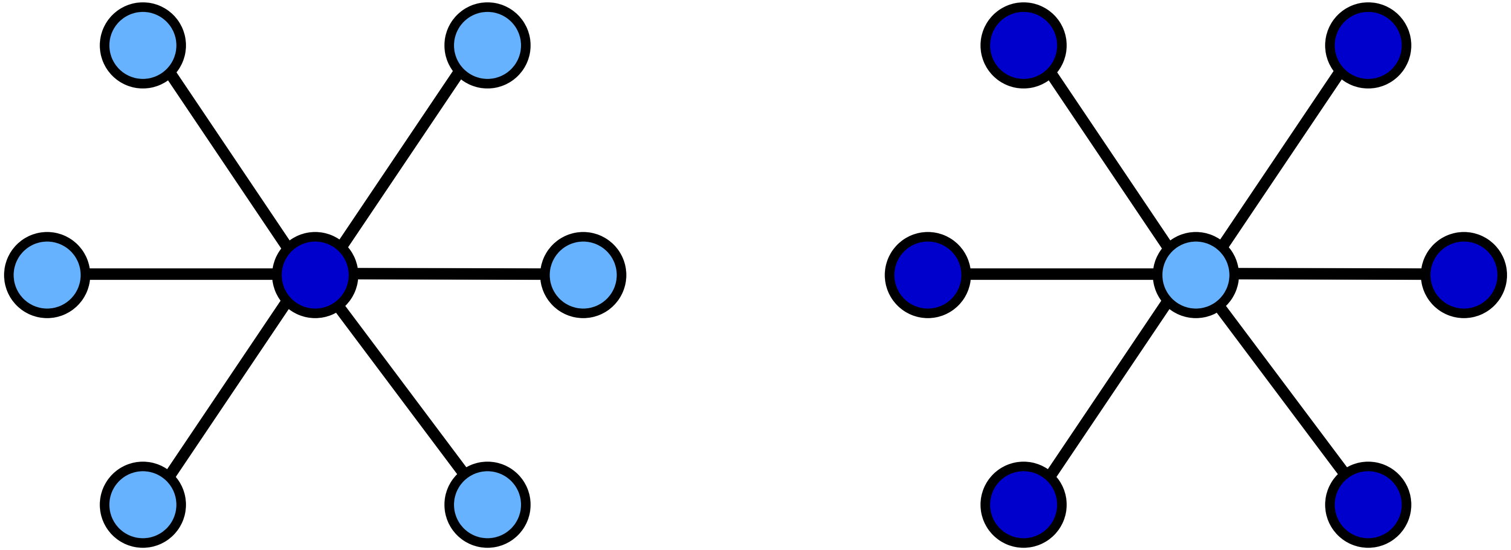
- Pure Strategy Nash Equilibria (PSNE):

$$u_i(a_i, \mathbf{a}_{-i}) \geq u_i(a_i^{'}, \mathbf{a}_{-i}) \ \forall i \in N, \ a_i^{'} \in A_i$$

# Finding equilibria

- *”What is an ideal outcome in this game?”*
- Equilibria correspond to Maximal Independent Sets (mIS) of graphs (Bramoullé & Kranton, 2007)
  - *independent set* is s.t. none of the vertices adjacent to each other
  - *maximal independent set:* IS not a proper subset of any other IS
- Finding *an* equilibrium (Jackson & Zenou, 2015)
  - Start with empty IS; incrementally add neighbours until IS is mIS
- Problem is NP-complete in general

# Finding equilibria



# Problem statement

- Given the set $\mathcal{E}$ of all PSNE and an objective function $f: \mathcal{E} \to [0,1]$
- Find PSNE profiles which satisfy $\operatorname{argmax}_{\mathbf{a} \in \mathcal{E}} f(\mathbf{a})$
- Example objectives: *social welfare* and *fairness*

$$SW(\mathbf{a}) = \frac{\sum_{i \in N} u_i(\mathbf{a})}{|N|} \qquad F(\mathbf{a}) = 1 - \frac{\sum_{i \in N} \sum_{j \in N} |u_i(\mathbf{a}) - u_j(\mathbf{a})|}{2n \sum_{j \in N} u_j(\mathbf{a})}$$

# Prior approaches

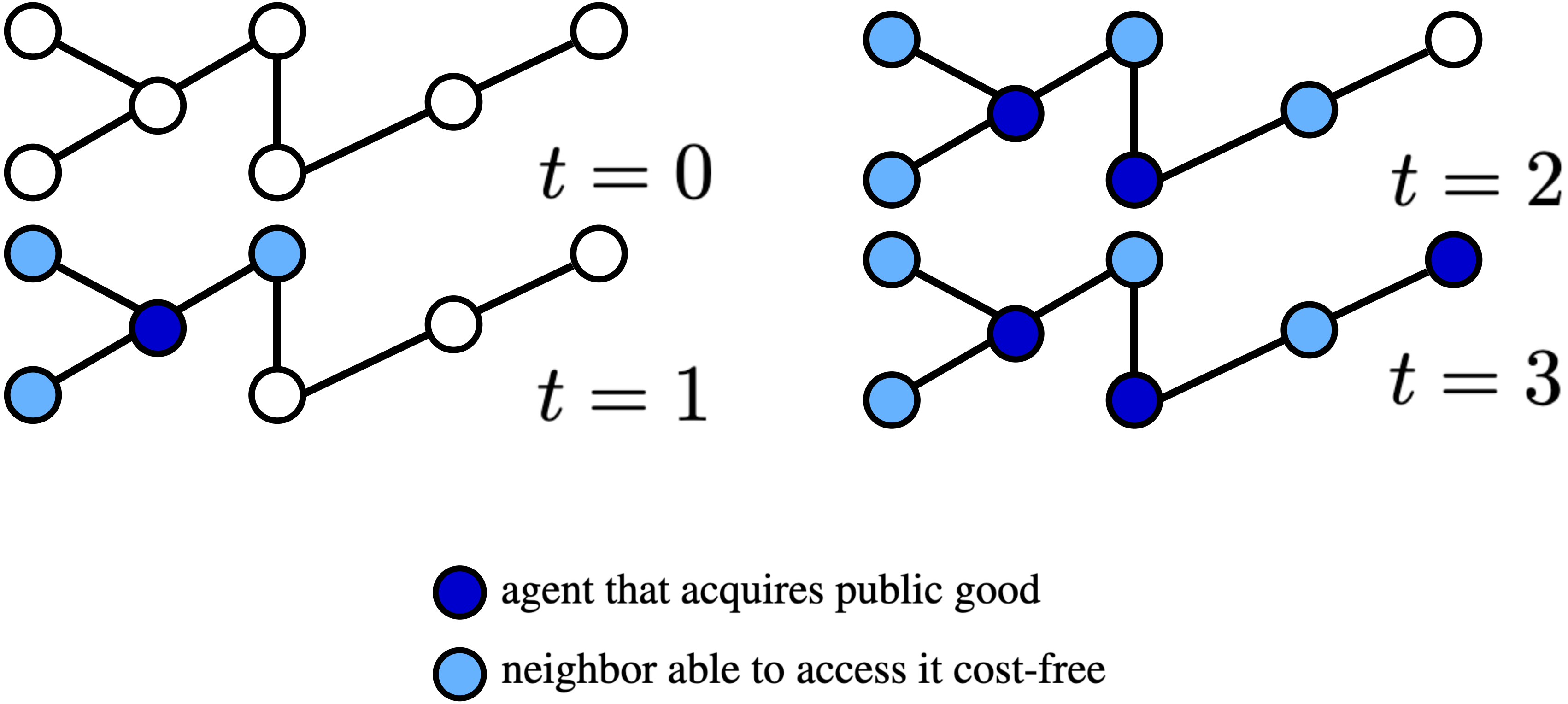
- Dall’Asta et al., 2011
  - Perturb configuration, play out the game to equilibrium
  - Accept new equilibrium according to simulated annealing rule
  - Ergodic Markov Chain, reaches optimal equilibrium in the limit
  - Approximate solution computationally feasible
- Levit al., 2018
  - Show general version of networked PGG is a potential game
  - Extend definition of utilities to include a payoff term
  - Players unhappy with outcome may convince neighbours to switch by offering a payoff (e.g., money)

# Our Approach for Finding Equilibria

# Our approach

1. Exploit connection with mIS property and formulate constructing an mIS as an MDP
2. Use Monte Carlo Tree Search to find optimal mIS using model of MDP
3. Collect a dataset of MCTS trajectories
4. Use dataset to train a GNN-parametrized policy by imitation learning

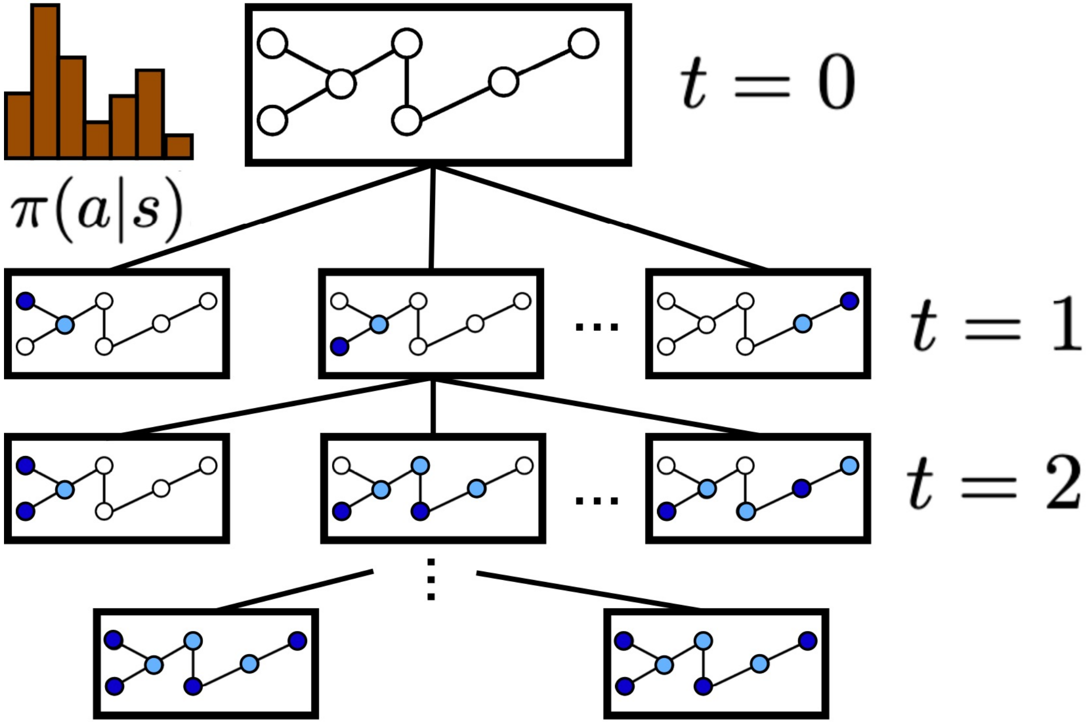
# 1. Formulating mIS construction as an MDP



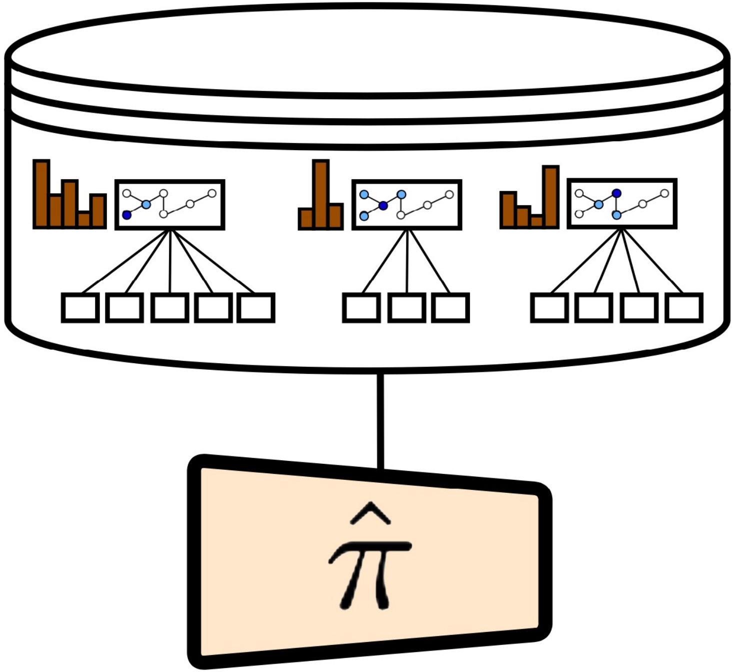
# 1. Formulating mIS construction as an MDP

- **State**: tuple $(G, I_t)$ formed of graph and IS
- **Action**: $\mathcal{A}_t = N \setminus \bigcup_{i \in I_t} \mathcal{N}_i$
- **Transitions**: deterministic; $I_t = I_{t-1} \cup \{a\}$
- **Rewards**: $f(\mathbf{a})$ at terminal states, 0 otherwise

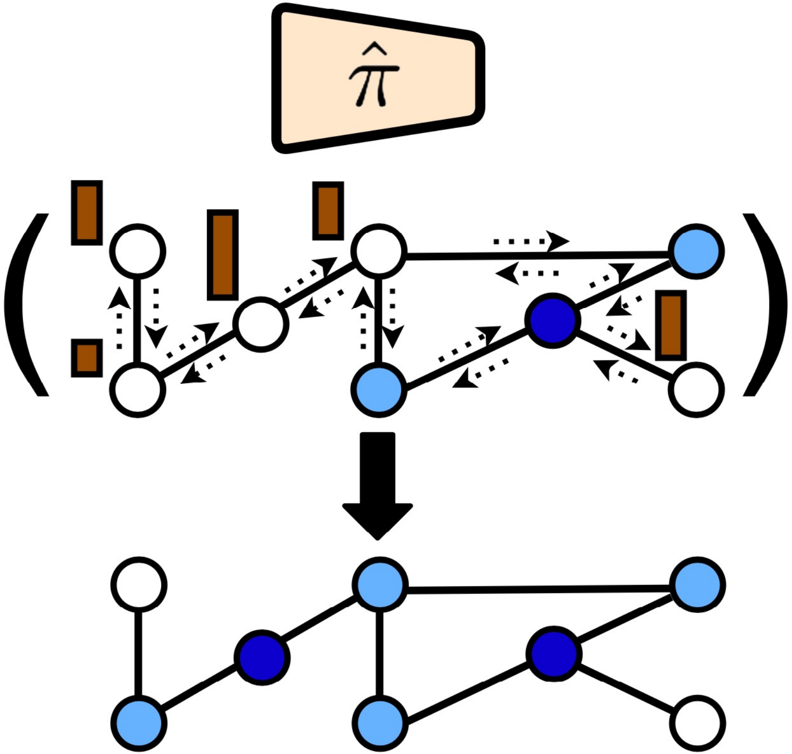
# 2. MCTS to search for optimal equilibria



# 3. Building a dataset of MCTS demonstrations



# 4. Imitation learning a GNN policy



# 4. Imitation learning a GNN policy

- Policy $\hat{\pi}$ parametrized by GNN (Dai et al. 2016)
- Outputs a *proto-action* $\phi(S_t)$
- Probabilities proportional to distance between proto-action and all available actions:

$$\hat{\pi}(A_t|S_t) = \frac{\exp(d(\mu_{A_t}, \phi(S_t))/\tau)}{\sum_{a \in \mathcal{A}(S_t)} \exp(d(\mu_a, \phi(S_t))/\tau)}$$

- Trained with KL loss: $\mathcal{L} = -\sum_{a \in \mathcal{A}(s)} \frac{C(s,a)}{C(s)} \log(\hat{\pi}(a|s))$

# Results & Discussion

# Experimental setup

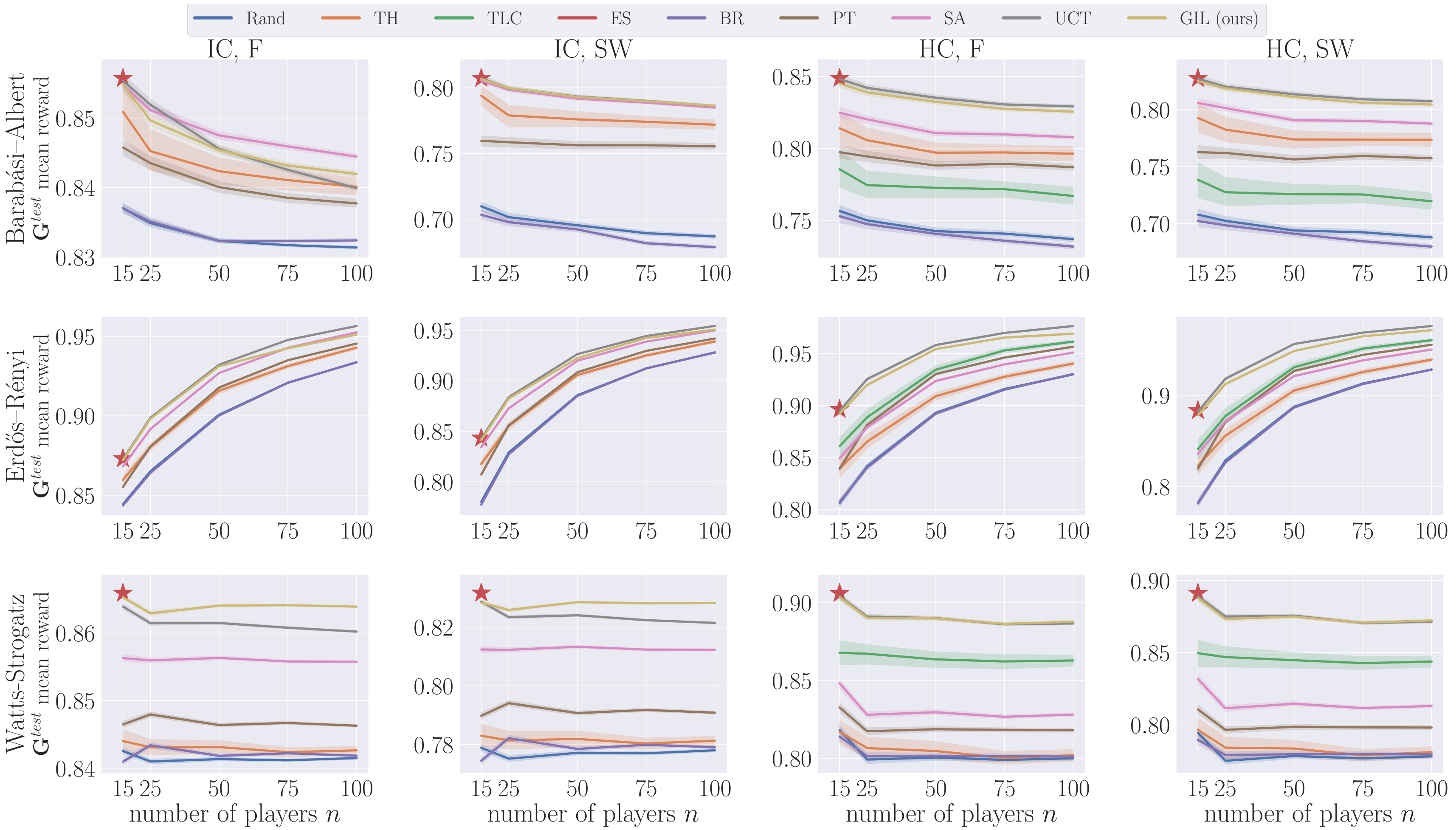
- Consider games with $n \in \{15, 25, 50, 75, 100\}$ players
- Take place over synthetic Barabási-Albert, Erdős-Rényi, Watts-Strogatz graphs
- Identical / heterogenous costs to acquire good (IC / HC)
- Constructing IL dataset: *separate, mixed, curriculum*
  - *separate*: only trajectories from same $n$
  - *mixed*: trajectories from all game sizes
  - *curriculum*: train in ascending order of $n$

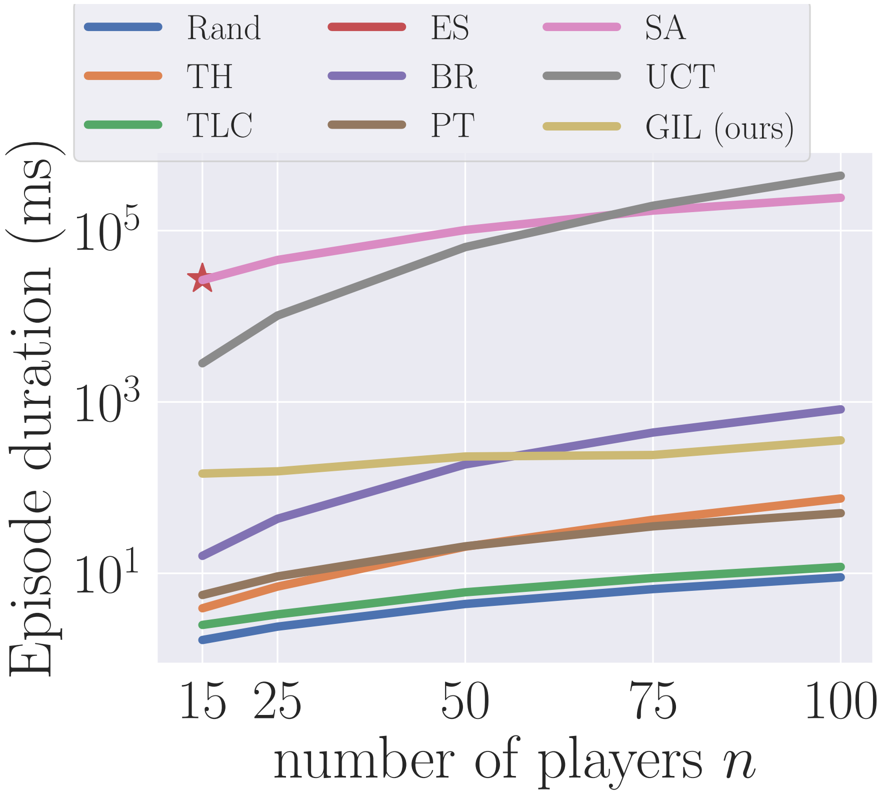
# Experimental setup

- Baselines:
    - **SA**: simulated annealing (Dall'Asta et al., 2011)
    - **PT**: payoff transfer (Levit et al., 2018)
    - **Random**: pick a mIS at random
    - **TH**: *target hubs* by placing public good on central nodes
    - **TLC**: place good on *lowest-cost* nodes in the network
    - **BR**: start from a random outcome, iteratively play *best response* until equilibrium reached
    - **ES**: *exhaustive search* over all action profiles (only applicable on very small graphs)

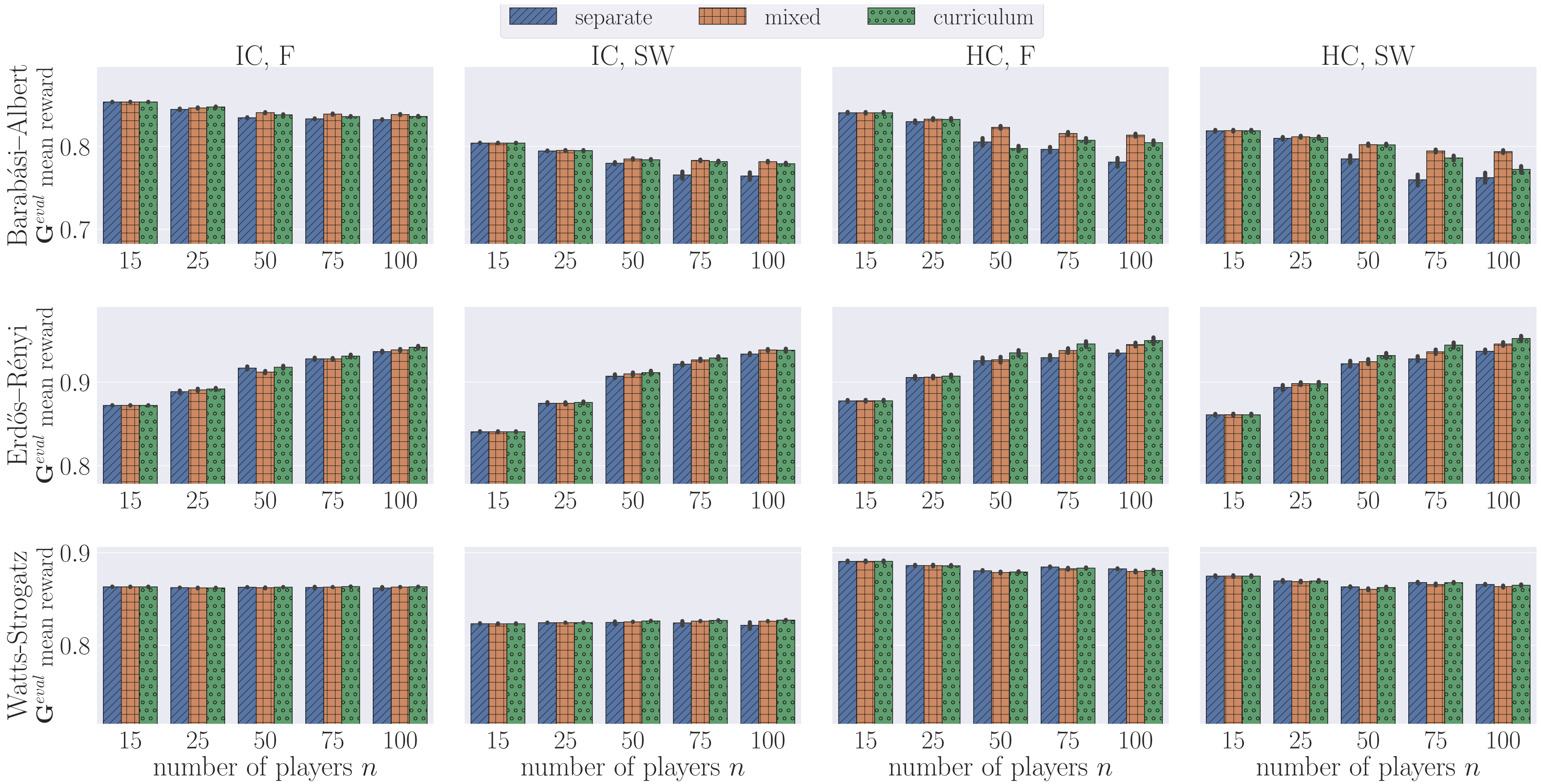
# Results

| c | G | $f$ | Random | TH | TLC | BR | PT | SA | UCT | GIL (ours) |
|---|---|---|---|---|---|---|---|---|---|---|
| HC | BA | F | 0.745$\pm$0.005 | 0.802 | 0.774 | 0.742$\pm$0.004 | 0.791$\pm$0.015 | 0.815$\pm$0.000 | **0.837**$\pm$0.000 | 0.834$\pm$0.001 |
| | | SW | 0.697$\pm$0.007 | 0.779 | 0.727 | 0.691$\pm$0.006 | 0.760$\pm$0.019 | 0.795$\pm$0.000 | **0.815**$\pm$0.000 | 0.813$\pm$0.000 |
| | ER | F | 0.877$\pm$0.001 | 0.896 | 0.920 | 0.877$\pm$0.000 | 0.911$\pm$0.002 | 0.908$\pm$0.001 | **0.945**$\pm$0.000 | 0.940$\pm$0.003 |
| | | SW | 0.868$\pm$0.001 | 0.890 | 0.912 | 0.867$\pm$0.000 | 0.903$\pm$0.002 | 0.903$\pm$0.001 | **0.940**$\pm$0.000 | 0.935$\pm$0.001 |
| | WS | F | 0.803$\pm$0.002 | 0.806 | 0.865 | 0.804$\pm$0.002 | 0.821$\pm$0.003 | 0.832$\pm$0.001 | **0.892**$\pm$0.000 | **0.892**$\pm$0.000 |
| | | SW | 0.781$\pm$0.002 | 0.785 | 0.846 | 0.782$\pm$0.003 | 0.800$\pm$0.004 | 0.817$\pm$0.001 | **0.876**$\pm$0.000 | **0.876**$\pm$0.000 |
| IC | BA | F | 0.833$\pm$0.000 | 0.844 | — | 0.834$\pm$0.000 | 0.841$\pm$0.005 | **0.849**$\pm$0.000 | 0.847$\pm$0.000 | 0.847$\pm$0.000 |
| | | SW | 0.697$\pm$0.007 | 0.779 | — | 0.691$\pm$0.006 | 0.757$\pm$0.019 | 0.794$\pm$0.000 | **0.795**$\pm$0.000 | **0.795**$\pm$0.000 |
| | ER | F | 0.893$\pm$0.000 | 0.906 | — | 0.892$\pm$0.000 | 0.907$\pm$0.001 | 0.916$\pm$0.000 | **0.922**$\pm$0.000 | 0.919$\pm$0.002 |
| | | SW | 0.867$\pm$0.000 | 0.889 | — | 0.866$\pm$0.001 | 0.889$\pm$0.002 | 0.903$\pm$0.000 | **0.910**$\pm$0.000 | 0.908$\pm$0.001 |
| | WS | F | 0.842$\pm$0.001 | 0.843 | — | 0.842$\pm$0.001 | 0.847$\pm$0.001 | 0.856$\pm$0.000 | 0.862$\pm$0.000 | **0.864**$\pm$0.000 |
| | | SW | 0.777$\pm$0.002 | 0.782 | — | 0.779$\pm$0.003 | 0.791$\pm$0.004 | 0.813$\pm$0.001 | 0.824$\pm$0.000 | **0.828**$\pm$0.000 |


Rand
TH
TLC
ES
BR
PT
SA
UCT
GIL (ours)
IC, F
IC, SW
HC, F
HC, SW
Barabási–Albert
Erdős–Rényi
Watts-Strogatz
$G^{test}$ mean reward
number of players $n$



Rand
ES
SA
TH
BR
UCT
TLC
PT
GIL (ours)
Episode duration (ms)
$10^5$
$10^3$
$10^1$
15
25
50
75
100
number of players $n$


separate
mixed
curriculum
IC, F
IC, SW
HC, F
HC, SW
Barabási–Albert
$G^{eval}$ mean reward
Erdős–Rényi
$G^{eval}$ mean reward
Watts-Strogatz
$G^{eval}$ mean reward
0.7
0.8
0.9
15
25
50
75
100
number of players $n$

# Summary of results

- Finds equilibria of higher social welfare and fairness than previous methods
  - Difference more substantial when costs differ between players
- IL policy preserves performance while 3 orders of magnitude cheaper to evaluate
- Best method for dataset construction depends on underlying network structure
  - BA: mixed; ER: curriculum; WS: no significant difference

# Outlook

- Related to ongoing efforts to study cooperation in multi-agent systems (Dafoe et al., 2020)
- While we consider a game theory application, method applies to maximal independent sets in general
  - see, e.g., Dall’Asta et al., 2009
- IL proto-action method of interest for graph combinatorial optimization and algorithmic reasoning
  - see, e.g., Cappart et al., 2021

# References


Bramoullé, Y., & Kranton, R. (2007). Public goods in networks. *Journal of Economic Theory*, *135*(1), 478–494.

Cappart, Q., Chételat, D., Khalil, E., Lodi, A., Morris, C., & Veličković, P. (2021). Combinatorial optimization and reasoning with graph neural networks. In: *IJCAI*.

Dafoe, A., Hughes, E., Bachrach, Y., Collins, T., McKee, K. R., Leibo, J. Z., Larson, K., & Graepel, T. (2020). Open Problems in Cooperative AI. *arXiv:2012.08630.*

Dai, Hanjun, Bo Dai, and Le Song (2016). "Discriminative Embeddings of Latent Variable Models for Structured Data". In: *ICML*.

Dall'Asta, L., Pin, P., & Ramezanpour, A. (2009). Statistical Mechanics of maximal independent sets. *Physical Review E*, *80*(6), 061136.

Dall'Asta, L., Pin, P., & Ramezanpour, A. (2011). Optimal Equilibria of the Best Shot Game. *Journal of Public Economic Theory*, *13*(6), 885–901.

Levit, V., Komarovsky, Z., Grinshpoun, T., & Meisels, A. (2018). Incentive-based search for efficient equilibria of the public goods game. *Artificial Intelligence*, *262*, 142–162.

Jackson, M. O., & Zenou, Y. (2015). Chapter 3—Games on Networks. In H. P. Young & S. Zamir (Eds.), *Handbook of Game Theory with Economic Applications* (Vol. 4, pp. 95–163). Elsevier.

# Solving Graph-based Public Good Games with Tree Search and Imitation Learning

Victor-Alexandru Darvariu, Stephen Hailes, Mirco Musolesi

https://arxiv.org/abs/2106.06762

v.darvariu@cs.ucl.ac.uk
@VictorDarvariu
https://victor.darvariu.me