

CNTT4899

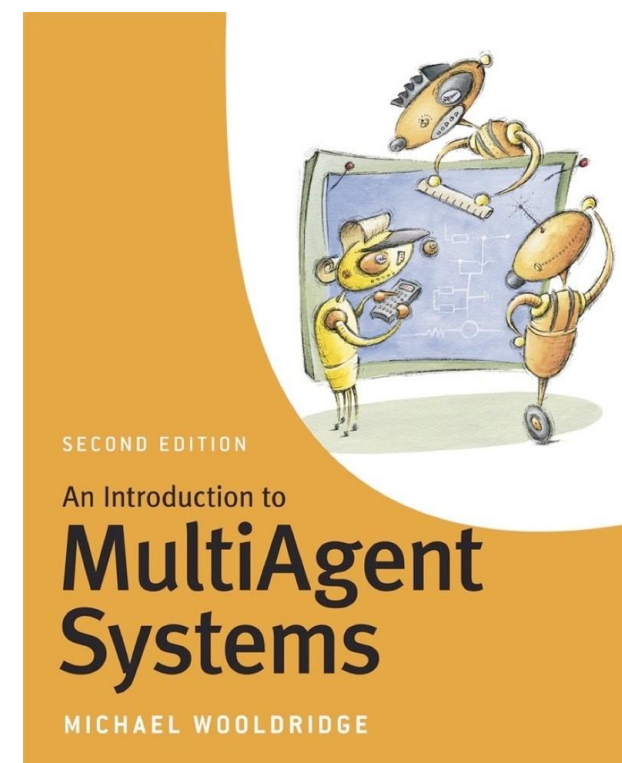
## Hệ thống đa tác nhân

Chương 11 - Tương tác đa tác tử

TS Nguyễn Bình Minh Bộ môn HTTT



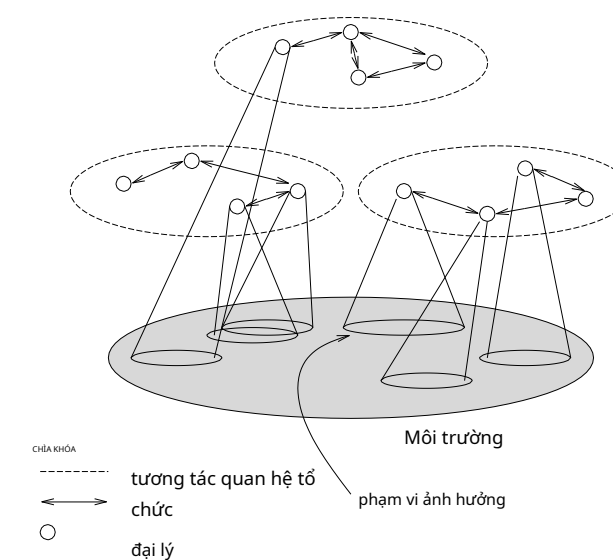
1



## Hệ thống đa tác nhân là gì?

- Một hệ thống đa tác nhân chứa một số tác nhân:
  - tương tác thông qua giao tiếp;
  - có khả năng hành động trong một môi trường;
  - có “phạm vi ảnh hưởng” khác nhau (có thể trùng nhau); Và
  - sẽ được liên kết bởi các mối quan hệ (tổ chức) khác.
- Chúng ta sẽ xem xét cách các đại lý quyết định cách tương tác trong *cạnh tranh* tình huống.

2



Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

2

1

# Tiện ích và Sở thích

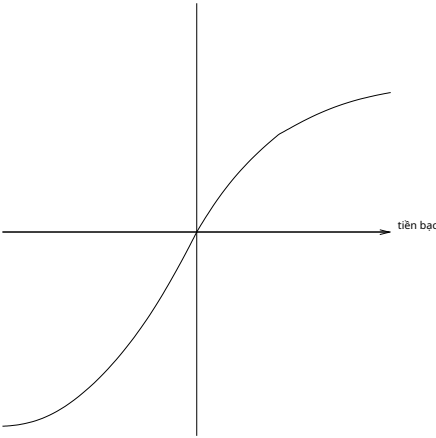
- Giả định của chúng tôi:
  - Giả sử chúng ta chỉ có hai đại lý:  $Ag = \{i, j\}$
  - Đại lý được coi là **tư lợi** tức là họ có sở thích về môi trường như thế nào.
- Cho rằng  $\Omega = \{\omega_1, \omega_2, \dots\}$  là tập hợp các "kết quả" mà các tác nhân có quyền ưu tiên hơn.
- Chúng tôi nắm bắt sở thích bằng cách **Các chức năng tiện ích**, được biểu diễn dưới dạng số thực ( $\mathbb{R}$ ):

$$b_{an}i: \Omega \rightarrow \mathbb{R}$$
$$b_{an}j: \Omega \rightarrow \mathbb{R}$$

- Các chức năng tiện ích dẫn đến **thứ tự ưu tiên** trên kết quả, ví dụ:

$$\omega \succ_{Ti} \omega' \text{ có nghĩa } b_{an}Ti(\omega) \geq b_{an}Ti(\omega') \text{ và}$$
$$\omega \succ_{Tj} \omega' \text{ có nghĩa } b_{an}Tj(\omega) > b_{an}Tj(\omega')$$

- trong đó  $\omega$  và  $\omega'$  đều là kết quả có thể xảy ra của  $\Omega$



Tiện ích không phải là tiền. chỉ là một cách để mã hóa sở thích.

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# Cuộc gặp gỡ đa tác nhân

- Chúng ta cần một mô hình môi trường trong đó các tác nhân này sẽ hành động...
  - các tác nhân đồng thời chọn một hành động để thực hiện và kết quả của các hành động họ chọn là kết quả trong  $\Omega$  sẽ dẫn đến
  - các **thật sự** kết quả phụ thuộc vào **sự kết hợp** hành động
  - giả sử mỗi tác nhân chỉ có hai hành động khả thi mà nó có thể thực hiện:
  - I  $EAc = \{C, D\}$ , Ở đâu
    - $C$  ("hợp tác") và
    - $D$  ("khuyết điểm")
- Hành vi môi trường được đưa ra bởi **hàm biến trạng thái  $\tau$** 
  - (đã giới thiệu ở Chương 2):

$$\tau : \underbrace{Ac}_{\text{agent } i\text{'s action}} \times \underbrace{Ac}_{\text{agent } j\text{'s action}} \rightarrow \Omega$$

4

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

Cuộc gặp gỡ đa tác nhân

- Đây là một hàm biến trạng thái  $\tau(tôi, j)$ 
  - Môi trường này nhạy cảm với hành động của cả hai tác nhân.
- Với bộ chuyển đổi trạng thái này, không có tác nhân nào có bất kỳ ảnh hưởng nào trong môi trường này.
- Với cái này, môi trường được kiểm soát bởi  $j$

$\tau(D, D) = \omega_1$  $\tau(D, C) = \omega_2$

$\tau(C, D) = \omega_3$  $\tau(C, C) = \omega_4$

$\tau(D, D) = \omega_1$  $\tau(D, C) = \omega_1$

$\tau(C, D) = \omega_1$  $\tau(C, C) = \omega_1$

$\tau(D, D) = \omega_1$  $\tau(D, C) = \omega_2$

$\tau(C, D) = \omega_1$  $\tau(C, C) = \omega_2$

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

hành động hợp lý

- Giả sử chúng ta có trường hợp cả hai tác nhân có thể ảnh hưởng đến kết quả và chúng có các hàm tiện ích sau:

$bạn_{Tôi}(\omega_1)=1$  $bạn_{Tôi}(\omega_2)=1$  $bạn_{Tôi}(\omega_3)=4$  $bạn_{Tôi}(\omega_4)=4$

$bạn_j(\omega_1)=1$  $bạn_j(\omega_2)=4$  $bạn_j(\omega_3)=1$  $bạn_j(\omega_4)=4$
- Với một chút lạm dụng ký hiệu:

$bạn_{Tôi}(D, D) = 1$  $bạn_{Tôi}(D, C) = 1$  $bạn_{Tôi}(C, D) = 4$  $u_{Tôi}(C, C) = 4$

$u_j(D, D) = 1$  $bạn_j(D, C) = 4$  $u_j(C, D) = 1$  $bạn_j(C, C) = 4$
- Sau đó đại lý  $Tôi$  sở thích của là  $(C, C) \succ_{Tôi} (Đ, Đ) \succ_{Tôi} (Đ, C) \succ_{Tôi} (C, D)$
- Trong trường hợp này, những gì nên  $Tôi$  LÀM?
- $Tôi$  thích tất cả các kết quả phát sinh thông qua  $C$  trên tất cả các kết quả phát sinh thông qua  $Đ$ .
- Như vậy  $C$  là *lựa chọn hợp lý* vì  $Tôi$ .

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# Ma trận xuất chi

- Chúng ta có thể mô tả kịch bản trước đó trong một *ma trận xuất chi* thể hiện ngược lại
  - Đại lý *Tôi* là *người chơi cột* và nhận được *phần thưởng cao hơn* trong một tế bào.
  - Đại lý là *người chơi hàng* và nhận được *phần thưởng thấp hơn* trong một tế bào.
- Trên thực tế có hai ma trận ở đây, một (gọi nó là *MỘT*) xác định số tiền chi trả cho *Tôi* và một cái khác *b* chỉ định tiền chi trả cho *j*.
- Đôi khi chúng ta sẽ viết trò chơi như  $(A, B)$  để công nhận điều này.

Trong trường hợp này, *Tôi* hợp tác và đạt được một tiện ích là 4; trong khi *khiếm khuyết* và lợi ích một tiện ích chỉ có 1.

		<i>Tôi</i>	
		khuyết điểm	chuồng
<i>j</i>	khuyết điểm	1   1   4	
	chuồng	1   1   4	4   4   4

$(C, C) \succ_{\text{Tôi}} (Đ, A) \succ_{\text{Tôi}} (Đ, C) \succ_{\text{Tôi}} (Đ, Đ)$

# Khái niệm giải pháp

- Một tác nhân hợp lý sẽ hành xử như thế nào trong bất kỳ tình huống nào?
- Chơi. . .
  - chiến lược ưu thế;
  - Chiến lược cân bằng Nash;
  - Các chiến lược tối ưu Pareto;
  - chiến lược tối đa hóa phúc lợi xã hội.

Chiến lược chiếm ưu thế

- Đưa ra bất kỳ chiến lược cụ thể  $S$  (hoặc  $C$  hoặc  $D$ .) đại lý đó  $Tôi$  có thể chơi, sẽ có là một số kết quả có thể xảy ra.
- Chúng tôi nói  $S$  thống trị  $S_2$  nếu mọi kết quả có thể bởi  $Tôi$  đang chơi  $S$  được ưa thích hơn mọi kết quả có thể bởi  $Tôi$  đang chơi  $S_2$ .
- Vì vậy, trong trò chơi ngược lại,  $C$  thống trị  $D$  cho cả hai người chơi.

		<i>Tôi</i>	
		khuyết điểm	chuồng
<i>j</i>	khuyết điểm	1	4
	chuồng	1	4

Chiến lược chiếm ưu thế

- Một tác nhân hợp lý sẽ không bao giờ chơi một *chiến lược thống trị*.
  - tức là, một chiến lược bị chi phối (và do đó kém hơn) bởi một chiến lược khác
- Vì vậy, khi quyết định phải làm gì, chúng ta có thể loại bỏ các chiến lược bị chi phối.
  - Thật không may, không phải lúc nào cũng có một chiến lược độc nhất vô nhị.

		<i>Tôi</i>	
		khuyết điểm	chuồng
<i>j</i>	khuyết điểm	1	4
	chuồng	1	4

## Trạng thái cân bằng Nash

- Nói chung, chúng ta sẽ nói rằng hai chiến lược  $S_1$  và  $S_2$  ở trạng thái cân bằng Nash (NE) nếu:
  - với giả định rằng đại lý  $Tôi$  ở kịch  $S_1$ , đại lý  $j$  không thể làm gì tốt hơn là chơi  $S_2$ ;
    - I Nếu như **TÔI** lái xe bên trái đường, **BẠN** không thể làm gì tốt hơn là lái xe bên trái!
  - với giả định rằng đặc vụ  $j$  chơi  $S_2$ , đại lý  $Tôi$  không thể làm gì tốt hơn là chơi  $S_1$ .
    - I Nếu như **BẠN** lái xe bên trái đường, **TÔI** không thể làm gì tốt hơn là lái xe bên trái!
- Không đại lý nào có bất kỳ động cơ nào để đi chệch khỏi một **Cân bằng Nash (NE)**.



Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

## Trạng thái cân bằng Nash

- Xem xét ma trận xuất chi ngược lại:
  - đây **Cân bằng Nash (NE)** là  $(Đ, Đ)$ .
  - Trong một trò chơi như thế này, bạn có thể tìm thấy NE bằng cách xem xét các kết quả, hỏi xem liệu một trong hai tác nhân có thể cải thiện phần thưởng của mình bằng cách chuyển đổi chiến lược hay không.
- Vì vậy, ví dụ,  $(ĐĨA CD)$  không phải là ĐB vì  $Tôi$  có thể chuyển số tiền chi trả từ 1 thành 5 bằng cách chuyển từ  $Đ$  ĐẾN  $Đ$ .

		Tôi	
		khuyết điểm	chuông
j	khuyết điểm	3 5	2 1
	chuông	2 0	1 0

		Tôi	
		khuyết điểm	chuông
j	khuyết điểm	3 5	2 1
	chuông	2 0	1 0

12

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

Trạng thái cân bằng Nash

- Chính thức hơn:
  - Một chiến lược  $(Tô_i, j_i)$  là một *chiến lược thuần túy Cân bằng Nash* giải pháp cho trò chơi  $(A, B)$  nếu như:

$-tô_i, một_{Tô_i,j} \geq một_{Tô_i,j}$   
 $-j, b_{Tô_i,j} \geq b_{Tô_i,j}$

- Không may thay:
  - Không phải mọi kịch bản tương tác đều có *chiến lược thuần túy* Cân bằng Nash (NE).
  - Một số kịch bản tương tác có *nhều hơn một chiến lược thuần túy* Cân bằng Nash (NE).

Trạng thái cân bằng Nash

- Trò chơi ngược lại (phía trên) có hai NE chiến lược thuần túy,  $(C, C)$  và  $(Đ, Đ)$ 
  - Trong cả hai trường hợp, một tác nhân đơn lẻ không thể đơn phương cải thiện lợi nhuận của mình.

		Tôi	
		khuyết điểm	chuồng
j	khuyết điểm	5 3	1 2
	chuồng	0 2	3 3

- Trong trò chơi ngược lại trò chơi (thấp hơn) không có chiến lược thuần túy NE
  - Đối với mọi kết quả, một trong các tác nhân sẽ cải thiện tiện ích của mình bằng cách chuyển đổi chiến lược.
  - Chúng ta có thể tìm thấy một dạng NE trong những trò chơi như vậy, nhưng chúng ta cần vượt ra ngoài những chiến lược thuần túy.

		Tôi	
		khuyết điểm	chuồng
j	khuyết điểm	2 1	1 2
	chuồng	0 2	1 1



# Chiến lược hỗn hợp Cân bằng Nash

- đồng xu phù hợp
  - người chơi *Tôi* và/hoặc đồng thời chọn mặt đồng xu, hoặc là “ngửa” hoặc “sấp”.
  - Nếu họ hiển thị cùng một khuôn mặt, sau đó *Tôi* thắng, trong khi nếu họ thể hiện những khuôn mặt khác nhau, thì *chị* chiến thắng.
- KHÔNG có cặp chiến lược nào tạo thành một chiến lược thuần túy NE:
  - bất kể cặp chiến lược nào được chọn, ai đó sẽ ước họ đã làm điều gì đó khác.
- Giải pháp là cho phép *chiến lược hỗn hợp*:
  - chơi “ngửa” với xác suất 0,5
  - chơi “sấp” với xác suất 0,5.
- Đây là một chiến lược cân bằng Nash hỗn hợp.

		<i>Tôi</i>	
		cái đầu	đuôi
<i>j</i>	cái đầu	1 - 1	- 1 1
	đuôi	- 1 1	1 - 1



Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# Chiến lược hỗn hợp Cân bằng Nash

- Cân nhắc trò chơi Rock/Paper/Kéo
  - Giấy bìa đá
  - Kéo cắt giấy
  - kéo cùn đá
- Điều này có ma trận xuất chi sau

		<i>Tôi</i>		
		đá	kéo	giấy
<i>j</i>	đá	0 0	1 0	0 1
	giấy	0 1	0 0	1 0
	kéo	1 0	0 1	0 0

- Những gì bạn nên làm?
  - *Chọn một chiến lược một cách ngẫu nhiên!*



Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018



# Chiến lược hỗn hợp

- Một chiến lược hỗn hợp có dạng
  - chơi  $a_1$  với xác suất  $P_1$
  - chơi  $a_2$  với xác suất  $P_2$
  - ...
  - chơi  $a_k$  với xác suất  $P_k$ .
  - như vậy mà  $P_1 + p_2 + \dots + p_k = 1$ .
- Nash đã chứng minh rằng:
  - *mọi trò chơi hữu hạn đều có trạng thái cân bằng Nash trong các chiến lược hỗn hợp.*

**Định lý Nash**

Nash đã chứng minh rằng *mọi trò chơi hữu hạn đều có trạng thái cân bằng Nash trong các chiến lược hỗn hợp.* (Không giống như trường hợp của các chiến lược thuần túy.)

Vì vậy kết quả này đã khắc phục được tình trạng thiếu giải pháp; nhưng vẫn có thể có nhiều hơn một trạng thái cân bằng Nash. . .

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# Tối ưu Pareto

- Một kết quả được cho là *tối ưu Pareto* (hoặc *Pareto hiệu quả*) nếu như:
  - không có kết quả nào khác khiến một tác nhân *tốt hơn* mà không làm cái khác *đại lý tồi tệ hơn* tắt.
    - Nếu một kết quả là tối ưu Pareto, thì ít nhất một tác nhân sẽ miễn cưỡng rời khỏi nó (vì tác nhân này sẽ tồi tệ hơn).
    - Nếu một kết cục không phải là tối ưu Pareto, thì có một kết quả khác  $\omega$  điều đó làm cho mọi người hạnh phúc, nếu không muốn nói là hạnh phúc hơn  $\omega$ .
- Các đại lý “hợp lý” sẽ đồng ý chuyển đến  $\omega$  trong trường hợp này.
  - *Ngay cả khi tôi không được hưởng lợi trực tiếp từ  $\omega$ , bạn có thể hưởng lợi mà không cần tôi đau khổ.*

Trò chơi này có một Pareto hiệu quả kết quả: (Đ, Đ)

		Tôi	
		chuông bị lỗi	
j	khuyết điểm	3 5	1
	chuông	0 2	0
		2	1

Không có giải pháp nào trong đó một trong hai đại lý làm tốt hơn

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# Phúc lợi xã hội

- Phúc lợi xã hội của một kết quả  $\omega$  là tổng các tiện ích mà mỗi đại lý nhận được từ  $\omega$ :

$$\sum_{i \in Ag} u_i(\omega)$$

- Hãy nghĩ về nó như là “tổng số tiền trong hệ thống”.
- Là một khái niệm giải pháp:
  - có thể phù hợp khi toàn bộ hệ thống (tất cả các đại lý) có một chủ sở hữu duy nhất (khi đó lợi ích chung của hệ thống mới quan trọng chứ không phải cá nhân).
  - Nó không xem xét lợi ích cho các cá nhân.
  - Một kết quả rất sai lệch có thể tối đa hóa phúc lợi xã hội.

Trong cả hai trò chơi này,  $(C, C)$  tối đa hóa phúc lợi xã hội

$j$ 

	Tôi	
	khuyết điểm	chuồng
khuyết điểm	2	1
chuồng	3	4

$j$ 

	Tôi	
	khuyết điểm	chuồng
khuyết điểm	2	1
chuồng	3	0

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# Tương tác cạnh tranh và tổng bằng không



- Khi sở thích của các đại lý hoàn toàn trái ngược nhau, chúng ta có *cạnh tranh nghiêm ngặt* kịch bản.

- Các cuộc chạm trán có tổng bằng 0 là những cuộc gặp gỡ mà các tiện ích có tổng bằng 0:

$$bạn_{\omega}(w)+u_j(w)=0 \text{ cho tất cả } \omega \in \Omega.$$

- Các cuộc chạm trán có tổng bằng 0 là một tin xấu: để tôi có được + tiện ích thì bạn phải có được tiện ích âm! *Kết quả tốt nhất cho tôi là tồi tệ nhất cho bạn!*
- *Những cuộc gặp gỡ tổng bằng 0 trong đời thực rất hiếm.* . . nhưng mọi người có xu hướng hành động trong nhiều tình huống như thể chúng bằng không.
- Hầu hết các trò chơi đều có một số chỗ trong tập hợp các kết quả để các tác nhân tìm ra (phần nào) kết quả đôi bên cùng có lợi.

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

Thế tiến thoái lưỡng nan của người tù

- Ma trận hoàn trả cho thế tiến thoái lưỡng nan của người tù

		Tôi	
		khuyết điểm	chuông
j	khuyết điểm	2	1
	chuông	4	3

Như không phải dân ông  
muốn thừa nhận  
có tội,  
sự hợp tác có nghĩa  
không xưng tội!

Thế tiến thoái lưỡng nan của người tù

đàn ông bị buộc tội chung với tôi và bị giam trong các phòng giam riêng biệt, với bất kỳ cuộc gặp gỡ hoặc giao tiếp nào.

được bảo rằng:

- nếu một người xưng tội và người kia không (C,D) hoặc (D,C), cha giải tội sẽ là tự do, và người kia sẽ bị bỏ tù vì ba năm;
- nếu cả hai đều thú nhận (D,D) thì mỗi người sẽ bị phạt tù hai năm.

Cả hai tù nhân đều biết rằng nếu không ai thú tội (C,C), thì mỗi người sẽ bị bỏ tù một năm.

- Trên cùng bên trái: Nếu cả hai đều phản bội, thì cả hai đều bị trừng phạt vì tội phản bội lẫn nhau.
- Trên cùng bên phải: Nếu tôi hợp tác và j thất bại, tôi nhận được phần thưởng của kẻ ngốc là 1, trong khi j nhận được 4.
- Dưới cùng bên trái: Nếu j hợp tác và tôi phản bội, j sẽ nhận được phần thưởng của kẻ ngốc là 1, trong khi tôi nhận được 4.
- Góc phải ở phía dưới: Phần thưởng cho sự hợp tác lẫn nhau (tức là không thú nhận).

Những gì bạn nên làm?

- Các hành động hợp lý cá nhân là **kiếm khuyết**.
  - Điều này đảm bảo số tiền chi trả không thấp hơn 2, trong khi hợp tác đảm bảo số tiền chi trả nhiều nhất là 1.
  - Vì vậy, phản bội là phản ứng tốt nhất đối với tất cả các chiến lược khả thi: cả hai tác nhân đều phản bội và nhận được phần thưởng = 2.
- Nhưng **trực giác** nói đây là **không** kết quả tốt nhất:
  - Chắc chắn cả hai nên hợp tác và mỗi người nhận được 3 phần thưởng!
- Đây là lý do tại sao trò chơi Prisoners Dilemma thú vị
  - Việc phân tích dường như cung cấp cho chúng ta một **ngịch lý** trả lời.

Khái niệm giải pháp

- Chiến lược chiếm ưu thế ở đây là đào thoát.
- (D, D) là cân bằng Nash duy nhất.
- Tất cả các kết quả ngoại trừ (D, D) là tối ưu Pareto.
- (C, C) tối đa hóa phúc lợi xã hội.

		Tôi	
		chuông bị lỗi	
j	khuyết điểm	2	1
	chuông	4	3

Thế tiến thoái lưỡng nan của người tù

- Nghịch lý rõ ràng này là vấn đề cơ bản của các tương tác đa tác nhân.
  - Nó dường như ngụ ý rằng sự hợp tác sẽ không xảy ra trong xã hội của những tác nhân tư lợi.
- Ví dụ trong thế giới thực:
  - cắt giảm vũ khí hạt nhân -“tại sao tôi không giữ của tôi”
  - hệ thống người lái miễn phí - giao thông công cộng, chia sẻ tệp;
  - ở Anh — giấy phép truyền hình.
  - ở Mỹ — tài trợ cho NPR/PBS.
- Thế tiến thoái lưỡng nan của tù nhân là phổ biến.
  - Chúng ta có thể phục hồi hợp tác không?

Khái niệm giải pháp

- Chiến lược chiếm ưu thế ở đây là đào thoát.
- $(D, D)$  là cân bằng Nash duy nhất.
- Tất cả các kết quả ngoại trừ  $(D, D)$  là tối ưu Pareto.
- $(C, C)$  tối đa hóa phúc lợi xã hội.

		Tôi	
		chuồng bị lỗi	
j	khuyết điểm	2	1
	2	4	3
chuồng		1	3

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

Lập luận để phục hồi hợp tác

- Kết luận mà một số người đã rút ra từ phân tích này:
  - khái niệm lý thuyết trò chơi về hành động hợp lý là sai!
  - bằng cách nào đó tiến thoái lưỡng nan đang được xây dựng sai
- Lập luận để phục hồi hợp tác:
  - lòng vị tha
  - Người tù kia là song sinh của tôi!
  - Cân bằng chương trình và hòa giải
  - Cái bóng của tương lai. . .

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

## Tất cả chúng ta không phải là Machiavelli

- Rất nhiều “lòng vị tha” là một cái gì đó khác:
  - Hoặc là có một số phản ứng có đi có lại bị trì hoãn; hoặc
  - Có cơ chế trừng phạt đào tẩu.
- Có một lý do tại sao HMRC (hoặc IRS ở Hoa Kỳ) kiểm tra thuế của mọi người :-)
- Lòng vị tha có thể là thứ khiến chúng ta cảm thấy dễ chịu
  - Đây là lý do tại sao chúng tôi sẵn sàng trả tiền cho nó.

“... Chúng tôi không quá cứng rắn, và bên cạnh đó, mọi người thực sự hành động vị tha...”

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

25

## Tù nhân khác là song sinh của tôi



- Lập luận rằng cả hai tù nhân sẽ suy nghĩ giống nhau và quyết định rằng tốt nhất là nên hợp tác.
  - Nếu họ là song sinh, họ phải suy nghĩ theo cùng một hướng, phải không?
  - (Hoặc họ có một số thỏa thuận rằng họ sẽ không nói chuyện.)
- Chà, nếu đây là trường hợp, chúng ta không thực sự chơi Thế tiến thoái lưỡng nan của người tù!
- Có thể quan trọng hơn là nếu bạn biết người khác sẽ hợp tác, bạn **vẫn** đào tẩu thì tốt hơn.

26

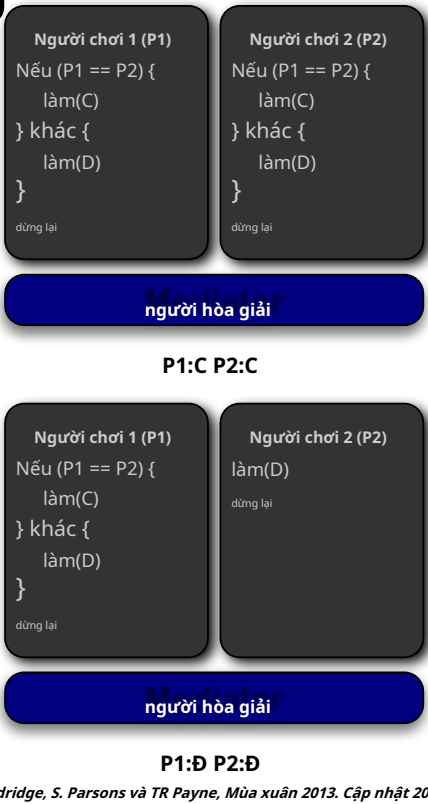
Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

26

# Chương trình cân bằng

- Chiến lược mà bạn thực sự muốn áp dụng trong thể tiến thoái lưỡng nan của tù nhân là:*Tôi sẽ hợp tác nếu anh ấy muốn*
  - Cân bằng chương trình cung cấp một cách để kích hoạt điều này.
- Mỗi đại lý gửi một*chiến lược chương trình* đến một*người hòa giải*cái mà*cùng thực hiện*các chiến lược.
  - Điều quan trọng là các chiến lược có thể được*có điều kiện dựa trên chiến lược của những người khác.*
  - Phản hồi tốt nhất cho chương trình này:
    - gửi cùng một chương trình*, đưa ra một kết quả của(C, C)

27



Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

## Thể tiến thoái lưỡng nan của người tù lặp đi lặp lại (Cái bóng của tương lai)

- Chơi trò chơi nhiều hơn một lần.
  - Nếu bạn biết bạn sẽ gặp lại đối thủ của mình, thì*động lực để đào tẩu* dường như bốc hơi.
    - Nếu bạn khiếm khuyết, bạn có thể*trừng phạt*(so với phần thưởng hợp tác.)
    - Nếu bạn bị hút, sau đó*những gì bạn mất có thể được khấu hao trong phần còn lại của các lần lặp lại*, làm cho nó một mất mát nhỏ.
- Hợp tác là (chứng minh)*sự lựa chọn hợp lý*trong thể tiến thoái lưỡng nan của tù nhân lặp đi lặp lại vô tận.
  - (Tiếng hoan hô!)
- Nhưng nếu có một số lần lặp lại hữu hạn thì sao?

28

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# cảm ứng ngược

- Nhưng. . . giả sử cả hai bạn đều biết rằng bạn sẽ chơi trò chơi đúng n lần.
  - Ở vòng thứ n - 1, bạn có động cơ để bỏ cuộc, để đạt được phần thưởng bổ sung đó.
  - Nhưng điều này làm cho vòng n - 2 trở thành vòng “thực” cuối cùng, và vì vậy bạn cũng có động cơ để bỏ qua vòng đó.
- Đây là *bài toán quy nạp ngược*.
- Chơi thể tiến thoái lưỡng nan của tù nhân với một số vòng cố định, hữu hạn, được xác định trước, thường được biết đến, đào thoát là chiến lược tốt nhất.
  - Điều đó dường như gợi ý rằng bạn không bao giờ nên hợp tác.
  - Vậy làm thế nào để hợp tác phát sinh? Tại sao nó có ý nghĩa?

29

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

Miễn là bạn có một số xác suất lặp lại thì sự hợp tác tương tác có thể có kết quả mong đợi tốt hơn.

Miễn là có đủ người hợp tác ngoài kia, bạn có thể tiến lên phía trước bằng cách hợp tác.


# Giải đấu của Axelrod

- Giả sử bạn chơi thể tiến thoái lưỡng nan của người tù (IPD) với một *phạm vi* của đối thủ.
- Bạn nên chọn cách tiếp cận nào để tối đa hóa lợi ích tổng thể của mình?
  - *đào ngũ có tốt hơn không*, và hy vọng tìm được những kẻ lừa đảo?
  - Hoặc *tốt hơn là hợp tác* và cố gắng tìm những người thân thiện khác để hợp tác cùng?

30

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

Robert Axelrod



Robert Axelrod (1984) đã điều tra vấn đề này, với một giải đấu máy tính dành cho các chương trình chơi song đề tù nhân lặp đi lặp lại.

Axelrod đã tổ chức giải đấu và nhiều nhà nghiên cứu đã gửi các phương pháp để chơi trò chơi.



# Chiến lược trong Giải đấu của Axelrod

- TIT-FOR-TAT đáng ngạc nhiên cho chiến thắng.
  - Nhưng đừng đọc quá nhiều về điều này :-)
- Trong các tình huống như giải đấu Thế lưỡng nan của người tù lặp đi lặp lại (IPD)...
- ...cách tiếp cận tốt nhất phụ thuộc rất nhiều vào *tập hợp đầy đủ các phương pháp là gì.*
- TIT-FOR-TAT đã hoạt động tốt vì có những người chơi khác mà nó có thể hợp tác.

**JOSS**

Như TIT-FOR-TAT, trừ khuyết điểm định kỳ.

**ALL-D**

“Luôn phản bội” - chiến lược điều hâu;

ăn miếng trả miếng

1. Trên vòng  $u = 0$ , hợp tác.  
2. Ở vòng  $u > 0$ , hãy làm những gì đối thủ của bạn đã làm ở vòng  $u - 1$ .

Kiểm thử

Vào vòng 1, khiếm khuyết. Nếu đối thủ trả đũa, hãy chơi TIT-FOR-TAT. Nếu không xen kẻ hợp tác & đào tẩu.

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# Bí quyết thành công trong Giải đấu của Axelrod

- **Đừng ghen tị:**
  - Đừng chơi như thể nó không có tổng!
- **Hãy tử tế:**
  - Bắt đầu bằng cách hợp tác và hợp tác có đi có lại.
- **Trả đũa thích hợp:**
  - Luôn trừng phạt hành vi phản bội ngay lập tức, nhưng sử dụng vũ lực “có chừng mực”
  - đừng lạm dụng nó.
- **Đừng giữ mối hận thù:**
  - Luôn luôn có đi có lại hợp tác ngay lập tức.

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# săn hươu

- Một nhóm thợ săn đi săn hươu.
  - Nếu như *tất cả họ đều tập trung vào con nai*, họ sẽ bắt được nó và tất cả đều có rất nhiều thức ăn.
  - Nếu như *một số trong số họ đi bắt thỏ*, con hươu sẽ trốn thoát.
- Trong trường hợp này, những người săn thỏ sẽ có một ít thức ăn và những người săn hươu (còn lại) sẽ bị đói.
- Mỗi thợ săn nên làm gì?

"...Bạn và một người bạn quyết định rằng sẽ là một trò đùa tuyệt vời khi xuất hiện vào ngày cuối cùng của năm học với một kiểu tóc lở bích nào đó. Cả hai đều được ủng hộ bởi nhóm của bạn ~~th~~ bạn sẽ được cắt tóc.

Sau đó là một đêm do dự. Khi bạn dự đoán phản ứng của cha mẹ và giáo viên [ . . . ] bạn bắt đầu tự hỏi liệu bạn mình có thực sự hoàn thành kế hoạch hay không.

Không phải là bạn không muốn kế hoạch thành công; kết quả tốt nhất có thể là cả hai bạn cùng đi cắt tóc.

Vấn đề là, nó sẽ là khủng khiếp nếu là **chỉ một** để xuất hiện với mái tóc. Đó sẽ là kết quả tồi tệ nhất có thể.

Bạn không thích thú với sự xấu hổ của bạn mình. nếu bạn **không** đi cắt tóc, nhưng người bạn đó đã làm, và trông giống như một thằng ngốc thực sự, điều đó sẽ tốt hơn nếu cả hai bạn cùng đi cắt tóc..."

Mike Wooldridge

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018

# săn hươu

- Hai giải pháp cân bằng Nash  $(C, C)$  và  $(\emptyset, \emptyset)$ .
  - Nếu bạn biết tôi sẽ hợp tác, điều tốt nhất bạn có thể làm là hợp tác.
  - Nếu bạn biết tôi sẽ đào ngũ, thì đó là điều tốt nhất bạn có thể làm.
- Phúc lợi xã hội được tối đa hóa bằng  $(C, C)$ .
- Kết quả hiệu quả Pareto duy nhất là  $(C, C)$ .
- Như thường lệ với trạng thái cân bằng Nash, lý thuyết không giúp chúng ta thực sự trong việc quyết định bên kia sẽ làm gì.
  - Do đó lo lắng về việc cắt tóc.
- Kịch bản tương tự cũng xảy ra trong các cuộc binh biến và đình công.
  - Tất cả chúng ta sẽ tốt hơn nếu vị thuyền trưởng đáng ghét của chúng ta bị phế truất, nhưng nếu một số người trong chúng ta nhượng bộ, tất cả chúng ta sẽ bị treo cổ.

Ma trận tiền thưởng săn hươu

		Tôi	
		khuyết điểm	chuông
j	khuyết điểm	2 2	3 1
	chuông	3 1	4 4

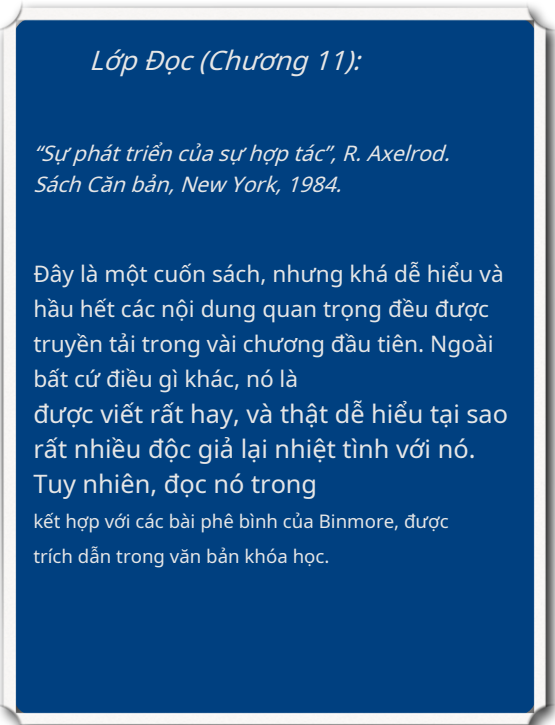
Sự khác biệt so với thể tiến thoái lưỡng nan của tù nhân là bây giờ sẽ tốt hơn nếu cả hai bạn hợp tác hơn là nếu bạn đào tẩu trong khi người kia hợp tác.

Bản quyền: MJ Wooldridge, S. Parsons và TR Payne, Mùa xuân 2013. Cập nhật 2018



# Bản tóm tắt

- Chương này đã xem xét các tương tác của tác nhân và một cách tiếp cận để mô tả chúng.
- Cách tiếp cận mà chúng tôi đã xem xét ở đây là *trò chơi lý thuyết*, một công cụ mạnh mẽ để phân tích các tương tác.
- Chúng tôi đã xem xét các khái niệm giải pháp về cân bằng Nash và tối ưu Pareto.
- Sau đó, chúng tôi xem xét Thế tiến thoái lưỡng nan của người tù cổ điển và cách trò chơi có thể được phân tích bằng lý thuyết trò chơi.
- Chúng tôi cũng đã xem xét Thế tiến thoái lưỡng nan của Người tù lặp đi lặp lại và các trò chơi 2 × 2 kinh điển khác.



Bản quyền: MJ Wooldridge, S. Parsons và TR Payme, Mùa xuân 2013. Cập nhật 2018