1. Tasks

- Chinese Whispers, noiseless channel, 1 player
- Chinese Whispers, noiseless channels, 2 players
- Chinese Whispers, noiseless channels, n players
- Chinese Whispers, noisy input channel 10%, 1 player
- ...

2. Input Strings

Length of strings and number of characters must be specified.

- iid (uniformely distributed)
- iid (non-uniformely distributed)
- n-grams
- other distributions?
- English words/text?

3. Intermediate Strings

Length of strings and number of characters must be specified.

4. Reward functions for the last agent

Let n be the length of strings, I be the input string to the first agent, O be the output string of the last agent. Was Skalierung angeht bin ich mir nicht ganz sicher...

- "Hamming reward": $r(I, O) = n - \text{Hamming}(I, O)$
- Levenshtein reward
- ...

5. Reward functions for intermediate agents

Scale of this reward must be specified.
Let n be the length of strings, O be the output string of one intermediate agent.

- Zero: $r(O) = 0$
- Shortness: $r(O) = \#\text{NULLTOKENS}(O)$

6. Neural network blocks

Intermediate dims etc. should be specified.

- 1-Convolution
- n-Convolution
- Fully Connected
- Self Attention

7. RL Algorithms

Learning rates, batch size etc. should be specified.

- Policy Gradients (REINFORCE)
- Deep Deterministic Policy Gradient ?
- Actor Critic ?

8. List of Experiments

8.1. **Experiment 1.**

Chinese Whispers, noiseless channel, 1 player
Length 5, Number of characters 2, iid (uniformely distributed)
(No intermediate strings)
Hamming reward, mean-std-normalized as batch
(No intermediate agents)
Embedding / 1-Convolution; Total params: 10
REINFORCE, SGD(lr=0.01, momentum=0), batch size 64
Code: https://github.com/Pausau/StillePost/blob/main/Experiment1.ipynb
Results: Task fully solved once after $\sim$6.26 parameter updates.

8.2. **Experiment 2.**

Chinese Whispers, noiseless channel, 1 player
Length 5, Number of characters **5**, iid (uniformely distributed)
(No intermediate strings)
Hamming reward, mean-std-normalized as batch
(No intermediate agents)
Embedding / 1-Convolution; Total params: 55
REINFORCE, SGD(lr=0.01, momentum=0), batch size 64
Code: https://github.com/Pausau/StillePost/blob/main/Experiment2.ipynb
Results: Task fully solved once after $\sim$22.62 parameter updates.

8.3. **Experiment 3.**
Chinese Whispers, noiseless channel, **2 players**
Length 5, Number of characters 2, iid (uniformely distributed)
Intermediate: Length 5, Number of characters 2
Hamming reward, mean-std-normalized as batch
Intermediate: No reward
Embedding / 1-Convolution; Params: 10 per player, total params: 20
REINFORCE, SGD(lr=0.01, momentum=0), batch size 64
Code: https://github.com/Pausau/StillePost/blob/main/Experiment3.ipynb
Results: Task fully solved once after ∼18.52 parameter updates.

8.4. **Experiment 4.**
Chinese Whispers, noiseless channel, 2 players
Length 5, Number of characters **3**, iid (uniformely distributed)
Intermediate: Length 5, Number of characters **3**
Hamming reward, mean-std-normalized as batch
Intermediate: No reward
Embedding / 1-Convolution; Params: 21 per player, total params: 42
REINFORCE, SGD(lr=0.01, momentum=0), batch size 64
Code: https://github.com/Pausau/StillePost/blob/main/Experiment4.ipynb
Results: 66% of the time, the task cannot by fully solved within 100 parameter updates. For the times where it was fully solved, the task was fully solved once after ∼35.65 parameter updates.

8.5. **Experiment 5.**
Chinese Whispers, noiseless channel, 2 players
Length 5, Number of characters **5**, iid (uniformely distributed)
Intermediate: Length 5, Number of characters **5**
Hamming reward, mean-std-normalized as batch
Intermediate: No reward
Embedding / 1-Convolution; Params: 55 per player, total params: 110
REINFORCE, SGD(lr=0.01, momentum=0), batch size 64
Code: https://github.com/Pausau/StillePost/blob/main/Experiment5.ipynb
Results: 100% of the time, the task cannot by fully solved within 150 parameter updates.

8.6. **Experiment 6.**
Chinese Whispers, noiseless channel, **3 players**
Length 5, Number of characters **2**, iid (uniformely distributed)
Intermediate: Length 5, Number of characters **2**
Hamming reward, mean-std-normalized as batch
Intermediate: No reward
Embedding / 1-Convolution; Params: 10 per player, total params: 30

REINFORCE, SGD(lr=0.01, momentum=0), batch size 64
Code: https://github.com/Pausau/StillePost/blob/main/Experiment6.ipynb
Results: Task fully solved once after ∼26.82 parameter updates.

8.7. **Experiment 7. (like Nr.3).**
Chinese Whispers **with parallel agents**, noiseless channel, **1 + 2** players
Length 5, Number of characters 2, iid (uniformely distributed)
Intermediate: Length 5, Number of characters 2
Hamming reward, mean-std-normalized as batch
Intermediate: No reward
Embedding / 1-Convolution; Params: 10 per player, total params: 30
REINFORCE, SGD(lr=0.01, momentum=0), batch size 64
Code: https://github.com/Pausau/StillePost/blob/main/Experiment7.ipynb
Results: Task fully solved once after ∼16.52 parameter updates.

8.8. **Experiment 8. (like Nr.4).**
Chinese Whispers **with parallel agents**, noiseless channel, **1 + 2** players
Length 5, Number of characters 2, iid (uniformely distributed)
Intermediate: Length 5, Number of characters **3**
Hamming reward, mean-std-normalized as batch
Intermediate: No reward
Embedding / 1-Convolution; Params: 10 per player, total params: 30
REINFORCE, SGD(lr=0.01, momentum=0), batch size 64
Code: https://github.com/Pausau/StillePost/blob/main/Experiment8.ipynb
Results: 84% of the time, the task cannot by fully solved within 100 parameter updates. For the times where it was fully solved, the task was fully solved once after ∼36.25 parameter updates.