

## GIỚI THIỆU LÝ THUYẾT TRÒ CHƠI VÀ MỘT SỐ ỨNG DỤNG TRONG KINH TẾ HỌC VI MÔ

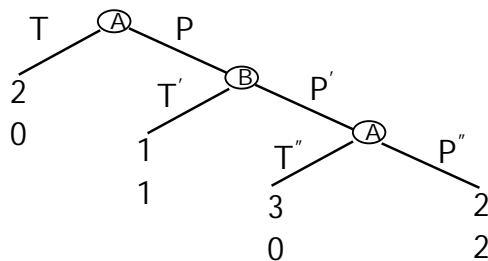
### Phần 2: Trò chơi động với thông tin đầy đủ

Trò chơi động (dynamic game) diễn ra trong nhiều giai đoạn, và một số người chơi sẽ phải hành động ở mỗi một giai đoạn. Trò chơi động khác với trò chơi tĩnh ở một số khía cạnh quan trọng. Thứ nhất, trong trò chơi động, thông tin mà mỗi người chơi có về những người chơi khác rất quan trọng. Nhờ ở Phần 1 đã phân biệt, một người có thông tin đầy đủ (complete information) khi người ấy biết hàm thỏa dụng (key - payoff) của những người chơi khác. Còn một người có thông tin hoàn hảo (perfect information) nếu họ tại mỗi bước phải ra quyết định (hành động), người ấy biết được toàn bộ lịch sử của các bước đi trước đó của trò chơi. Thứ hai, khác với các trò chơi tĩnh, trong trò chơi động một số năng lực (credibility) của những lời hứa (promises) hay đe dọa (threats) là yếu tố then chốt. Và cuối cùng, để tìm kiếm cân bằng cho các trò chơi, chúng ta phải vận dụng phương pháp quy nạp ngược (backward induction).

#### Trò chơi động với thông tin đầy đủ và hoàn hảo

*Ví dụ 1: Một trò chơi tổng cộng*

Trong tổng cộng một trò chơi động với thông tin đầy đủ và hoàn hảo và có cấu trúc nhớ



hình vẽ. Tại mỗi nút hoặc A hoặc B phải ra quyết định. Không gian hành động của họ chứa gồm hai khả năng: hoặc chọn trái (T), hoặc chọn phải (P). Những con số ở ngoài của các nhánh trong cây quyết định thể hiện kết quả thu được của hai người chơi, trong đó số đầu tiên là kết quả của A.

Để tìm kiếm cân bằng của trò chơi này, chúng ta không thể bắt đầu từ giai đoạn nào tiên, mà ngược lại, chúng ta sẽ dùng phương pháp quy nạp ngược, tức là bắt đầu từ giai đoạn cuối cùng của trò chơi.

Lưu ý là phương án tối ưu cho người chơi thứ nhất là kết quả T'', dù người A nhận 3 và B không được gì. Còn phương án tối ưu cho B là kết quả P'', trong đó B nhận 2 và A không được gì. Không cái nào xảy ra nếu không xảy ra. Tại sao vậy?

Nếu trò chơi kết thúc ở giai đoạn 3 thì A chắc chắn sẽ chọn T'' (vì  $3 > 2$ ). Còn nếu B nhận được ra quyết định ở giai đoạn 2 và biết điều này chắc chắn sẽ không chọn P' mà chọn

T' (vì  $1 > 0$ ). Vào giai đoạn 1, A dời nước trước nước những hành động kế tiếp của cái hai người nên chắc chắn chọn T (vì  $2 > 1$ ).<sup>1</sup>

Bây giờ chúng ta quay lại thảo luận về niềm tin của lời hứa hay đe dọa. Giả sử trước khi bắt đầu chơi, A hẹn hò với B nhờ sau. Trong lần chơi đầu tiên anh nên chọn P. Nếu thế khi nên lo lắng thì tôi sẽ chọn P', và rồi trong giai đoạn cuối cùng anh sẽ chọn P" nếu mọi chúng ta cùng nước 2. Liệu A có nên tin vào lời hẹn hò (hứa hẹn) bằng miệng này của B hay không?<sup>2</sup> Nếu này là trò chơi xảy ra một lần và mức ích của mỗi người chơi nên tuân theo luật của ích của mình thì câu trả lời hiển nhiên là không. Lý do là trong giai đoạn 2, B biết chắc là nếu A nói "yêu cầu chọn T" thì anh ta sẽ không nước gì, còn A sẽ nước 3 (là kết cục tồi nhất của A). Lòng trước này, B chắc nói A chọn P là sẽ chọn T' nếu nước 1. Những trước tình huống này, với những thông tin cho trước và nếu A là người duy lý thì chắc chắn A sẽ không dại gì nghe theo lời hứa hẹn ngon ngọt của B. Kết quả là A sẽ chọn T trong giai đoạn đầu tiên nhờ chúng ta đã phân tích ở trên. Nói một cách ngắn gọn, những hứa hẹn và đe dọa trong tương lai mà không mang tin cậy sẽ không hề có tác dụng gì, dù là nhỏ nhất, tới ông xã của những người chơi trong giai đoạn hiện tại. Trong một phần khác, chúng ta sẽ nghiên cứu tình huống trong đó lời hứa/ đe dọa mang tin cậy và do đó có ảnh hưởng nên hành vi của những người chơi ngay trong giai đoạn hiện tại.

### Ví dụ 2: Mô hình nước quyết song phòng Stackelberg (1934)

Nhìn lại trình tự thời gian của trò chơi này nhờ sau:

- 1) Hãng 1 chọn sản lượng  $q_1 \geq 0$
- 2) Hãng 2 quan sát  $q_1$  rồi sau đó chọn sản lượng  $q_2 \geq 0$
- 3) Hai hãng sản xuất với sản lượng  $q_1, q_2$  và lợi nhuận tổng cộng là  $\pi_1$  và  $\pi_2$

$$\pi_1(q_1, q_2) = q_1[P(Q) - c] ; Q = q_1 + q_2$$

$$\pi_2(q_1, q_2) = q_2[P(Q) - c] ; P(Q) = a - Q = a - (q_1 + q_2)$$

trong đó hàng số là chi phí biến, nước thời là chi phí trung bình của cái 2 hãng.

Nếu tìm kiếm cân bằng của trò chơi này, chúng ta lại áp dụng phương pháp quy nạp ngược bằng cách bắt đầu với hãng thời 2. Đầu tiên chúng ta phải tìm hàm phản ứng tốt nhất của hãng 2 nói với quyết định sản lượng  $q_1^*$  của hãng thời nhất trong giai đoạn 1 :

$$\begin{aligned} \text{Max } \pi_2(q_1, q_2) &= q_2[a - c - q_1^* - q_2] \Rightarrow q_2 = (a - c - q_1^*)/2 \\ q_2 &\geq 0 \end{aligned}$$

<sup>1</sup> Nếu dùng phương pháp quy nạp ngược sẽ dùng ở đây một cách dễ dàng là nhờ cấu trúc thông tin này phù hợp với hai của bài toán (tổng tổng) này. Trong các bài toán thức tế cấu trúc thông tin thông phức tạp hơn nhiều.

<sup>2</sup> Vì là hợp đồng miệng nên nó không thể bị che giấu nhờ trong tài.

Lưu ý rằng về mặt hình thức thì hàm phản ứng  $q_2(q_1^*)$  ở đây giống như trong mô hình Cournot. Tuy nhiên, có một điểm khác biệt quan trọng là trong mô hình Cournot,  $q_1^*$  là một giá trị giải trình, còn trong mô hình này, khi ra quyết định  $q_2$  hãng 2 đã quan sát được và biết giá trị của  $q_1^*$ .

Vì vậy là bài toán với thông tin này không hoàn hảo nên hãng 2 không thể nhất định vào vị trí của hãng 1 và do vậy biết rằng nếu mình quyết định sản lượng là  $q_1^*$  thì hãng 2 sẽ sản xuất  $q_2 = (a - c - q_1^*)/2$ . Vì vậy, trong giai đoạn 1, hãng 1 sẽ chọn  $q_1$  sao cho

$$\begin{aligned} \text{Max } \pi_1(q_1, q_2(q_1)) &= q_1[a - c - q_1 - q_2(q_1)] = q_1 \frac{a - c - q_1}{2} & \Rightarrow q_1^* &= \frac{a - c}{2} \\ & & \Rightarrow q_2^* &= \frac{a - c}{4} \end{aligned}$$

Lợi nhuận tổng cộng là:

$$\begin{aligned} \pi_{S1}^* &= \frac{(a - c)^2}{8} > \pi_{c1}^* &= \frac{(a - c)^2}{9} \\ \pi_{S2}^* &= \frac{(a - c)^2}{16} > \pi_{c2}^* &= \frac{(a - c)^2}{9} \end{aligned}$$

Câu hỏi đặt ra là tại sao hãng 1 có thể đạt được một sản lượng và lợi nhuận tổng cộng với một sản lượng và lợi nhuận nhỏ hơn trong khi hãng 2 thậm chí còn không đạt được một lợi nhuận trong một quyết định song phương Cournot? Câu trả lời không thuần túy chỉ nằm ở trình tự thời gian mà quan trọng hơn là do thông tin. Trong ví dụ này, cái hai hãng nếu biết nhiều thông tin hơn so với trường hợp một quyết định song phương Cournot: Hãng 2 có thể quan sát quyết định về sản lượng của hãng 1, còn hãng 1 biết là hãng 2 biết sản lượng của mình. Tuy nhiên hãng 1 có thể sử dụng thông tin bổ sung này để làm lợi cho mình trong khi hãng 2 khi có thêm thông tin lại bị thiệt hại. Hay nói một cách chính xác hơn, việc hãng 2 làm cho hãng 1 biết là hãng 2 biết sản lượng của hãng 1 làm cho hãng 2 bị thiệt. Nếu thấy như vậy, giải thích bằng một cách nào đó hãng 2 gây nhiều thông tin làm cho hãng 1 không biết được là liệu hãng 2 có biết sản lượng của mình hay không. Khi ấy, bài toán trở thành tổng tối ưu với trường hợp một quyết định Cournot trong đó 2 bên quyết định sản lượng mà không hề biết sản lượng thực tế của bên kia (thông tin không hoàn hảo)

**Ví dụ 3:** *Ma trận đàm phán* (Rubinstein sequential bargaining) – xem bài học thêm.

**Trò chơi năng với thông tin này như không hoàn hảo** (xem bài học thêm)

**Trò chơi lặp lại** (repeated games)

Mức kích của tiêu mức này là xem xét liệu các đề nghị hay hứa hẹn tổng lại năng tin cậy ảnh hưởng thế nào tới hành vi hiện tại của những người chơi.

**Ví dụ 1: Thế lưỡng nan trong trò chơi lặp hai giai đoạn**

Quay lại bài toán lưỡng nan của người tuồng trình bày dưới dạng chuẩn tắc nhỏ trong bảng bên.

Cân bằng Nash duy nhất là (không hợp tác, không hợp tác) và kết cục là (1, 1). Bây giờ giải số trò chơi này (gọi là trò chơi giai đoạn – stage game) một lần lại lần nữa, bằng kết quả một trình bày trong bảng dưới đây.

		<b>Người 1</b>	
		<i>Không hợp tác</i>	<i>Hợp tác</i>
<b>Người 2</b>	<i>Không hợp tác</i>	1, 1	5, 0
	<i>Hợp tác</i>	0, 5	4, 4

Cân bằng Nash duy nhất vẫn là (không hợp tác, không hợp tác) và kết cục hợp tác vẫn không đạt một lần nữa một niềm cân bằng

		<b>Người 1</b>	
		<i>Không hợp tác</i>	<i>Hợp tác</i>
<b>Người 2</b>	<i>Không hợp tác</i>	2, 2	6, 1
	<i>Hợp tác</i>	1, 6	5, 5

**Nhận xét:**

- Nếu trò chơi giai đoạn (stage game) chỉ có một cân bằng Nash duy nhất thì nếu trò chơi này một lần lại nhiều lần thì cũng sẽ chỉ có một cân bằng Nash duy nhất, nó là số lần lặp lại cân bằng Nash của trò chơi giai đoạn.
- Rõ ràng là nếu trò chơi này một lần lại nhiều lần thì thiết hai tờ viết không hợp tác sẽ rất lớn. Câu hỏi đặt ra là liệu có cách nào để thiết lập số hợp tác hay không? Ở đây chúng ta tạm thời không quan tâm tới khía cạnh này nữa và lòng tin của mỗi người chơi mà chỉ xem xét thuận tiện về mặt kinh tế của họ.

**Ví dụ 2: Thế lưỡng nan trong trò chơi lặp vĩnh viễn**

Bây giờ giải số trò chơi một lần lại một cách vĩnh viễn. Chúng ta sẽ xem xét khả năng một đề dọa hay hứa hẹn thông lại năng tin cậy ảnh hưởng thế nào tới hành vi hiện tại của những người chơi?

Nhìn lại cùng thời tính hiện giá của thu nhập, trong đó mỗi người nhận một  $\pi_1$  trong giai đoạn 1,  $\pi_2$  trong giai đoạn 2 v.v. Tổng thu nhập của người nói tính theo giá hiện tại là  $\sum PV = \pi_1 + \delta\pi_2 + \delta^2\pi_3 + \dots$ ; trong đó  $\delta$  là nhân tố chiết khấu (discount factor).

Bây giờ chúng ta sẽ chứng minh rằng ngay cả khi trò chơi giai đoạn chỉ có một cân bằng Nash duy nhất thì vẫn có cách để buộc những người chơi duy lý hợp tác với nhau, với niềm tin  $\delta$  đủ lớn. Cách thức để thiết lập số hợp tác này là thực hiện chiến lược “trừng phạt” (trigger strategy) mà thực chất là một lời đề dọa trả thù năng tin cậy nói với những hành vi vi phạm hợp đồng. Chiến lược trừng phạt này một thực hiện như sau:

- Trong giai đoạn 1, chọn "hợp tác"
- Trong giai đoạn t, tiếp tục chọn "hợp tác" chừng nào trong (t-1) giai đoạn trước người kia cũng chọn "hợp tác"
- Chuyển sang chơi "không hợp tác" nếu trong giai đoạn (t-1), người kia phản bội hợp đồng chơi "hợp tác"

Giả sử trong suốt (t-1) giai đoạn đầu tiên, cả hai người chơi đều tuân thủ thỏa ước và chọn "hợp tác". Nhưng tại giai đoạn thứ t, một người toan tính việc vi phạm thỏa ước vì thấy cái lợi trước mắt. Khi ấy, người này phải so sánh 2 giá trị thu nhập kỳ vọng của hợp tác và không hợp tác.

Nếu trong giai đoạn t người ấy không hợp tác thì người ấy nhận 5, và tới (t+1) trở lại người kia sẽ chọn không hợp tác nên trong phần người này, và khi ấy phần ứng trả nhất tổng ứng của người này cũng sẽ là không hợp tác. Nhờ vậy, tổng giá trị kỳ vọng thu nhập của người ấy theo hiện giá là

$$PV_{\bar{c}} = \delta^{t-1} \cdot 5 + \delta^t \cdot 1 + \delta^{t+1} \cdot 1 + \dots$$

$$PV_{\bar{c}} = \delta^{t-1} \left[ 5 + \frac{\delta}{1-\delta} \right] \quad (1)$$

Khả năng thời 2 là người ấy tiếp tục chọn hợp tác. Khi ấy, tổng thu nhập của anh ta theo hiện giá là

$$PV_C = \delta^{t-1} \cdot 4 + \delta^t \cdot 4 + \delta^{t+1} \cdot 4 + \dots$$

$$PV_C = \delta^{t-1} + \frac{4}{1-\delta} \quad (2)$$

So sánh (1) và (2) ta thấy  $PV_C \geq PV_{\bar{c}} \Leftrightarrow \frac{4}{1-\delta} \geq 5 + \frac{\delta}{1-\delta}$

$$\Leftrightarrow 4 \geq 5(1-\delta) + \delta = 5 - 4\delta$$

$$\Leftrightarrow \delta \geq 1/4$$

Nhờ vậy, nếu  $\delta \geq 1/4$  thì chiến lược trông chờ là một cân bằng Nash. Nói cách khác, với  $\delta$  đủ lớn (tức là những người chơi biết khâu tổng lại nhiều lần) thì khi theo đuổi mục tiêu và kẻ lười nhả hòa lợi ích của mình thì tất cả người chơi đều có những cơ hội trong thỏa ước hợp tác.

### Ví dụ 3: Trò chơi với nước quyền song phòng Cournot

Chúng ta hãy biết rằng trong trường hợp nước quyền song phòng Cournot:

$q_{c1}^* = q_{c2}^* = (a-c)/3$  và do vậy  $Q_C^* = 2(a-c)/3 > Q_m^* = (a-c)/2$  (= một tổng cao khi hai doanh nghiệp cùng kết luận nước quyền). Nhờ vậy, hai hãng này có thể

áp dụng chiến lược trừng phạt nếu hai nước sẽ hợp tác trong sản xuất. Nếu kiểm tra lại một số nhiều các nội dung trình bày ở ví dụ 2, chúng ta có thể làm một bài tập như sau. Giả sử trò chơi Cournot này được lặp lại mãi mãi, hãy tìm giá trị tối thiểu của  $\delta$  để giải pháp hợp tác là một cân bằng Nash (SPNE)?

Chiến lược trừng phạt như sau:

- Bất cứ ai chơi bằng việc chọn một sản lượng  $Q_{m/2}^*$  ( $= (a-c)/4$ ) trong giai đoạn 1
- Nếu trong (t-1) giai đoạn nào tiên, bên kia chọn  $Q_{m/2}^*$  thì tiếp tục chọn  $Q_{m/2}^*$ . Bằng không thì chuyển sang  $Q_{c/2}^*$  ( $= (a-c)/3$ ) mãi mãi.

Giả sử ở giai đoạn t, hãng 1 toàn tính chuyển phải với thỏa ước ban đầu. Hãng này biết là hãng 2 sẽ chuyển sang chọn  $q_2^* = q_{c2}^*$  kể từ giai đoạn thời (t+1). Vì vậy, hãng 1 đang trước hai lựa chọn:

- Phải với thỏa ước:

$$\begin{aligned} \pi^{\bar{C}} &= \delta^{t-1} \pi_d + \delta^t \pi_C + \delta^{t+1} \pi_C + \dots \\ &= \delta^{t-1} (\pi_d + \delta \pi_C + \delta^2 \pi_C + \dots) \end{aligned}$$

$$\pi^{\bar{C}} = \delta^{t-1} \left( \pi_d + \frac{\delta}{1-\delta} \pi_C \right)$$

Nếu hãng 2 tiếp tục chọn hợp tác trong giai đoạn t, tức là tiếp tục chọn  $q_2^* = Q_{m/2}^* = (a-c)/4$  thì  $q_{d1}^*$  sẽ  $\max q_{d1} [a - c - q_{d1} - (a-c)/4] \Rightarrow q_{d1}^* = 3(a-c)/8 \Rightarrow \pi_d = 9(a-c)^2/64$

- Phải với thỏa ước:

$$\pi^C = \delta^{t-1} \pi_m + \delta^t \pi_m + \delta^{t+1} \pi_m + \dots$$

$$\pi^C = \delta^{t-1} \frac{\pi_m}{1-\delta}$$

So sánh  $\pi^C \geq \pi^{\bar{C}}$  :

$$\begin{aligned} \Leftrightarrow \frac{\pi_m}{1-\delta} &\geq \pi_d + \frac{\delta}{1-\delta} \pi_C \\ \Leftrightarrow \frac{(a-c)^2}{8(1-\delta)} &\geq \frac{9(a-c)^2}{64} + \frac{\delta}{1-\delta} \frac{(a-c)^2}{9} \\ \Leftrightarrow \frac{1}{8} &\geq \frac{9(1-\delta)}{64} + \frac{\delta}{9} \\ \Leftrightarrow 72 &\geq 81(1-\delta) + 64\delta = 81 - 17\delta \\ \Leftrightarrow \delta &\geq \frac{9}{17} \end{aligned}$$

Một lần nữa chúng ta lại thấy rằng nếu  $\delta$  đủ lớn (tức là những người chơi chơi nhiều lần) thì khi theo đuổi mục tiêu vì kẻ lợi ích của mình thì hai công ty cũng có những cơ hội trong thỏa ước hợp tác.

---

*Tài liệu tham khảo*

Robert Gibbons, "Game Theory for Applied Economists", Princeton University Press, 1992