

## REINFORCEMENT LEARNING COMPETITION - THE COMPETITION RULES

### ◆ GOAL

You need to dig for as much gold as possible using the least amount of energy. There are many traps and obstacles in the forest so move wisely and use the right command in time.

### ◆ RULES OF GAME

- **n** treasure hunters ( $1 \leq n \leq 4$ ) start at the same point in the forest.
- The forest map is represented as a **21\*9** two-dimensional matrix.
- The original amount of energy of each hunter is **50**.
- The game has **100** turns and for each turn, the player has to perform one of the following actions:

GoLeft	Go to the left ←
GoRight	Go to the right→
GoUp	Go up ↑
GoDown	Go down ↓
Free	Take a rest
Craft	Dig for gold

- After each turn, the system will send the actions performed by all players and update each player's gold and energy.
- To win, players must pass the obstacles and dig for the most amount of gold.

### ◆ ENERGY CONSUMED:

Going to one of the following cells or digging for gold consumes the following amount of energy:



Land: **-1** energy



Trap: **-10** energy. Use only 1 time (per turn). If in a turn, more than one player goes into a trap box (not used), then all will be deducted 10 energy. After being used, the trap turns into land.



Forest: is the random value between **[-5;-20]** energy



Swamp: **-5 -> -10 -> -20 -> -40 -> -100** energy. The amount of energy needed will increase with the number of times the treasure hunters enter the swampy box. The first time will be subtract 3 energy, the second will be 5 ... (the number of times counted by turns)



Mine: **-4** energy



Dig for gold: **-5** energy

#### ◆ ENERGY REGAINED

In order to regain energy, the adventurers must take time to rest:

<i>Number of turns</i>	<i>Energy regained</i>
1 turn	<b>[E/4]</b>
2 consecutive turns	<b>[E/4] + [E/3]</b>
3 consecutive turns	<b>[E/4] + [E/3] + [E/2]</b>
In case energy exceeds E	<b>No regain</b>

*Note: the result is rounded down to the nearest integer*

#### ◆ RULES OF GOLD CRAFT:

Assuming that the amount of gold in the mine is  $G$  and there are  $m$  people digging for gold.

The amount of gold each person collects is calculated as in the table below:

<i>Gold</i>	<i>The Gold each person gains</i>	<i>Remaining Gold</i>	<i>Energy consumed</i>
$G \geq m*50$	50	$G-m*50$	5
$G < m*50$	$[G/m]$	0 (this gold mine turns into land)	5

#### ◆ OTHER RULES:

- Not standing at a gold mine but digging for gold: **-10E**
- Not having enough energy to perform the action and not using a rest command: **Eliminated**
- No action is taken at the current turn: **Eliminated**
- **E = 0: Eliminated**

#### ◆ FORM OF COMPETITION AND BOT TRAINING

##### ● Form of competition:

- The competition is structured as a Gold Miner game.
- Teams compete online with the Organizers bots (in the Screening Stage - Navigation Challenge) or compete directly with each other (in the Final Stage - Treasure Island).
- Teams are required to write a program that uses machine learning algorithms to solve the problems set out by the Organizers, in which using Reinforcement Learning is definitely a plus.
- Teams then put their code and trained model (hereinafter referred to as "Agent") on the system (<https://codelearn.io/>).
- The Organizers arrange matches upon reception of Agents from teams.

- **Code writing and Agent training**

- Each team is required to write a program in Python with popular tools like Tensorflow or Torch based on the Organizers' sample code and use the reinforcement learning algorithm to train the Agent. The Agent, when submitted to the Organizers, will be used to compete on the server system.
- Teams upload their Agent versions to the system (up to 30MB) after running through testing, gold mining matches against the Organizers' bots in the Screening Stage - Navigation Challenge, and matches against the opponent teams in the Final Stage - Treasure Island.
- Each team is allocated an amount of up to 1024 Mb RAM to run the code during the competition.
- In the Screening Stage - Navigation Challenge, teams are allowed to update unlimited Agent versions on the server system, but can only play maximum 20 matches with the Organizer's Bot. (The match with the Bot will be automatically activated after the code is verified successfully).
- In the Final Stage - Treasure Island, each match will have four teams competing each other. The teams run their Agents on the Organizers' system. The winning team will proceed to the next stage. The team that wins the final match is the champion.
- After each round of competition, teams are allowed to update a new code. For the Final Round, the team can submit a maximum of 5 or a minimum of 1 code version to the Organizers. The deadline for submitting (5) new version codes is 18:59:59 18/09/2020 (GMT+7). Before every match, each team needs to confirm with the Organizers the code version will use (by active code version). In case the competition team doesn't activate the code, they will not be able to participate in this match and the default score of this match will be 0 point.
- Using the Organizers' bots to submit results is prohibited. Any instances of copying will be disqualified.
- During the competition, the teams can track the amount of gold mined after each stage on the competition's website.

◆ **FORM OF COMPETITION AND RANKING OF EACH STAGE**

Reinforcement Learning Competition includes:

- Screening Stage - Navigation Challenge: 08/08 - 03/09/2020
- Friendly Stage - Stepping Stone: 05/09 - 17/09/2020
- Final Stage - Treasure Island (Day 1): 19/09/2020
- Final Stage - Treasure Island (Day 2): 26/09/2020

## 1. SCREENING STAGE - NAVIGATION CHALLENGE

### 1.1. Schedule: 08/08 – 03/09/2020

- The Screening Stage - Navigation Challenge lasts for four weeks.
- The stage is open at 8:00 on 8/8 and closed at 23:00 on 3/9.
- The start time is counted from the time the topic is opened. Teams develop and submit their Agents at any time during the valid period.
- Teams can only submit their Agents during the valid period, once the topic is closed, teams will not be able to make any change.

### 1.2. Form of competition:

- The teams compete online on the system <https://codelearn.io>
- In this stage, each team is given five maps. These maps will be used in matches during the Navigation Challenge stage.
- The mission of the Agent is to dig as much gold as possible with the amount of available energy.
- Each team is allowed to update unlimited Agent versions on the server system, but can only play maximum 20 matches with the Organizer's Bot. (The match with the Bot will be automatically activated after the code is verified successfully).
- For each Agent version update, teams will compete in five matches on five given maps.
- In this stage, the positions of the obstacles, boxes of gold and the amount of gold **will not** change after each match. Each team will compete against three bots of the Organizers (not against other teams).

### 1.3. Scoring:

- The score of each Agent update is the total amount of gold mined after five matches.
- The score in the Navigation Challenge stage of a team is the score of **the Agent update with the highest amount of gold mined.**

#### **1.4. Final ranking**

- The team with the best score in the Navigation Challenge stage will rank first.
- In the event that two or more teams gain the same score, the ranking is determined as follows:
  - In case two or more teams gain the same scores, the team that submits to the system at the earliest will rank first.  
  
For example: The best score of Team A is 1,300 gold and is recorded at 16:05:30 on 8/8 on the system. The best score of Team B is also 1,300 gold, but is recorded at 9:25:00 on 16/8 on the system. Team A ranks above team B.
- In the event the teams gain the same score and submit to the system at the same time, they will rank equally. In this case, the Organizers will consider drawing lots to select the team to proceed to the next stage. The Organizers' decision is final and binding on the participants in any circumstance.
- After a maximum of 20 submissions, the teams with the best score in the Navigation Challenge stage will be qualified to the Treasure Island stage.
- The teams pledge to compete honestly. Upon any detected fraud, the Organizers cancel the result and select the teams with the next best score in the Navigation Challenge stage.

## **2. FRIENDLY STAGE - STEPPING STONE**

### **2.1. Schedule:** 5/9 - 17/9 (AoE time)

The Organizers hold friendly matches at 12:00 pm every day during the Stepping Stone stage.

### **2.2. Form of competition:**

- Teams that win the Navigation Challenge stage are randomly divided into 8 groups by the Organizers. Each group has 8 numbers of teams.
- The Organizers sequentially arrange friendly matches for the teams. After each day, teams can view the match history on the system.

## **3. FINAL STAGE - TREASURE ISLAND**

### **3.1. GENERAL INFORMATION**

#### **3.1.1. Schedule:** 19/09/2020 (Day 1) - 26/09/2020 (Day 2)

### **3.1.2. Form of competition:**

- The teams compete online on the system <https://codelearn.io>
- The Final Stage - Treasure Island consists of 4 rounds: Round 1 - Sudden Death, Round 2 - Moving Forward, Quarter Final - The True Heir and Final - Mysterious Treasure.
- Round 1 - Sudden Death & Round 2 - Moving Forward will take place on Saturday, 19/9/2020. Quarter Final - The True Heir & Final - Mysterious Treasure will be held a week later.

### **3.1.3. Topic questions and agent training instruction:**

- In this stage, the Organizers will slightly change the map.
- There are 12 maps will be used in the Final Round. They will be announced to all teams by the organizers on September 3rd.
- Before every match, each team needs to confirm with the Organizers the code version will use (by active code version). In case the competition team doesn't activate the code, they will not be able to participate in this match and the default score of this match will be 0 point.

### **3.1.4. Scoring and ranking:**

#### **Rules:**

**First place: 3 points**

**Second place: 2 points**

**Third place: 1 point**

**Fourth place: 0 point**

**\*Applicable to all Treasure Island matches**

- The ranking of each team is determined based on the amount of gold mined. In case that more than 1 team gain the same amount of gold, each team's score is determined as follows:

**\*Ranking score = (Total scores of drawal teams) / (Total number of drawal teams)**

**Example 1:** The amount of gold 4 teams mined are: Team A: 1500 (gold), Team B: 1500 (gold), Team C: 1000 (gold), Team D: 800 (gold). So the points that each team

gains is: Team A = Team B =  $(3 + 2) / 2 = 2.5$  points; Team C: 1 point; Team D: 0 point.

**Example 2:** The amount of gold 4 teams mined are: Team A: 1500 (gold), Team B: 1000 (gold), Team C: 1000 (gold), Team D: 1000 (gold). So the points that each team gains is: Team A: 3 points; Team B = Team C = Team D =  $(2 + 1 + 0) / 3 = 1$  point.

- The final score in each round is the total score of all the matches in that round.
- The team that gained the highest final score in each round will rank first.
- In the event that two or more teams gain the same score, the ranking is determined as follows:
  - If two or more teams have the same score, the team with the higher amount of gold mined after matches will rank higher.
  - If two or more teams have the same score and the same amount of gold mined, the team with the higher amount of "remaining energy" in the last match of that round will rank higher.
  - If two or more teams have the same score, the same amount of gold mined and the same amount of "remaining energy" in the last match of that round, they will rank equally. In this case, the Organizers will consider drawing lots to select the team to proceed to the next stage.
  - The teams pledge to compete honestly. Upon any detected fraud, the Organizers cancel the result and select the teams with the next best score to proceed.

\* **Note:** The winner is the team that remains the longest in the match and has mined the most gold.

Example: Team A mines 2000 (gold) but dies in step 98; Team B mines 1000 (gold) and finishes the full game; Team C mines 1500 (gold) and dies in step 99; Team D lives until the end of the game and mines 800 (gold), the ranking points each team will receive are: Team B: 3 points; Team D: 2 points; Team C: 1 point, Team A: 0 point.

## **3.2. QUALIFYING ROUND 1: SUDDEN DEATH**

### **3.2.1. Form of competition:**

- Teams with the highest score in the Navigation Challenge stage will compete in Round 1.

- Based on the result of the Navigation Challenge stage, the Organizers divide the teams into four seed groups, the number of teams in each seed group is 16.
- Teams will then be divided randomly into groups A, B, C, D,... Each group must include four teams from four different seed groups. (draw time is expected to take place on September 18)
- Each group competes in five matches. The total number of matches in this round is 80.
- Each match is competed by all four teams of the group.

#### **3.2.2. Scoring and ranking:**

- **The final score in Round 1** of a team is the total score gained after five matches.
- The Organizers select a team with **the highest final score in Round 1** in each group to enter Round 2 (total 16 teams).

### **3.3. QUALIFYING ROUND 2: MOVING FORWARD**

#### **3.3.1. Form of competition:**

- 16 teams qualified for Round 2 are divided equally into four groups: I, II, III, and IV.
- Each group will compete in five matches (20 matches in total).

#### **3.3.2. Scoring and ranking:**

- **The final score in Round 2** of a team is the total score gained after five matches.
- The Organizers select 02 teams with **the highest final score in Round 2** in each group to enter Quarter Final: The True Heir (total 08 teams).

### **3.4. QUARTER-FINAL: THE TRUE HEIR**

#### **3.4.1. Form of competition:**

- Eight teams qualified for the Quarter Final are grouped as follows:
  - The first-rank team in group I, first-rank team in group II, second-rank team in group III, second-rank team in group IV compete in the Quarter Final match 1.
  - The second-rank team in group I, second-rank team in group II, first-rank team in group III, first-rank team in group IV compete in the Quarter Final match 2.
  - Each Quarter Final match is played out in five small matches (10 small matches in total)

#### **3.4.2. Scoring and ranking:**

- **The final score of a team** is the total score gained after five matches.
- The Organizers selects 02 teams with **the highest final score** in each Quarter Final match to enter the Final (total 04 teams).

### **3.5. FINAL STAGE - MYSTERIOUS TREASURE**

#### **3.5.1. Form of competition:**

- Four winning teams in the Quarter Final will advance to the Final Stage.
- The teams compete in a total of five matches.

#### **3.5.2. Scoring and ranking:**

- **The final score of a team** is the total score gained after five matches.
- Awards are given based on **the final score**, specifically:
  - The team with the highest final score wins the First Prize
  - The team with the second highest final score wins the Second Prize
  - The team with the third highest final score wins the Third Prize
  - The other team wins the Fourth Prize
  - The Talent and Prospective awards are determined and given after the Final.