

AIWolf - Werewolf playing Agent Competition

Challenge:

Build an agent that is able to build consensus with other actors (who are potentially deceptive), identify agents of the opposing team and coordinate to vote them out of the game.

Werewolf (also called mafia) is social game focusing on discussion and debate, which puts an uninformed majority team (villagers) against an informed minority team (werewolf). The goal of each team is to eliminate all the members of the opposing team. The main tool for both teams to achieve this goal is voting: At each game phase, all players must collectively vote on one player that will be removed from the game.

Agents in the villager team are the uninformed majority. If each villager know who the other villagers are, they could coordinate to form a majority voting block and win the game. However, they do not know which of the other agents are in the same team as themselves. Therefore they must communicate with other agents to build a coalition and find a consensus on who is the wolf.

Agents in the werewolf team are the informed minority. They know the alignment of every agent (but not their exact roles). However, they can be outvoted if their own team membership is revealed. Therefore they must pretend to be a villager and lead the consensus towards the removal of other villager-aligned players.

The game progresses alternately through two phases: Day and Night. During the day Phase, the agents communicate to exchange arguments and build a consensus on which player is most likely to be werewolf-aligned. At the end of the day phase, a voting takes place, and one of the agents is chosen to be eliminated. During the Night phase, agents with special roles can acquire information about the game state.

The werewolf competition at ANAC 2020 will be composed of two parallel tasks. In the “Protocol Division”, the participants will create an agent can play werewolf with any role and uses a protocol with a fixed regular language to communicate to each other. In the “Natural Language Division”, the agents will use natural language to communicate to each other.

The winning agent will be able to form successful coalitions with other agents, in order to vote players of opposing teams out of the game, and achieve a high winning ratio.

Task for the Entrants

Entrants to the both tasks must develop and submit an AIWolf agent using Java, Python or C#. The agents communicate with the game server using sockets, through a standard interface provided by the competition organizers. A time limit for each query from the server is enforced.

In the Protocol Division task, the agents use a fixed communication protocol (AIWolf Protocol, Version 6.*) that allows them to communicate their intent, their beliefs about the game state, inquire other agents about their beliefs, and make requests to other agents. They will play on games with 15 players, on a two-tiered competition. The 15 best agents of the preliminary tier will advance to the finalist tier. The winner of each tier is decided by victory ratio.

In the Natural Language Division task, the agents communicate using sentences in English. Agents must be able to parse the meaning of each other's sentences, as well as communicating their own intent clearly. They will play on games with 5 players, and the winner is decided by analysis of the game logs.

More Information:

More information about the game rules, contest rules, agent preparation and sample code can be found on the contest webpage. Please see the following links:

- Website of the 2nd International Werewolf Competition at ANAC 2020: <http://aiwolf.org/en/2nd-international-aiwolf-contest>

Protocol Division task:

- AIWolf project page: <http://aiwolf.org/en/>
- AIWolf Developer's Mailing List: <https://groups.google.com/group/aiwolfdev/>
- AIWolf Contest Organizer's Mailing List: gm@aiwolf.org
- AIWolf Slack Channel: <https://forms.gle/EPsUud7jAiFEEYyF7>

Natural Language Division task:

- AIWolfDial (Natural Language Division) sample agent code and server runtimes: <https://github.com/aiwolfdial/RandomTalkAgent>

Submission:

Protocol Division task:

Participants must submit the source code and binary files for their agents, following the format specified in the competition website (<http://contest.aiwolf.org/en>). The deadlines are:

- Submission Deadline: June 15th, 2020
- Finalist Announcement: June 30th, 2020
- Finalist Submission deadline (Protocol Division): TBA (around 10 days before IJCAI)

Finalists and those with significant novel contributions will be invited to briefly present their work during the competition workshop.

Natural Language Division task:

Participants should run their agents to connect with our server, details will be provided in <https://github.com/aiwolfdia/RandomTalkAgent> .

- Pre-run (self match games): TBA
- Formal run (games between participants): TBA (tentative, depends on ACAN workshop paper submission deadline)

Participants are encouraged to submit their papers to the ACAN workshop.

Organizing Committee:

Claus Aranha, Hirotaka Osawa (University of Tsukuba), Fujio Toriumi (University of Tokyo), Yoshinobu Kano (Shizuoka University), and members of the AIWolf Project.