

Reinforce Lupus

Cooperation through Emergent Communication in *The Werewolf* Social Deduction Game

Nicolo' Brandizzi ¹

Luca Iocchi ¹

Davide Grossi ²

¹ Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Sapienza University of Rome

² Bernoulli Institute for Maths, CS and AI, University of Groningen

Problem : Social Deduction Games

M opposing parties:

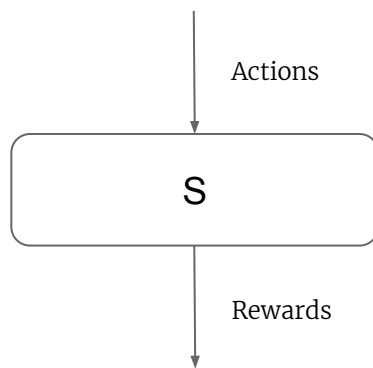
$$P^{(1)}, P^{(2)}, \dots, P^{(M)}$$

With m number of players:

$$P^{(m)} = \{p_1^{(m)}, p_2^{(m)}, \dots, p_n^{(m)}\}$$



Environment S
implementing the
game logic



Action Set A :

$$\{a_1, a_2, \dots, a_n\} = A$$

Unidirectional
communication
actions:

$$C_{i,j}(b) : p_j \rightarrow p_i$$

$$\forall p_j, p_i \quad j \neq i$$

Problem : Communication actions

Dedicated Communication policy π_C

$$\pi_C(C^{(t)}(b)) \rightarrow C^{(t+1)}(b)$$



Action policy π_a

$$\pi_a(a_t) \rightarrow a_{t+1}$$

Embedded communication

$$\pi(\{a_t, C^{(t)}(b)\}) \rightarrow \{a_{t+1}, C^{(t+1)}(b)\}$$

Methods: *The Werewolf Game*

Roles



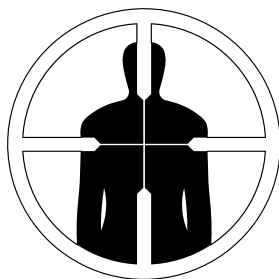
Phases



Methods: Communication actions

Target

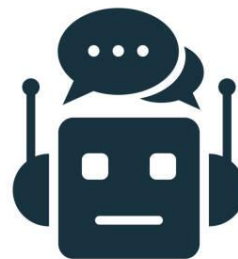
Each agent outputs an id in range $(0, \text{num_players})$ which is the agent they want to vote for, aka target.



Signal

A modular vector used for communication.

Signal Range [SR] and Signal Length [SL] decided by user.



Methods: Policies

Static Policies

Random: random target for each wolf

Random Unite: same random target for all wolves

Revenge: target villagers who voted for wolves or random

Trainable Policy

Proximal policy optimization [PPO]

$$\max_{\theta} E_t[L^{CLIP}(\theta) - c_1 L^{VF}(\theta) + c_2 S[\pi_{\theta}](s_t)]$$

Methods: Metrics

- **Vil Win:** villagers winning ratio.
- **Suicide:** average number of times a player vote for itself.
- **Average days:** total number of days before a match is over.
- **Accord :** percentage of agents voting for the same target.

Results : Nine players

Design Choice				Results			
Comm	Rnd	Unt	Rvg	Vil Win	Accord	Suicide	Days
0SL	X			0.044	0.478	0.086	1.55
0SL		X		0.03	0.695	0.059	1.5
0SL			X	0.12	0.482	0.078	1.64
1SL-2SR	X			0.19	0.47	0.078	1.58
1SL-2SR		X		0.08	0.685	0.055	1.52
1SL-2SR			X	0.4	0.479	0.065	1.9

Legend

Rnd: random
Unt : Unite
Rvg : revenge

SL: signal length
SR : signal range

Vil Win : villager's winning rate.
Accord : percentage of agents voting for the same target.
Suicide : number of agents voting for themselves.
Days : average number of day before a match ends.

Results : Twenty one players

Comm	Win	Suicide	Days	Accord
0SL	0.42	0.072	7.84	0.56
1SL-2SR	0.25	0.075	7.74	0.57
9SL-2SR	0.98	0.04	7.3	0.56
21SL-2SR	0.94	0.05	7.6	0.56
1SL-21SR	0.72	0.062	8	0.56
21SL-21SR	0.61	0.066	7.9	0.56

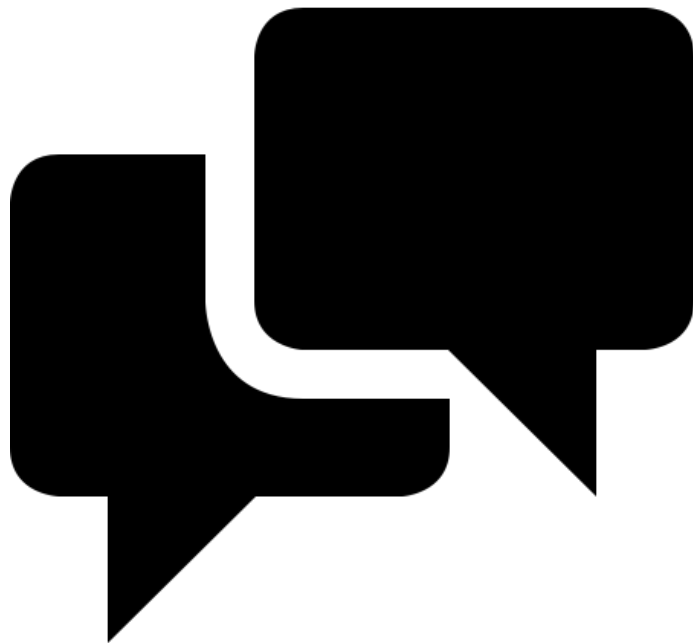
Legend

Rnd: random
Unt : Unite
Rvg : revenge

SL: signal length
SR : signal range

Vil Win : villager's winning rate.
Accord : percentage of agents voting for the same target.
Suicide : number of agents voting for themselves.
Days : average number of day before a match ends.

Conclusion



Reinforce Lupus

Cooperation through Emergent Communication
in *The Werewolf* Social Deduction Game

Thank you

Nicolo' Brandizzi
Luca Iocchi
Davide Grossi