

LUẬT CHƠI CHÍNH THỨC CUỘC THI ĐẤU TRƯỜNG REINFORCEMENT LEARNING

◆ MỤC TIÊU

Bạn cần đào được số vàng nhiều nhất với số năng lượng tiêu tốn là ít nhất. Trong khu rừng nơi bạn đang đứng có rất nhiều cạm bẫy và chướng ngại vật, vì vậy hãy di chuyển một cách thông minh và sử dụng lệnh nghỉ đúng lúc.

◆ LUẬT CHƠI

- n thợ săn kho báu ($1 \leq n \leq 4$) sẽ cùng xuất phát tại cùng một điểm trong khu rừng.
- Bản đồ khu rừng được biểu diễn dưới dạng một ma trận 2 chiều **21*9**.
- Số năng lượng lúc đầu của mọi người là **50**.
- Lượt chơi mỗi trận đấu là **100**, mỗi lượt người chơi sẽ phải thể hiện một trong các hành động:

GoLeft	Đi sang trái ←
GoRight	Đi sang phải →
GoUp	Đi lên ↑
GoDown	Đi xuống ↓
Free	Nghỉ ngơi
Craft	Đào vàng

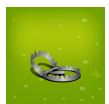
- Sau mỗi 1 lượt, hệ thống sẽ gửi lại phương án đi của những người chơi khác đồng thời cập nhật số vàng và năng lượng của mỗi người chơi.
- Người chơi phải vượt qua các chướng ngại vật và đào thật nhiều vàng để chiến thắng.

◆ QUY TẮC TIÊU THỤ NĂNG LƯỢNG:

Đi vào một trong các ô sau hoặc đào vàng sẽ tiêu tốn số năng lượng như sau:



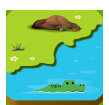
Đất: **-1** năng lượng



Bẫy: **-10** năng lượng. Chỉ được sử dụng 1 lần (tính theo lượt chơi. Nếu trong 1 lượt chơi, nhiều người chơi cùng đi vào 1 ô bẫy (chưa được sử dụng), thì tất cả đều sẽ bị trừ 10 năng lượng. Sau khi được sử dụng, bẫy chuyển thành đất.



Rừng: là giá trị ngẫu nhiên từ trong đoạn **[-5, -20]** năng lượng



Đầm lầy: **-5 -> -20 -> -40 -> -100** năng lượng. Năng lượng cần dùng sẽ tăng dần theo số lần vào ô có đầm lầy. Lần đầu tiên sẽ là trừ 5 năng lượng, lần thứ 2 là 20... (số lần vào tính theo lượt chơi)



Mỏ vàng: **-4** năng lượng



Đào vàng: **-5** năng lượng

◆ QUY TẮC HỒI NĂNG LƯỢNG:

Để hồi năng lượng, nhà thám hiểm phải dành thời gian ngồi nghỉ:

<i>Số lần nghỉ</i>	<i>Số năng lượng hồi</i>
1 lượt	[E/4]
2 lượt liên tiếp	[E/4] + [E/3]
3 lượt liên tiếp	[E/4] + [E/3] + [E/2]
Năng lượng hồi quá E	Không hồi thêm

Lưu ý: kết quả của phép chia được làm tròn xuống số nguyên gần nhất

◆ QUY TẮC ĐÀO VÀNG:

Giả sử tại thời điểm hiện tại, lượng vàng trong mỏ là **G** và có **m** người cùng đào vàng. Số lượng vàng mỗi người thu được được tính toán như bảng dưới đây:

Lượng vàng	Lượng vàng mỗi người đạt được	Lượng vàng còn lại ở mỏ	Năng lượng tiêu tốn
$G \geq m \cdot 50$	50	$G - m \cdot 50$	5
$G < m \cdot 50$	$[G/m]$	0 (ô này chuyển thành ô đất)	5

◆ CÁC QUY TẮC KHÁC:

- Không đứng tại mỏ mà dùng lệnh đào vàng: **-10E**
- Không đủ năng lượng mà không nghỉ ngơi: **Loại**
- Đi ra khỏi bản đồ: **Loại**
- Không đưa ra hành động khi đến lượt chơi: **Loại**
- Năng lượng **E = 0**: **Loại**

◆ HÌNH THỨC THI ĐẤU VÀ HUẤN LUYỆN AGENT

● Hình thức thi đấu:

- Cuộc thi được mô phỏng dưới dạng game đào vàng (Gold miner).
- Các đội thi thi đấu online với Bot của BTC (tại Vòng bảng - Đón đầu thách thức) hoặc thi đấu trực tiếp với các đội thi khác (tại Vòng Chung kết - Đào giấu vàng).
- Các đội thi cần viết một chương trình sử dụng các thuật toán học máy để đưa ra các chuỗi quyết định sử dụng Reinforcement Learning để giải quyết vấn đề (có thể sử dụng các thuật toán khác nhưng không khuyến khích).

- Đội thi đưa đoạn code và mô hình đã được huấn luyện của mình (sau đây gọi là Agent) trực tiếp lên hệ thống (<https://codelearn.io/>).
- Ban tổ chức tiến hành xếp trận đấu sau khi nhận được Agent từ các đội thi.

- **Hình thức viết code và huấn luyện Agent**

- Mỗi đội cần viết một đoạn chương trình bằng ngôn ngữ lập trình Python với các công cụ phổ biến như Tensorflow, Torch dựa trên code mẫu của ban tổ chức, sử dụng thuật toán học tăng cường huấn luyện Agent của mình. Agent sau khi nộp lên ban tổ chức được dùng để thi đấu trên hệ thống server được chuẩn bị sẵn.
- Các đội thi upload các phiên bản Agent của mình lên hệ thống, dung lượng tối đa 30MB, chạy qua các bước kiểm thử, các lượt thi đào vàng với Bot của BTC tại Vòng bảng - Đón đầu thách thức và các trận đấu với các đội đối thủ tại Vòng Chung kết - Đào giấu vàng
- Mỗi đội chơi được cấp tối đa 1024 Mb RAM để chạy mã nguồn trong khi thi đấu.
- Tại Vòng bảng - Đón đầu thách thức, các đội thi được phép cập nhật không giới hạn số phiên bản Agent, nhưng chỉ được thi đấu tối đa 20 lần với Bot của BTC (Trận đấu với bot sẽ được tự động kích hoạt sau khi code của đội thi được xác thực thành công). Ban tổ chức sử dụng kết quả các trận đấu tại từng lần cập nhật phiên bản để xếp hạng Vòng bảng - Đón đầu thách thức.
- Tại Vòng Chung kết - Đào giấu vàng, mỗi trận đấu bao gồm 04 đội thi đấu với nhau. Các đội thực thi Agent trên hệ thống thi đấu của Ban tổ chức. Đội thắng cuộc được vào vòng thi đấu tiếp theo. Đội thắng trận cuối cùng là đội vô địch.
- Các đội thi không được sử dụng bot của BTC để nộp thành kết quả. Mọi trường hợp phát hiện sự sao chép, đội thi sẽ bị loại khỏi cuộc thi.
- Sau mỗi vòng thi, các đội được phép cập nhật code mới. Đối với Vòng Chung kết, các đội được nộp tối đa là 5, tối thiểu là 1 phiên bản code cho BTC. Thời hạn cuối cùng cho việc nộp code là đến hết 18:59:59 ngày 18/09/2020. Trước mỗi trận đấu, các đội cần chủ động active version code muốn tham gia thi đấu. Trường hợp đội thi không có version code nào được active thì sẽ không được tham gia thi đấu, và mặc định số điểm xếp hạng đạt được là 0 điểm cho trận đấu đó.
- Trong quá trình thi, các đội thi có thể theo dõi thông tin lượng vàng đào được sau mỗi lượt đấu trên website của cuộc thi.

◆ HÌNH THỨC THI ĐẤU VÀ CÁCH TÍNH ĐIỂM TỪNG VÒNG

Cuộc thi Đấu trường Reinforcement Learning bao gồm:

- Vòng bảng - Đón đầu thách thức: 08/08 - 03/09/2020
- Vòng giao hữu - Stepping Stone: 05/09 - 17/09/2020
- Vòng Chung kết - Đào giấu vàng (Lượt trận 1): 19/09/2020
- Vòng Chung kết - Đào giấu vàng (Lượt trận 2): 26/09/2020

1. VÒNG BẢNG - ĐÓN ĐẦU THÁCH THỨC

1.1. Thời gian thi đấu: 08/08 – 03/09/2020 (giờ AoE)

- Vòng bảng - Đón đầu thách thức sẽ kéo dài trong 4 tuần.
- Thời gian mở đề Vòng bảng là 8:00 ngày 8/8, thời gian đóng đề: 23:59' ngày 3/9.
- Thời gian bắt đầu được tính từ lúc mở đề. Các đội thi làm và nộp Agent tại bất kỳ thời điểm nào trong thời gian mở đề.
- Đội thi chỉ có thể nộp Agent trong thời gian đề thi mở, khi vòng thi đã kết thúc, đội thi sẽ không thể quay lại để tham gia thi đấu.

1.2. Hình thức thi đấu:

- Đội thi thi đấu online trên hệ thống <https://codelearn.io>
- Tại vòng này, đội thi được cung cấp 5 bản đồ địa hình cho trước. Các bản đồ này sẽ được sử dụng trong các trận đấu tại Vòng Đón đầu thách thức.
- Nhiệm vụ của các Agent là đào được số vàng nhiều nhất với số năng lượng cho phép.
- Các đội thi được phép cập nhật không giới hạn số phiên bản Agent, nhưng chỉ được thi đấu tối đa 20 lần với Bot của BTC (Trận đấu với bot sẽ được tự động kích hoạt sau khi code của đội thi được xác thực thành công).
- Với mỗi lượt cập nhật phiên bản Agent, các đội sẽ thi đấu 5 trận trên 5 bản đồ được cho trước.
- Tại vòng thi này, vị trí của các vật cản, ô chứa vàng và số lượng vàng sẽ không thay đổi sau mỗi ván chơi, các đội chơi sẽ thi đấu độc lập với 3 Bot của BTC (không có sự xuất hiện của các đội thi khác.) Trong trận đấu nếu các bot của ban tổ chức chết, các Agent vẫn tiếp tục đào vàng đến khi hết lượt thi đấu được ban tổ chức quy định.

1.3. Tính điểm thành tích:

- Thành tích của mỗi lượt cập nhật Agent bằng tổng số vàng đào được sau 5 trận đấu.

- Thành tích Vòng Đón đầu thách thức của một đội thi là thành tích của **lượt cập nhật Agent có số vàng cao nhất** mà đội thi đào được.

1.4. Xếp hạng chung cuộc

- Đội thi có Thành tích Vòng Đón đầu thách thức tốt hơn sẽ xếp trên.
- Trong trường hợp có 2 hoặc nhiều đội thi đạt cùng thành tích, xếp hạng được xác định như sau:

- Trường hợp có 2 hoặc nhiều đội thi đạt cùng điểm số, đội thi nào có thời gian submit lên hệ thống sớm hơn thì đội đó xếp trên.

Ví dụ: Thành tích tốt nhất trong các lượt cập nhật Agent của đội A là 1300 lượng vàng, lượt cập nhật đó được ghi nhận trên hệ thống vào 16:05:30 ngày 8/8. Trong khi đó, thành tích tốt nhất trong các lượt cập nhật Agent của đội B cũng là 1300 lượng vàng, nhưng lượt cập nhật đó được ghi nhận trên hệ thống vào 9:25:00 ngày 16/8, thì đội A được xếp trên đội B.

- Trường hợp các đội thi cùng điểm số, cùng thời gian submit lên hệ thống, các đội thi sẽ có thứ hạng bằng nhau. Trong trường hợp này, BTC xem xét bóc thăm cho các đội thi có thứ hạng bằng nhau để lấy đội đi tiếp. Kết quả của BTC là kết quả cuối cùng.
- Sau tối đa 20 lượt submit, các đội thi có thành tích Vòng Đón đầu thách thức tốt nhất sẽ được chọn vào Vòng Đào giấu vàng.
- Các đội thi cam kết thi đấu trung thực. Khi có bất kỳ phát hiện nào về gian lận, BTC xem xét hủy kết quả thi đấu và lựa chọn các đội thi có xếp hạng tiếp theo tại bảng xếp hạng Vòng Đón đầu thách thức để thay thế.

2. VÒNG GIAO HỮU - BƯỚC ĐỆM

2.1. Thời gian thi đấu: 5/9 - 17/9 (giờ AoE)

BTC sẽ mở hệ thống đấu giao hữu vào 12:00 pm mỗi ngày trong suốt khoảng thời gian diễn ra Vòng Stepping Stone.

2.2. Hình thức thi đấu:

- Các đội vượt qua Vòng Đón đầu thách thức được BTC chia ngẫu nhiên vào 8 bảng, mỗi bảng có 8 đội. Mỗi ngày, danh sách đội của mỗi bảng được BTC ngẫu nhiên thay đổi.
- Ban tổ chức lần lượt xếp các trận đấu giao hữu cho các đội thi. Sau mỗi ngày, các đội thi có thể tải lại lịch sử trận đấu trên hệ thống.

3. VÒNG CHUNG KẾT - ĐÀO GIÁU VÀNG

3.1. THÔNG TIN CHUNG

3.1.1. Thời gian thi đấu: 19/09/2020 (Lượt trận 1) - 26/09/2020 (Lượt trận 2)

3.1.2. Hình thức thi đấu:

- Thi đấu online trên hệ thống <https://codelearn.io>
- Vòng Chung kết - Đào giấu vàng bao gồm Vòng loại số 1: Sudden Death - Cái chết bất ngờ, Vòng loại số 2: Moving Forward - Không lùi bước, Vòng Tứ kết: The True Heir - Người xứng đáng và Trận Chung kết: Mysterious Treasure - Báu vật huyền bí.
- Vòng loại số 1 và Vòng loại số 2 được diễn ra vào ngày 19/09/2020. Vòng Tứ kết và Trận Chung kết diễn ra vào ngày 26/09/2020.

3.1.3. Đề bài và huấn luyện Agent:

- Từ Vòng Đào giấu vàng, Ban tổ chức sẽ thay đổi bản đồ với số lượng thay đổi nhỏ. Có tất cả 12 bản đồ được sử dụng cho Vòng Chung kết. BTC sẽ công bố bản đồ cho các đội thi vào ngày 3/9.
- Trước mỗi trận đấu, các đội cần chủ động active version code muốn tham gia thi đấu. Trường hợp đội thi không có version code nào được active thì sẽ không được tham gia thi đấu, và mặc định số điểm xếp hạng đạt được là 0 điểm cho trận đấu đó.

3.1.4. Tính điểm và xếp hạng:

Quy ước:

Xếp hạng 1: 3 điểm

Xếp hạng 2: 2 điểm

Xếp hạng 3: 1 điểm

Xếp hạng 4: 0 điểm

***Quy ước này được áp dụng cho tất cả các Vòng đấu của Vòng Đào giấu vàng**

- Xếp hạng của mỗi trận đấu được xác định dựa trên số lượng vàng mà các đội đào được. Trong trường hợp có nhiều hơn 1 đội đào được cùng một lượng vàng, điểm xếp hạng của mỗi đội được xác định như sau:

***Điểm xếp hạng = (Tổng điểm xếp hạng hòa) / (Tổng số đội hòa)**

Ví dụ 1: Số lượng vàng 4 đội đào được lần lượt là: Đội A: 1500 (vàng), Đội B: 1500 (vàng), Đội C: 1000 (vàng), Đội D: 800 (vàng). Vậy số điểm xếp hạng mà mỗi đội nhận được là: $\text{Đội A} = \text{Đội B} = (3+2)/2 = 2,5$ điểm; $\text{Đội C} = 1$ điểm; $\text{Đội D} = 0$ điểm

Ví dụ 2: Số lượng vàng 4 đội đào được lần lượt là: Đội A: 1500 (vàng), Đội B: 1000 (vàng), Đội C: 1000 (vàng), Đội D: 1000 (vàng). Vậy số điểm xếp hạng mà mỗi đội nhận được là: Đội A: 3 điểm; $\text{Đội B} = \text{Đội C} = \text{Đội D} = (2+1+0)/3 = 1$ điểm.

- Điểm chung cuộc của từng vòng là tổng điểm xếp hạng của các trận đấu trong vòng đó.
- Đội có Điểm chung cuộc của từng vòng lớn hơn thì xếp trên.
- Trong trường hợp có 2 hoặc nhiều đội thi đạt cùng thành tích, xếp hạng được xác định như sau:
 - Trường hợp có 2 hoặc nhiều đội thi đạt cùng tổng điểm xếp hạng, đội thi nào có tổng số vàng đào được sau các trận lớn hơn sẽ xếp trên.
 - Trường hợp có 2 hoặc nhiều đội thi đạt cùng tổng điểm xếp hạng, cùng số vàng đào được, đội thi nào có “Mức năng lượng còn lại” tại trận đấu cuối cùng của Vòng thi đó nhiều hơn sẽ xếp trên.
 - Trường hợp các đội thi cùng tổng điểm, cùng số vàng đào được, cùng “Mức năng lượng còn lại” tại trận đấu cuối cùng của Vòng thi đó, các đội sẽ có thứ hạng bằng nhau. Trong trường hợp này, BTC xem xét bốc thăm cho các đội thi có thứ hạng bằng nhau để lấy đội đi tiếp
 - Các đội thi cam kết thi đấu trung thực. Khi có bất kỳ phát hiện nào về gian lận, BTC xem xét hủy kết quả thi đấu và lựa chọn các đội thi có xếp hạng tiếp theo tại bảng xếp hạng để thay thế.

***Lưu ý:** Đội chiến thắng là đội còn lại ở trận đấu lâu nhất và giành được lượng vàng nhiều nhất.

VD: Đội A đào được 2000 (vàng) nhưng chết ở step 98; đội B đào được 1000 (vàng) và sống đến hết trận; Đội C đào được 1500 (vàng) và chết ở step 99; Đội D sống đến cuối trận và đào được 800 (vàng) thì điểm xếp hạng mà mỗi đội nhận được là: Đội B: 3 điểm; Đội D: 2 điểm; Đội C: 1 điểm, Đội A: 0 điểm.

3.2. VÒNG LOẠI SỐ 1: SUDDEN DEATH - CÁI CHẾT BẤT NGỜ

3.2.1. Hình thức thi đấu:

- 64 đội có điểm cao nhất Vòng Đón đầu thách thức sẽ tiếp tục thi đấu Vòng loại số 1.
- Dựa trên kết quả của Vòng Đón đầu thách thức, BTC chia các đội thành 4 nhóm hạt giống, mỗi nhóm hạt giống gồm 16 đội.

- Tiếp đó các đội thi được BTC sắp xếp vào các bảng đấu, các bảng được đánh thứ tự A, B, C, D,... mỗi bảng phải bao gồm 4 đội ở 4 nhóm hạt giống khác nhau. Việc lựa chọn các đội cho 1 bảng được BTC bốc thăm ngẫu nhiên trên các nhóm hạt giống (dự kiến thời gian bốc thăm sẽ diễn ra vào ngày 18/9)
- Mỗi bảng thi đấu 5 lượt trận. Tổng số lượt trận của vòng này là 80.
- Mỗi trận đấu sẽ có sự xuất hiện của cả 4 đội (ở cùng một bảng) tham gia thi đấu.

3.2.2. Tính điểm và xếp hạng:

- **Điểm chung cuộc Vòng loại số 1** của một đội thi bằng tổng số điểm xếp hạng mà đội đó đạt được sau 5 lượt trận.
- BTC chọn 01 đội có **Điểm chung cuộc Vòng loại số 1** cao nhất ở mỗi bảng vào Vòng loại số 2.
- Tổng số đội được chọn sau Vòng loại số 1 là **16**.

3.3. VÒNG LOẠI SỐ 2: MOVING FORWARD - KHÔNG LÙI BƯỚC

3.3.1. Hình thức thi đấu:

- 16 đội đi tiếp vào Vòng loại số 2 được BTC chia đều vào 4 bảng, được đánh số I, II, III, IV.
- Mỗi bảng gồm 4 đội, thi đấu 5 lượt. Tổng số lượt thi đấu của vòng này là 20.

3.3.2. Tính điểm và xếp hạng:

- **Điểm chung cuộc Vòng loại số 2** của một đội thi bằng tổng số điểm xếp hạng mà đội đó đạt được sau 5 lượt trận.
- BTC chọn 2 đội có **Điểm chung cuộc Vòng loại số 2** cao nhất ở mỗi bảng vào Vòng The True Heir. Tổng số đội được chọn là **8**.

3.4. VÒNG TỨ KẾT: THE TRUE HEIR - NGƯỜI XỨNG ĐÁNG

3.4.1. Hình thức thi đấu:

- 8 đội được lựa chọn vào Vòng Tứ kết được BTC phân nhóm như sau:
 - Đội nhất bảng I, nhất bảng II, nhì bảng III, nhì bảng IV thi đấu trận Tứ kết 1.
 - Đội nhì bảng I, nhì bảng II, nhất bảng III, nhất bảng IV thi đấu trận Tứ kết 2.
- Mỗi trận Tứ kết gồm 4 đội, thi đấu tổng cộng 5 lượt. Tổng số trận đấu của vòng này là 2, tổng số lượt đấu là 10.

3.4.2. Tính điểm và xếp hạng:

- **Điểm chung cuộc Vòng Tứ kết** của một đội thi bằng tổng số điểm xếp hạng mà đội đó đạt được sau 5 lượt trận
- BTC chọn 2 đội có **Điểm chung cuộc Vòng Tứ kết** cao nhất ở mỗi trận Tứ kết vào Trận Chung kết. Tổng số đội được lựa chọn vào Trận Chung kết là 4 đội.

3.5. TRẬN CHUNG KẾT - MYSTERIOUS TREASURE

3.5.1. Hình thức thi đấu:

- Trận Chung kết là trận đấu giữa 4 đội được lựa chọn từ Vòng Tứ kết.
- Các đội thi đấu tất cả 5 lượt trận.

3.5.2. Tính điểm và xếp hạng

- **Điểm chung cuộc Trận Chung kết** của một đội thi bằng tổng số điểm xếp hạng mà đội đó đạt được sau 5 lượt trận.
- Giải thưởng mà các đội giành được dựa trên **Điểm chung cuộc Trận Chung kết** của đội thi đó, cụ thể:
 - Đội có điểm số chung cuộc cao nhất giành giải Nhất
 - Đội có điểm số chung cuộc cao thứ hai giành giải Nhì
 - Đội có điểm số chung cuộc cao thứ ba giành giải Ba
 - Đội còn lại giành giải Tư.
 - Các Giải Tài năng, Giải Triển vọng được Ban tổ chức lựa chọn và công bố sau Trận Chung kết.