

## 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 (uniformly distributed)
- iid (non-uniformly 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 (uniformly 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 (uniformly 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 (uniformly 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  $\sim 18.52$  parameter updates.

### 8.4. Experiment 4.

Chinese Whispers, noiseless channel, 2 players

Length 5, Number of characters **3**, iid (uniformly 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 be fully solved within 100 parameter updates. For the times where it was fully solved, the task was fully solved once after  $\sim 35.65$  parameter updates.

### 8.5. Experiment 5.

Chinese Whispers, noiseless channel, 2 players

Length 5, Number of characters **5**, iid (uniformly 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 be fully solved within 150 parameter updates.

### 8.6. Experiment 6.

Chinese Whispers, noiseless channel, **3 players**

Length 5, Number of characters **2**, iid (uniformly 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  $\sim 26.82$  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 (uniformly distributed)  
 Intermediate: Length 5, Number of characters 2  
 Hamming reward, mean-std-normalized as batch  
 Intermediate: No reward  
 Embedding(2) / 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  $\sim 16.52$  parameter updates.

#### 8.8. Experiment 8. (like Nr.4).

Chinese Whispers **with parallel agents**, noiseless channel, **1 + 2** players  
 Length 5, Number of characters **3**, iid (uniformly distributed)  
 Intermediate: Length 5, Number of characters **3**  
 Hamming reward, mean-std-normalized as batch  
 Intermediate: No reward  
 Embedding(3) / 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 be fully solved within 100 parameter updates. For the times where it was fully solved, the task was fully solved once after  $\sim 36.25$  parameter updates.

#### 8.9. Experiment 9. (like Nr.4).

Chinese Whispers, noiseless channel, 2 players  
 Length 5, Number of characters **3**, iid (uniformly distributed)  
 Intermediate: Length 5, Number of characters **3**  
 Hamming reward, mean-std-normalized as batch  
 Intermediate: No reward  
 Embedding(**64**) / 1-Convolution; Params: 387 per player, total params: 774  
 REINFORCE, SGD(lr=0.01, momentum=0), batch size 64  
 Code: <https://github.com/Pausau/StillePost/blob/main/Experiment9.ipynb>  
 Results: 22% of the time, the task cannot be fully solved within 100 parameter updates. For the times where it was fully solved, the task was fully solved once after  $\sim 23.56$  parameter updates.

#### 8.10. Experiment 10. (like Nr.4).

Chinese Whispers, noiseless channel, 2 players  
 Length 5, Number of characters **3**, iid (uniformly distributed)  
 Intermediate: Length 5, Number of characters **6**  
 Hamming reward, mean-std-normalized as batch  
 Intermediate: No reward  
 Embedding(**3**) / 1-Convolution; Params: 33, 30; total params: 60  
 REINFORCE, SGD(lr=0.01, momentum=0), batch size 64  
 Code: <https://github.com/Pausau/StillePost/blob/main/Experiment10.ipynb>  
 Results: 66% of the time, the task cannot be fully solved within 100 parameter updates. For the times where it was fully solved, the task was fully solved once after  $\sim 43.41$  parameter updates.

#### 8.11. Experiment 11. (like Nr.4).

Chinese Whispers, noiseless channel, 2 players  
 Length 5, Number of characters **3**, iid (uniformly distributed)  
 Intermediate: Length 5, Number of characters **6**  
 Hamming reward, mean-std-normalized as batch  
 Intermediate: No reward  
 Embedding(**64**) / 1-Convolution; Params: 582, 579; total params: 1161  
 REINFORCE, SGD(lr=0.01, momentum=0), batch size 64  
 Code: <https://github.com/Pausau/StillePost/blob/main/Experiment11.ipynb>  
 Results: 6% of the time, the task cannot be fully solved within 100 parameter updates. For the times where it was fully solved, the task was fully solved once after  $\sim 28.04$  parameter updates.

#### 8.12. Experiment 12. (like Nr.9).

Chinese Whispers, noiseless channel, 2 players  
 Length 5, Number of characters **3**, iid (uniformly distributed)  
 Intermediate: Length 5, Number of characters **3**  
 Hamming reward, mean-std-normalized as batch  
 Intermediate: No reward  
 Embedding(**256**) / 1-Convolution; Params: 1539 per player; total params: 3078  
 REINFORCE, SGD(lr=0.01, momentum=0), batch size 64  
 Code: <https://github.com/Pausau/StillePost/blob/main/Experiment12.ipynb>  
 Results: 22% of the time, the task cannot be fully solved within 100 parameter updates. For the times where it was fully solved, the task was fully solved once after  $\sim 22.72$  parameter updates.

#### 8.13. Experiment 13. (like Nr.11).

Chinese Whispers, noiseless channel, 2 players  
 Length 5, Number of characters **3**, iid (uniformly distributed)

Intermediate: Length 5, Number of characters **30**

Hamming reward, mean-std-normalized as batch

Intermediate: No reward

Embedding(**256**) / 1-Convolution; Params: 8448, 8451; total params: 16,899

REINFORCE, SGD(lr=0.01, momentum=0), batch size 64

Code: <https://github.com/Pausau/StillePost/blob/main/Experiment13.ipynb>

Results: 34% of the time, the task cannot be fully solved within 200 parameter updates (one above 100). For the times where it was fully solved, the task was fully solved once after  $\sim 56.73$  parameter updates.