

---

# Learning to communicate concisely

## *Deep Learning Course Project Proposal*

---

**Lukáš Jendele\***  
ETH Zürich  
jendelel@ethz.ch

**Ondrej Skopek\***  
ETH Zürich  
oskopek@ethz.ch

**Vignesh Ram Somnath\***  
ETH Zürich  
vsomnath@ethz.ch

Deep Reinforcement Learning has been the driving force behind several achievements in game-playing agents over the past few years. Notable examples are AlphaGo [1], AlphaGo Zero [2], AlphaZero [3], Capture The Flag [4], and many others. In this project, we aim to extend the idea behind AlphaZero to train AI agents to play the game of Codenames<sup>2</sup>.

Codenames is a multiplayer card game originally designed for four agents: two teams (red, blue) with two agents — a “spymaster” and “guesser”. The game board is a  $5 \times 5$  grid of nouns, each word belonging to one of four groups: red team, blue team, neutral, and an assassin. The teams alternate with taking turns: a spymaster gives a “clue” (single word that is not contained in the words on the board and a number designating the number of words associated with the clue). The team of the spymaster then goes on to iteratively guess words on the board. If they guessed a card of their team, they get another turn (with a maximum number of turns given by the clue word number + 1). If they guess a neutral card or a card of the other team, their turn ends. Guessing an assassin card makes the team lose immediately. This constitutes one round of the game.

We model the game in the following. We assume a fixed vocabulary  $V_B$  for the words on the board<sup>3</sup>, and a different (larger, more variant) fixed vocabulary  $V_C$  for giving/receiving clues by the agents<sup>4</sup> [5]. We assume  $V_B \subseteq V_C$ . For practical purposes, we only train the two agents for the red team, and let them play one round per training step, because Codenames is easier to evaluate than e.g. chess on a turn by turn basis (guessing a word of your team gives positive rewards, otherwise you get a negative reward).

Our proof of concept implementation in PyTorch will be made available, and the API will enable the creation of bots to play a game of humans vs. bots, bots vs. bots, or humans with bots, with pretrained or custom weights. We expect the game to train faster than AlphaZero, since our rounds are shorter and the rewards are more expressive.

To evaluate our trained agents, we propose the following experiments. First, we compare our agents to a baseline that optimizes a custom loss function in a pretrained word embedding space to find clues, and looks for nearest neighbors in the same space for guessing. Secondly, we compare our spymaster to the implementation of a spymaster by [6]. Optionally, we play a round of games using a human control group: both human vs. bot games and mixed human-bot games.

To summarize, our contributions will be the following:

1. Design and implement deep RL agents in PyTorch that learn to play Codenames by self-play.
2. Implement a simple game interface to play Codenames against/with trained agents.
3. Evaluate the performance of our agents in several diverse experiments, optionally followed by a short ablation study.

---

\*Equal contributions.

<sup>2</sup><https://boardgamegeek.com/boardgame/178900/codenames>

<sup>3</sup>From: <https://github.com/jbowens/codenames/>

<sup>4</sup><https://github.com/facebookresearch/fastText/blob/master/pretrained-vectors.md>

## References

- [1] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484, 2016.
- [2] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *Nature*, 550(7676):354, 2017.
- [3] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dhharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*, 2017.
- [4] Max Jaderberg, Wojciech M Czarnecki, Iain Dunning, Luke Marris, Guy Lever, Antonio Garcia Castaneda, Charles Beattie, Neil C Rabinowitz, Ari S Morcos, Avraham Ruderman, et al. Human-level performance in first-person multiplayer games with population-based deep reinforcement learning. *arXiv preprint arXiv:1807.01281*, 2018.
- [5] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *TACL*, 5:135–146, 2017.
- [6] Avi Kejriwal. A Spymaster AI for playing Codenames. <https://github.com/avikejriwal/Codenames>, 2018.