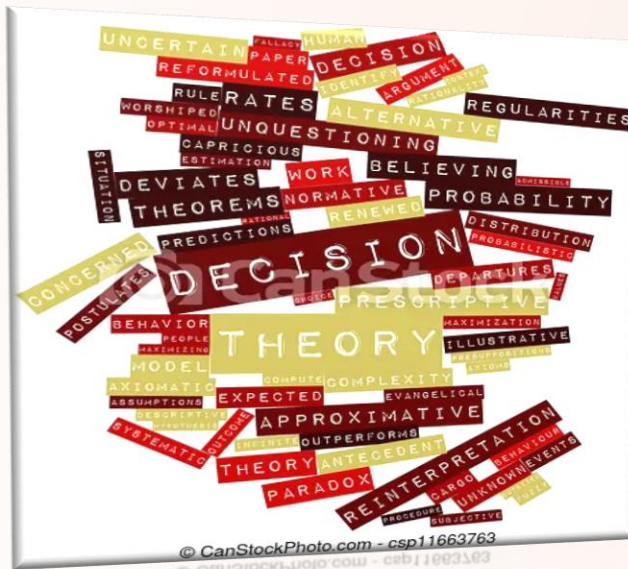


LÝ THUYẾT QUYẾT ĐỊNH

(Decision theory)



Biên soạn: ThS. Phạm Đình Tài
pdtai@ntt.edu.vn
0985.73.39.39

LÍ THUYẾT TRÒ CHƠI

- *Sự khó xử của tù nhân*
- *Phân loại trò chơi*
- *Kiến thức chung và lập luận ưu thế*
- *Trò chơi hai người tổng 0*
- *Các mục tiêu hỗn hợp và định lí minimax*

1. Tình trạng khó xử của tù nhân

Ví dụ:

- Cảnh sát đã bắt được hai kẻ buôn bán ma túy, Row và Col.
- Họ đang ngồi trong hai phòng giam riêng biệt trong đồn cảnh sát. Đương nhiên, họ không thể giao tiếp với nhau. Họ có một giờ để quyết định nên thú nhận tội ác của mình hay bác bỏ cáo buộc. Công tố viên, không có đủ bằng chứng để kết tội họ về tất cả các hành vi phạm tội của họ trừ khi ít nhất một trong số họ quyết định thú tội.





1. Tình trạng khó xử của tù nhân

Ví dụ:

- Tình huống pháp lý như sau:
 - ✓ Nếu **cả hai** tù nhân đều **thú tội**, họ sẽ bị **10 năm mỗi người**.
 - ✓ Tuy nhiên, nếu **một người thú tội** và người kia không thú nhận, thì tù nhân thú nhận sẽ được thưởng và chỉ được **miễn 1 năm tù**, trong khi người kia sẽ **nhận được 20 năm**.
 - ✓ Hơn nữa, nếu **cả hai** tù nhân đều **phủ nhận** cáo buộc, họ sẽ bị kết án **2 năm tù**.

1. Tình trạng khó xử của tù nhân

		Col	
		Thú tội	Tù chối
Row	Thú tội	- 10, - 10	- 1, - 20
	Tù chối	- 20, - 1	- 2, - 2

- Row có thể chọn giữa thú nhận hoặc từ chối các cáo buộc. Giả sử Row biết rằng Col sẽ thú nhận. Sau đó, Row về cơ bản sẽ đưa ra lựa chọn giữa 10 hoặc 20 năm tù.
- Vì vậy, trong trường hợp đó rõ ràng sẽ tốt hơn nếu thú nhận và nhận được 10 năm. Hơn nữa, nếu Row biết rằng Col sẽ phủ nhận các cáo buộc, Row thực tế sẽ phải đưa ra quyết định giữa việc nhận 1 hoặc 2 năm tù. Trong trường hợp đó, tất nhiên, tốt hơn là nên thú nhận và nhận 1 năm thay vì 2.
- Do đó, Row nên thú nhận bất kể Col quyết định làm gì.



1. Tình trạng khó xử của tù nhân

Có hai cách hơi khác nhau để Col có thể lập luận:

- **Đầu tiên**, Col có thể cho rằng Row sẽ thú nhận hoặc không. Sau đó, vì kết quả cho Col sẽ tốt hơn bất kể Row quyết định làm gì - vì 10 năm tốt hơn 20, và 1 tốt hơn 2 năm → nên Col nên thú nhận bất kể Row quyết định làm gì. Do đó, cả hai người chơi sẽ tự thú và lãnh án tù mỗi người 10 năm. Điều này đúng, mặc dù thực tế là cả hai tù nhân sẽ chỉ nhận được 2 năm nếu họ quyết định không thú tội.
- **Thứ hai**: Thay vì xem xét phản ứng tốt nhất của mình đối với tất cả các chiến lược được thực hiện bởi Row, Col có thể bắt đầu đơn giản từ giả định rằng Row là hợp lý. Nếu Row hợp lý, → sẽ thú nhận. Nếu Col biết Row có lý trí, ông ta cũng biết Row sẽ thú nhận.
→ có thể cả hai người chơi sẽ thú nhận.



1. Tình trạng khó xử của tù nhân

- Tình trạng tiến thoái lưỡng nan của các tù nhân cho thấy rằng những gì tối ưu cho mỗi cá nhân không cần phải trùng với những gì tối ưu cho nhóm.
- Tính hợp lý của cá nhân đôi khi mâu thuẫn với tính hợp lý của nhóm.
- Trong tình huống khó xử của tù nhân, tính hợp lý buộc cả hai người chơi phải chọn một kết quả tồi tệ hơn, cho cả hai người, so với kết quả mà họ đã đồng ý nếu họ hoạt động như một nhóm tìm cách giảm thiểu tổng số năm tù.
- Trong trường hợp này, cả hai đều nhận được mười năm thay vì hai.

1. Tình trạng khó xử của tù nhân

- Các tình huống khó xử của tù nhân nảy sinh bất cứ khi nào một trò chơi đối xứng, về cơ bản đều phải đối mặt với các chiến lược và kết quả giống nhau, và xếp hạng của các kết quả:

	Họ làm X	Họ không làm
Tôi làm X	2, 2	4, 1
Tôi không làm	1, 4	3, 3

- Lưu ý rằng các con số chỉ đề cập đến một số xếp hạng thứ tự của các kết quả. Tất cả những gì quan trọng là 4 tốt hơn 3 và 2 tốt hơn 1.*
- Trong nhiều trò chơi, người ta cho rằng tiện ích được đo trên một khoảng thời gian quy mô, nhưng trong tình trạng tiến thoái lưỡng nan của tù nhân, giả định khá mạnh mẽ này là thừa.*



2. Phân loại trò chơi

Những điểm khác biệt sau:

- Trò chơi có tổng bằng 0 và không có tổng
- Trò chơi không hợp tác so với trò chơi hợp tác
- Trò chơi di chuyển đồng thời so với di chuyển tuần tự
- Trò chơi có thông tin hoàn hảo so với trò chơi có thông tin không hoàn hảo.
- Trò chơi không đối xứng và đối xứng
- Trò chơi hai người so với n người
- Trò chơi không lặp lại so với trò chơi lặp lại

2. Phân loại trò chơi

a. Trò chơi có tổng bằng 0 và không có tổng

- *Trong trò chơi có tổng bằng không, bạn thắng chính xác bằng (các) đối thủ của bạn thua.*
- *Ví dụ điển hình bao gồm trò chơi sòng bạc và trò chơi trong phòng khách như cờ vua. Tổng số tiền hoặc đơn vị tiện ích là cố định cho dù điều gì xảy ra. Tuy nhiên, tình thế tiến thoái lưỡng nan của tù nhân không phải là một trò chơi có tổng bằng không, đơn giản là vì tiêu chí này không được thỏa mãn.*

2. Phân loại trò chơi

b. Trò chơi bất hợp tác và hợp tác

- Không được định nghĩa về việc người chơi có thực sự hợp tác hay không.
- Định nghĩa chính thức của một trò chơi bất hợp tác là một trò chơi trong đó người chơi không thể hình thành các thỏa thuận ràng buộc, bởi vì trong một số trò chơi bất hợp tác, người chơi thực sự hợp tác.
- *Các trò chơi không mang tính bất hợp tác được gọi là trò chơi hợp tác.* Theo định nghĩa, người chơi chơi một trò chơi hợp tác có thể đồng ý về các hợp đồng ràng buộc buộc họ phải tôn trọng bất cứ điều gì họ đã đồng ý.

2. Phân loại trò chơi

c. Trò chơi di chuyển đồng thời và trò chơi tuần tự

- Là trò chơi trong đó người chơi quyết định chiến lược của họ mà không cần biết (những) người chơi khác sẽ làm gì. Do đó, người chơi không nhất thiết phải chọn chiến lược của họ vào cùng một thời điểm.
- Tất cả những điều quan trọng là không người chơi nào biết (những) người chơi kia đã quyết định gì trước khi thực hiện động tác
- Trong các trò chơi tuần tự, những người chơi có một số (hoặc đầy đủ) thông tin về các chiến lược của những người chơi khác trong các vòng trước đó.

2. Phân loại trò chơi

Trò chơi di chuyển đồng thời

Ví dụ:

Kéo, giấy và đá, một trò chơi được chơi bởi trẻ em ở Mỹ và châu Âu, là một ví dụ về trò chơi di chuyển đồng thời. Cả hai người chơi thông báo một trong ba chiến lược thay thế, đó là **giấy**, **đá** hoặc **kéo**, mà không biết người chơi kia sẽ làm gì. Theo quy tắc của trò chơi này, **giấy đập đá, đá đập cây kéo và cây kéo đập giấy**

Trò chơi di chuyển tuần tự

Hai nhà sản xuất ô tô muốn tối đa hóa lợi nhuận của họ bằng cách giảm giá khác nhau cho khách hàng tiềm năng của họ. Cả hai công ty đều đưa ra quyết định mới về chính sách chiết khấu của mình ngay sau khi họ biết về các ưu đãi mới nhất của công ty khác; để đơn giản, giả định rằng công ty đầu tiên quyết định giảm giá vào các ngày Thứ Hai, Thứ Tư và Thứ Sáu, trong khi công ty kia đưa ra quyết định vào Thứ Ba, Thứ Năm và Thứ Bảy.

2. Phân loại trò chơi

d. Trò chơi có thông tin hoàn hảo

- Tập hợp con thú vị nhất của tập hợp các trò chơi tuần tự là những tập mà những người chơi có đầy đủ thông tin về các chiến lược được chơi bởi (những) người chơi khác.
- Trò chơi đáp ứng tiêu chí này được gọi là trò chơi có thông tin hoàn hảo. Do đó, tất cả các trò chơi có thông tin hoàn hảo đều là trò chơi tuần tự, nhưng không phải tất cả các trò chơi tuần tự đều là trò chơi có thông tin hoàn hảo.
- Cờ vua là một ví dụ điển hình về một trò chơi có thông tin hoàn hảo, cũng như trò chơi con rết được thảo luận trong

2. Phân loại trò chơi

e. Trò chơi đối xứng và bất đối xứng

- Nhiều trò chơi được các nhà lý thuyết trò chơi nghiên cứu là đối xứng, bao gồm hầu hết các phiên bản của tình thế tiến thoái lưỡng nan của tù nhân.

	Tầng 1	Tầng 2		Tầng 1	Tầng 2
Tầng 1	0, 1	0, 0	Tầng 1	0, 1	0, 0
Tầng 2	0, 0	1, 0	Tầng 2	0, 0	0, 1
(1)			(2)		

- Trong ma trận (1) minh họa trò chơi đối xứng, trong khi ma trận (2) minh họa trò chơi không đối xứng.
- Trò chơi đối xứng là trò chơi trong đó tất cả người chơi phải đối mặt với các chiến lược và kết quả giống nhau. Điều này có nghĩa là danh tính của người chơi không quan trọng.

2. Phân loại trò chơi

g. Trò chơi hai người so với n người

- Nhiều trò chơi tương đối dễ phân tích miễn là chỉ có hai người chơi, nhưng chúng có thể trở nên khó phân tích hơn đáng kể nếu có n người chơi.
- **Ví dụ**, nếu bạn chơi trò chơi nào đó với 5 người thay vì chỉ một người, bạn phải suy nghĩ kỹ hơn một chút trước khi quyết định chọn chiến lược nào. Tất nhiên, vì trò chơi n người không phù hợp với ma trận hai chiều truyền thống, nên những trò chơi như vậy cũng khó minh họa.

2. Phân loại trò chơi

Trò chơi hai người so với n người

- Sự khác biệt tiếp theo là giữa chiến lược hỗn hợp và thuần túy. Nói một cách ngắn gọn, chơi một chiến lược hỗn hợp có nghĩa là chơi một chiến lược thuần túy với một số xác suất nằm trong khoảng từ 0 đến 1.
- Ví dụ, để chơi một chiến lược hỗn hợp trong tình huống tiến thoái lưỡng nan của tù nhân có nghĩa là thú tội với xác suất p và từ chối cáo buộc với xác suất $1 - p$, trong đó $p > 0$. Nghĩa là, người chơi chọn chiến lược thuần túy để chơi bằng cách sử dụng một số ngẫu nhiên.

2. Phân loại trò chơi

h. Trò chơi không lặp lại so với trò chơi lặp lại

- Một số trò chơi có thể kéo dài mãi mãi mà không có hồi kết, trong khi các trò chơi khác luôn kết thúc sau một loạt chuyển động hữu hạn. Loại trò chơi trước đây được gọi là trò chơi lặp lại vô hạn.
- **Ví dụ**, những người đi làm lựa chọn giữa việc tự lái ô tô của họ đến nơi làm việc hoặc đi xe buýt sẽ chơi một trò chơi lặp đi lặp lại vô hạn. (Thực tế là con người chỉ sống trong một số ngày hữu hạn không được quan tâm nhiều từ quan điểm lý thuyết thuần túy.)

3. Kiến thức chung và lập luận ưu thế

- Tất cả các nhà lý thuyết trò chơi đều đồng ý rằng sẽ **thật tuyệt nếu có một lý thuyết thống nhất**, duy nhất có thể được áp dụng để giải quyết tất cả các loại trò chơi khác nhau. Không có lý thuyết thống nhất như vậy.
- Cho đến nay, các nhà lý thuyết trò chơi chỉ đưa ra các lý thuyết giải quyết chính xác các lớp con cụ thể của trò chơi, chẳng hạn như trò chơi hai người có tổng bằng không.
- Ngay cả khi không có một giải pháp chung nào có thể áp dụng cho tất cả các trò chơi, nhiều trò chơi trên thực tế có thể được giải quyết bằng cách chỉ áp dụng nguyên tắc thống trị một cách khéo léo.

3. Kiến thức chung và lập luận ưu thế

- Để giải quyết một trò chơi là tìm ra chiến lược mà người chơi hợp lý sẽ chọn nếu đối đầu với trò chơi được đề cập. Các chiến lược đã chọn quyết định kết quả của trò chơi.

	C1	C2	C3
R1	1, 3	2, 2	1, 0
R2	3, 2	3, 3	2, 7 *
R3	1, 1	8, 2	1, 4

→ cặp chiến lược được đánh dấu bằng dấu * trong ma trận, là một giải pháp cho trò chơi.

Trong trò chơi này, Row chắc chắn sẽ không chơi chiến lược R1, vì chiến thuật đó bị chi phối bởi chiến lược R2. Bất kể Col quyết định theo chiến lược nào, Row sẽ tốt hơn nếu chơi R2 hơn là R1. Do đó, cả hai người chơi biết chắc chắn rằng Row sẽ chơi R2 hoặc R3.

Điều này có nghĩa là chiến lược C3 là một chiến lược hấp dẫn đối với Col, vì Col biết rằng Col sẽ không kết thúc bằng 0. Nói chính xác hơn, vì Row sẽ không chơi R1, chiến lược C3 chiếm ưu thế hơn C2 và C1. Do đó, cả hai người chơi có thể kết luận rằng Col sẽ chơi chiến thuật C3. Hơn nữa, vì Col sẽ chơi C3, Row sẽ tốt hơn khi chơi chiến thuật R2 hơn là R3. Do đó, Col sẽ chơi chiến thuật C3 trong khi Row sẽ chơi chiến lược R2.



3. Trò chơi 2 người có tổng bằng 0

- Loại trò chơi được nghiên cứu rộng rãi nhất là trò chơi hai người có tổng bằng không. Nó đã được phân tích kỹ lưỡng bởi von Neumann và Morgenstern vào những năm 1940.
- Như đã giải thích trong phân loại, trò chơi hai người có tổng bằng không được đặc trưng bởi hai tính năng.
- Đầu tiên, nó luôn được chơi bởi chỉ hai người chơi (thay vì 3, hoặc 4, hoặc n người chơi).
- Thứ hai, bất kỳ số lượng tiện ích nào mà một người chơi thu được sẽ bị mất bởi người kia.

3. Trò chơi 2 người có tổng bằng 0

Định nghĩa: Một cặp chiến lược ở trạng thái cân bằng nếu và chỉ khi nó cho rằng một khi cặp chiến lược này được chọn, không người chơi nào có thể đạt được kết quả tốt hơn bằng cách đơn phương chuyển sang chiến lược khác.

	C1	C2	C3		C1	C2	C3
R1	9	8	7 *	R1	4	1 *	2
R2	7	- 5	6	R2	5	1 *	9
R3	4	1	- 2	R3	2	- 3	- 2

(1)

(2)

Trong (2), việc Row chơi chiến lược R1 hay R2 không quan trọng; cho rằng Col chơi chiến lược C2 cả hai lựa chọn thay thế tạo thành điểm cân bằng.

Rõ ràng, (R1, C3) trong (1) ở trạng thái cân bằng theo tiêu chí, vì 7 là giá trị nhỏ nhất của hàng 1, cũng như giá trị lớn nhất của cột thứ 3. (dấu *). Cũng khá dễ hiểu tại sao tiêu chí minimax là tiêu chí đủ: Vì 7 là giá trị cao nhất của cột được đề cập, Row sẽ không có lý do gì để chọn bất kỳ chiến lược khác cho rằng Col chơi chiến lược C3. Hơn nữa, vì 7 là giá trị nhỏ nhất của hàng, Col không thể làm tốt hơn khi cô biết Row sẽ chơi chiến thuật R1. → chính là lý do tại sao (R1, C3) ở trạng thái cân bằng.

4. Các chiến lược hỗn hợp và định lí minimax

	C1	C2	C3
R1	0	- 100	+100
R2	+100	0	- 100
R3	- 100	+100	0

- Ma trận minh họa một trò chơi do hai đội quân vĩ đại chơi. Row, chỉ huy quân đội R, chuẩn bị chỉ thị cho quân của mình tấn công quân C bằng cách sử dụng 1 trong 3 chiến lược thay thế từ R1 đến R3. Col phải tự vệ trước cuộc tấn công của R bằng cách chọn chính xác một trong 3 chiến lược phòng thủ từ C1 đến C3.
- *Tìm ra giải pháp cho trò chơi này là vô cùng quan trọng !?*

4. Các chiến lược hỗn hợp và định lí minimax

Giải pháp:

- Col: Nếu tôi ném một con súc sắc và chọn mỗi chiến lược thuần túy với xác suất là $1/3$, thì tiện ích mong đợi mà tôi lựa chọn sẽ bằng 0, bất kể xác suất mà Row chỉ định cho các chiến lược của mình. Như trước đây, nếu Col biết Row sẽ chơi chiến lược hỗn hợp tương tự, thì Col sẽ không có lý do gì để chuyển sang bất kỳ chiến lược hỗn hợp nào khác, vì 0 khi đó là tiện ích mong đợi cao nhất mà ông ta có thể có được.

	C1	C2	C3
R1	0	- 100	+100
R2	+100	0	- 100
R3	- 100	+100	0

Row có thể kết luận rằng tiện ích mong đợi khi chơi chiến thuật hỗn hợp

$[R1 \ 1/3, R2 \ 1/3, R3 \ 1/3]$ có tiện ích mong đợi là 0, bất kể người kia quyết định làm gì.



4. Các chiến lược hỗn hợp và định lí minimax

Gọi p là xác suất Col chọn C1, và gọi q là xác suất chọn C2, và gọi r là xác suất Col chọn C3. Sau đó, tiện ích dự kiến của Row khi chơi $[R1 \ 1/3, R2 \ 1/3, R3 \ 1/3]$ là:

$$\begin{aligned} & 1/3(0p - 100q + 100r) + 1/3(100p + 0q - 100r) + 1/3(-100p + 100q + 0r) \\ &= 1/3 (0p - 100p + 100p - 100q + 0q + 100q + 100r + 0r) \\ &= 0 \end{aligned}$$

Mọi trò chơi có tổng bằng 0 của hai người đều có một giải pháp, tức là một cặp chiến lược ở trạng thái cân bằng.

Định lí minimax

Mọi trò chơi có tổng bằng 0 của hai người đều có một nghiệm, tức là luôn có một cặp chiến lược cân bằng và nếu có nhiều hơn một cặp thì tất cả chúng đều có cùng công dụng mong đợi.

THANK YOU !

