



The Strategy of Conflict

# CHIẾN LƯỢC XUNG ĐỘT

Với lời nói đầu mới của tác giả

# THOMAS C. SCHELLING



NHÀ XUẤT BẢN TRẺ

# CHIẾN LƯỢC XUNG ĐỘT

(Tái bản lần thứ nhất)

335.1      19-56      /TRE  
TRE - 2007

The Strategy of Conflict

# CHIẾN LƯỢC XUNG ĐỘT

Tác giả:

THOMAS C. SCHELLING

Người dịch:  
VŨ TIẾN PHÚC  
NGUYỄN VĂN THẮNG

Hiệu đính:  
TS kinh tế NGUYỄN VINH DỰ  
NGUYỄN QUANG THẮNG  
(Đại học Tổng hợp Texas - Austin, Hoa Kỳ)

TÀI LIỆU THAM KHẢO

NHÀ XUẤT BẢN TRẺ  
THỜI BÁO KINH TẾ SÀI GÒN  
TRUNG TÂM KINH TẾ CHÂU Á-THÁI BÌNH DƯƠNG

© Copyright 1960 by The President and Fellows of Harvard College

Nhà xuất bản Trẻ độc quyền xuất bản bản tiếng Việt theo Hợp đồng nhượng quyền với  
Harvard University Press tháng 3.2006

The Translation and Publication of this book in Vietnamese language were made possible  
through the financial support of the Public Affairs Section, Consulate General of the  
United States of America in Ho Chi Minh City.

Sách được dịch và xuất bản bằng tiếng Việt với sự hỗ trợ về tài chính của Phòng Thông  
tin - Văn hóa, Tổng Lãnh sự quán Hoa Kỳ tại Thành phố Hồ Chí Minh.

## LỜI GIỚI THIỆU

TS. VŨ THÀNH TỰ ANH

*Chương trình Giảng dạy Kinh tế Fulbright*

Xung đột và hợp tác là một phần bản chất của cuộc sống loài người. Thuở ban sơ, xung đột và hợp tác mang tính tự nhiên và cá nhân, được thúc đẩy bởi bản năng sinh tồn. Khi con người biết quần cư thành bầy đàn, và càng về sau này, khi phương thức tổ chức cộng đồng và quốc gia – dân tộc ngày càng phát triển thì các hình thái xung đột càng trở nên phức tạp, mang tính duy lý và đấu trí nhiều hơn. Nhìn ra thế giới ngày nay thì có thể thấy hợp tác và xung đột hiện diện trong mọi lĩnh vực của đời sống, từ kinh tế, văn hóa, chính trị đến ngoại giao và an ninh v.v...

Các nhà khoa học xã hội từ lâu đã cố gắng hiểu và lý giải những nguyên nhân cơ bản của hợp tác và xung đột. Tuy nhiên, chỉ đến khi lý thuyết trò chơi ra đời – với tác phẩm nền tảng và kinh điển của John von Neumann and Oscar Morgenstern nhan đề *Lý thuyết trò chơi và hành vi kinh tế* xuất bản năm 1944 – và phát triển sau đó<sup>1</sup> thì các nhà nghiên cứu mới có trong tay những công cụ sắc bén

---

<sup>1</sup> Với sự đóng góp to lớn của John Nash, Reinhard Selten, John Harsanyi, Robert Aumann, và bản thân Thomas Schelling cùng nhiều học giả khác.

và chặt chẽ để phân tích các hiện tượng xung đột và hợp tác. Đóng góp của Schelling là ông đã góp phần phát triển lý thuyết trò chơi bất hợp tác (non-cooperative game theory) và đi tiên phong trong việc vận dụng lý thuyết này để trả lời những câu hỏi quan trọng của các ngành khoa học xã hội. Trong những thành công của Schelling, quan trọng nhất có lẽ là việc ông chỉ ra một cách thuyết phục rằng các tương tác xã hội có thể được nhìn nhận như là những trò chơi bất hợp tác ẩn chứa cả sự đồng thuận lẫn xung đột về lợi ích.

Cụ thể hơn, một câu hỏi xuyên suốt các công trình nghiên cứu và hoạt động cố vấn của Schelling là: *"Tại sao một số nhóm người, tổ chức, hay quốc gia có thể làm này nở và duy trì sự hợp tác, trong khi ở những nơi khác lại tồn tại xung đột giữa các bên?"* Những công trình của Schelling (và của các nhà nghiên cứu sau này) đã cho thấy lý thuyết trò chơi – hay lý thuyết ra quyết định trong các điều kiện tương tác chiến lược – là một cách tiếp cận hiệu quả (và hiện đang chiếm ưu thế) trong việc trả lời cho câu hỏi muôn thuở này. Bài diễn văn của Viện khoa học Hoàng gia Thụy Điển nhiệt thành tuyên dương ông và Robert Aumann (nhà toán học gốc Do Thái) đã có những đóng góp to lớn trong việc "*nâng cao hiểu biết của chúng ta về các hành vi xung đột và hợp tác thông qua việc mở rộng và áp dụng lý thuyết trò chơi – một phương pháp được sử dụng để phân tích các tương tác chiến lược giữa các tác nhân với nhau. Những công trình của họ đã vượt xa khỏi ranh giới của kinh tế học và góp phần chuyển hóa các môn khoa học xã hội khác. Các nghiên cứu của Aumann và Schelling*

*tiếp tục định hướng cho các thảo luận về sự hình thành các thể chế xã hội."*

Cuốn sách *Chiến lược xung đột* (The Strategy of Conflict) mà bạn đang cầm trên tay là sự tổng kết quá trình nhiều năm nghiên cứu và trải nghiệm của Schelling. Không phải ngẫu nhiên mà cuốn sách được hình thành và ra đời trong nỗi ám ảnh của chiến tranh lạnh, trong bối cảnh cuộc chạy đua hạt nhân giữa hai cường quốc thế giới lúc bấy giờ là Liên Xô và Mỹ ngày càng trở nên gay cấn. Schelling đã rất ngạc nhiên (và có phần thích thú nữa) khi phát hiện ra rằng ưu thế thương lượng (hay mặc cả) của mỗi bên chịu ảnh hưởng bởi nhiều yếu tố, đặc biệt là các lựa chọn khả hữu cũng như khả năng chủ động thay đổi những lựa chọn này của bản thân họ và của đối thủ. Bằng các công cụ toán học, ông đã chứng minh được điều tưởng chừng là nghịch lý rằng, trước một tình huống xung đột với đối thủ, một bên có thể chiếm ưu thế bằng cách "phé bỏ" một số lựa chọn đang có của mình. Kết luận này làm chúng ta nhớ lại câu "đặt vào chỗ chết thì sống, đặt vào chỗ mất thì còn" trong binh pháp Tôn Tử hay tích "qua sông đốt thuyền" của Hàn Tín. Trong thời đại ngày nay, những chiến lược này vẫn tiếp tục được vận dụng một cách linh hoạt. Bill Gates nổi tiếng của nước Mỹ cũng thường nói: "Chúng ta đặt cược cả công ty vào Internet!" Còn ở bán cầu bên kia của trái đất, Triệu Tử Dương đã kiên quyết "tự buộc" chính phủ Trung Quốc vào những cam kết nghiêm ngặt với cộng đồng thế giới khi gia nhập WTO để đẩy mạnh cải cách trong nước vì ông hiểu rõ rằng, khó khăn lớn nhất của Trung Quốc không phải là trên bàn đàm phán quốc

tế mà là ở các thế lực chống đối của các nhóm lợi ích trong nước. Thế mới thấy trực giác của người xưa và trí tuệ thời này có một sự tương thông tuyệt diệu, chỉ khác là ngày nay các nhà khoa học sử dụng những công cụ phân tích thông qua các mô hình toán ngày càng tinh vi và chặt chẽ hơn.

Trong cuốn sách này, Schelling cũng chỉ ra rằng trong các tình huống có xung đột, chính khả năng “trả đũa” (chứ không phải là khả năng kháng cự hay tấn công) mới là điều quan trọng. Hơn thế, ông còn cho rằng trả đũa “xuất kỳ bất ý” sẽ làm cho đối thủ “ngán” hơn nhiều, và vì vậy giúp tăng cường độ tin cậy và hiệu quả cho các biện pháp trả đũa.

Không chỉ quan tâm tới xung đột, Schelling còn nghiên cứu các quá trình thiết lập môi trường cho sự tin cậy và cam kết chiến lược, nhờ đó sự hợp tác trong dài hạn có thể được duy trì. Trong quá trình hợp tác này, các bên cũng sẽ nhận thấy về lâu về dài họ cùng có thể được lợi nếu như ban đầu họ chịu hy sinh một phần quyền lợi của mình. Những phân tích này của Schelling đã giúp giải thích một phổ rộng lớn những hiện tượng thường gặp, từ chiến lược cạnh tranh của công ty cho tới việc ủy thác quyết định về chính trị hay ngoại giao.

Schelling cũng quan tâm tới khả năng hợp tác của các bên trong hoàn cảnh không tồn tại xung đột đáng kể về lợi ích giữa các bên, nhưng tất cả các bên sẽ cùng phải trả giá khá đắt khi sự hợp tác bị phá vỡ. Trong nghiên cứu của mình, đặc biệt là trong những thí nghiệm ở trên lớp

với sinh viên, Schelling phát hiện ra rằng các kết cục hợp tác thường xảy ra nhiều hơn so với dự đoán của lý thuyết. Năng lực hợp tác xem ra phụ thuộc vào khung quy chiếu chung của các bên. Các quy ước và chuẩn mực xã hội (tục lệ, tập quán, giá trị, quan niệm về đạo đức v.v...) là một bộ phận hữu cơ của khung quy chiếu chung này. Công trình của Schelling trong lĩnh vực này là nguồn cảm hứng cho nhà triết học David Lewis (mặc dù không phải là người đầu tiên) đi đến ý tưởng cho rằng ngôn ngữ được hình thành như là phương tiện để phối hợp các hoạt động có tính tập thể của cộng đồng.

Rõ ràng là sẽ không thực tế nếu cố gắng trình bày, dù chỉ một cách vấn tắt, những vấn đề Schelling khảo cứu trong cuốn *Chiến lược xung đột* bằng một bài giới thiệu nhỏ này. Để kết thúc bài giới thiệu, chúng tôi xin gợi ý một tinh thần – và do vậy – một cách đọc cuốn sách của Schelling mà bản thân chúng tôi cho là hữu ích. Tôn Tử – nhà cầm quân và binh pháp huyền thoại của Trung Hoa đã từng nói “trận địa phi tướng chi địa”, nghĩa là “đã là nơi đánh nhau thì không có gì tốt đẹp cả”. Thế nhưng chính Tôn Tử cũng lại là tác giả của cuốn *Binh pháp* nổi tiếng mà có lẽ đến bản thân ông cũng không ngờ là hơn 2.000 năm sau bản dịch cuốn sách của ông lại là một trong những đầu sách bán chạy nhất ở thế giới phương Tây. Điều đáng nói ở đây là ở chỗ, có lẽ chính vì nỗi chán ghét chiến tranh mà Tôn Tử đã nghiên cứu chiến tranh và nghệ thuật quân sự, đơn giản chỉ vì theo ông “chiến tranh là đại sự của quốc gia, quan hệ tới việc sống chết của nhân dân, sự mất mát của nhà nước, không thể không khảo sát nghiên

cứu cho thật kỹ." Chán ghét chiến tranh nhưng vẫn phải nghiên cứu, nghiên cứu một cách tường tận để nhờ đó mà tránh được những hậu quả tàn khốc của nó. Quay lại với những công trình của Schelling, có lẽ không phải ngẫu nhiên mà Viện khoa học Hoàng gia Thụy Điển đã nhận định rằng, những đóng góp khoa học của Schelling đã góp một phần quan trọng trong việc giải quyết xung đột và nỗ lực để tránh chiến tranh. Hiểu rõ xung đột là cơ sở để nảy nở sự hợp tác. Mong rằng các bạn hãy đọc cuốn sách này của Schelling với tinh thần như vậy.

*Thành phố Hồ Chí Minh,  
11.9.2007*

## LỜI NÓI ĐẦU CHO LẦN XUẤT BẢN NĂM 1980

Khi tôi biết Nhà xuất bản Harvard University Press chuẩn bị cho in bản bìa mềm của cuốn sách này, tôi băn khoăn không biết chỗ nào trong đó đã trở nên quá lạc hậu mà tôi cần lược bỏ hoặc viết lại, ít ra cũng viết lời xin lỗi trong phần lời nói đầu mới. Cuốn sách *Chiến lược xung đột* ra đời đã được 20 năm. Tôi không thường xuyên đọc lại cuốn sách này. Có những phần trong cuốn sách tôi đã không ngó ngàng tới trong hơn một thập kỷ. Có thể có những điều tôi nói đã trở nên nhảm chán, lạc lõng, hoặc không đúng nữa.

Quả thật là có một vài điều như vậy. Nhưng về tổng thể tôi có thể vui mừng nói rằng, mặc dù thỉnh thoảng gặp phải một vài ví dụ đã lạc hậu, hầu hết nội dung cuốn sách vẫn còn nguyên giá trị. Những bình luận trong Chương 1 về sự kém phát triển của chiến lược quân sự trong các trường đại học và các lực lượng quân đội giờ đây đã xa thực tế đến mức chỉ còn lại giá trị lịch sử. Một vấn đề lớn hơn là liệu các sinh viên – mà các sinh viên có thể là những người duy nhất ngày nay đọc cuốn sách này lần đầu tiên – có nhận ra được các tên tuổi như Kim Môn (Quemoy), Khrushchev, và Mossadeq, hay có biết Hoa hậu Rheingold đã từng được lựa chọn như thế nào hay không.

Tất cả chúng ta nên vui mừng vì Phụ lục A không bị lạc hậu. Phần này được viết dựa trên tiền đề rằng vũ khí nguyên tử không được sử dụng lại từ sau vụ Nagasaki. Mong rằng cuốn sách này được in lại nhiều lần với tiền đề đó không thay đổi.

Một số ý tưởng mà tôi cho là xuất phát từ Chương 10 đã trở nên thịnh hành. Một số sau đó lại trở nên lỗi thời. Ngày nay có rất nhiều sách báo nói về hạn chế vũ khí, bao gồm cả những điều tôi đã viết. Nhưng Chương 10 vẫn nói đầy đủ về Hiệp ước Hạn chế vũ khí chiến lược (Strategic Arms Limitation Treaty – SALT) một cách rõ ràng như bất cứ 25 trang sách nào khác. Bạn đọc nào muốn theo đuổi các suy nghĩ của tôi về chiến lược và kiểm soát vũ khí có thể đọc cuốn sách với tên gọi như thế mà tôi viết cùng với Morton H. Halperin, do Twentieth Century Fund xuất bản năm 1961, hoặc cuốn *Vũ khí và Ảnh hưởng* (*Arms and Influence*), Yale University Press, 1966.

Có thể nội dung lý thuyết, chứ không phải chính sách ngoại giao, là điều khiến hầu hết bạn đọc quan tâm đến cuốn sách này vào thời điểm hiện nay. Với việc tập hợp các bài tiểu luận trong cuốn sách, tôi hy vọng góp phần thiết lập một lĩnh vực liên ngành đã được diễn đạt theo các cách khác nhau như "lý thuyết mặc cả," "lý thuyết xung đột," hay "lý thuyết chiến lược." Tôi muốn chứng tỏ rằng một số lý thuyết cơ bản, vận dụng các khoa học chính trị, xã hội, và kinh tế, thậm chí luật và triết học, kể cả nhân chủng học, có thể rất hữu ích không chỉ cho các nhà lý thuyết thuần túy mà còn cho những người quan tâm đến các vấn đề thực tiễn. Tôi cũng mong như vậy, và bây giờ tôi nghĩ, có thể không đúng, rằng lý thuyết trò chơi nên

được định hướng lại theo chiều ứng dụng trong nhiều lĩnh vực này. Trừ các trường hợp ngoại lệ như Howard Raiffa, Martin Shubik, và Nigel Howard, các nhà lý thuyết trò chơi có khuynh hướng dùng lại nhiều hơn trong ranh giới toán học. Lĩnh vực liên ngành mà tôi hy vọng sẽ được thiết lập vẫn tiếp tục phát triển nhưng không bùng nổ và chưa có tên riêng của mình.

Một vài tạp chí, đặc biệt là tờ *Journal of Conflict Resolution*, đã đóng vai trò quan trọng trong việc phát triển lĩnh vực này. Nhưng ngoại trừ một vài thuật ngữ như "trò chơi có tổng lợi ích khác không" (non-zero-sum game) và "phản thưởng" (payoff), ngay cả lý thuyết cơ bản nhất cũng rất ít được sử dụng trong các tạp chí chuyên ngành dành cho các nhà hoạch định chính sách và những người thực hành. (Chỉ một vài năm trước đây, trong bài viết về thái độ khác nhau của Liên Xô và Mỹ đối với một số vũ khí đặc biệt có thể đưa vào kiểm soát, tôi sử dụng một vài ma trận 2 x 2 để giúp người đọc nhìn thấy những sự khác biệt. Tổng biên tập của tạp chí nọ, mà tôi xin miễn nêu tên ra đây, khăng khăng đòi tôi xóa bỏ các ma trận đó để không làm cho độc giả chán chường. Ông cho rằng độc giả sẽ cảm thấy dễ chịu hơn với những mô tả bằng lời vì chúng ít tính chất tra tấn hơn, mặc dù họ sẽ khó hiểu đúng ý tôi hơn.)

Cuốn sách đã được đón nhận rộng rãi, và nhiều người khuyến khích tôi bằng cách nói rằng họ thích cuốn sách này hoặc học được cái gì đó từ nó. Nhưng sự phản hồi làm ấm lòng tôi nhất trong 20 năm qua là từ John Strachey, người đã qua đời. John Strachey, tác giả các cuốn sách mà tôi đã đọc khi còn học đại học, là một nhà kinh tế Mác-xít nổi tiếng vào những năm 1930. Sau chiến tranh

Ông làm Bộ trưởng Quốc phòng trong Chính phủ Công Đảng Anh. Một vài người trong chúng tôi tại Trung tâm các vấn đề quốc tế của Harvard có mời ông đến thăm vì ông đang viết một cuốn sách về giải trừ quân bị và kiểm soát vũ khí. Khi ông ghé thăm tôi, ông đã kêu lên cuốn sách này giúp ích cho suy nghĩ của ông như thế nào, và trong khi ông đang nói một cách say sưa, tôi cố đoán xem những ý tưởng mới mẻ nào của tôi ở chương nào đã có tác động đến ông như vậy. Hóa ra là chẳng phải vì một ý tưởng đặc biệt ở chương nào đó. Trước khi đọc cuốn sách này, ông không hiểu rằng một cuộc xung đột có tổng lợi ích khác không có thể tồn tại được. Ông hiểu xung đột có thể cùng tồn tại với lợi ích chung nhưng nghĩ, hay cho là hiển nhiên, rằng chúng tách biệt nhau về cơ bản, chứ không phải là những khía cạnh của một cấu trúc thống nhất. Khi đọc cuốn sách của tôi, một học giả, người chỉ quan tâm đến chủ nghĩa tư bản độc quyền và đấu tranh giai cấp, chiến lược nguyên tử và chính trị đồng minh, người dành phần cuối của sự nghiệp đời mình cho kiểm soát vũ khí và lập lại hòa bình, đã ngã gục trước một ý tưởng sơ đẳng đến nỗi thậm chí tôi cũng không biết nó không phải là điều hiển nhiên. Với sự khiêm tốn và tự trọng của mình ông đã thú nhận điều đó với tôi. Bạn sẽ không bao giờ biết trước được việc viết một cuốn sách có thể có ảnh hưởng như thế nào.

THOMAS C. SCHELLING  
*Cambridge, Massachusetts*

# MỤC LỤC

## PHẦN I

### CÁC YẾU TỐ CỦA MỘT LÝ THUYẾT VỀ CHIẾN LƯỢC

1. MÔN KHOA HỌC CHÂM PHÁT TRIỂN VỀ CHIẾN LƯỢC QUỐC TẾ	3
2. MỘT TIỂU LUẬN VỀ MẶC CẢ	30
3. MẶC CẢ, TRAO ĐỔI THÔNG TIN, VÀ CHIẾN TRANH HẠN CHẾ	78

## PHẦN II

### ĐỊNH HƯỚNG LẠI LÝ THUYẾT TRÒ CHƠI

4. TIỀN TỐI MỘT LÝ THUYẾT VỀ QUYẾT ĐỊNH PHÙ THUỘC LẮN NHAU	123
5. SỰ THỰC THI, LIÊN LẠC VÀ CÁC NƯỚC ĐI CHIẾN LƯỢC	177
6. LÝ THUYẾT TRÒ CHƠI VÀ NGHIÊN CỨU THỰC NGHIỆM	240

## PHẦN III

### CHIẾN LƯỢC CÓ THÀNH PHẦN NGẪU NHIÊN

7. NGẪU NHIÊN HÓA HỨA HẸN VÀ ĐE DỌA	259
8. ĐE DỌA CÓ YẾU TỐ NGẪU NHIÊN	278
9. LO SỢ LẮN NHAU VỀ TẤN CÔNG BẤT NGỜ	307
10. TẤN CÔNG BẤT NGỜ VÀ GIẢI TRỪ QUÂN BỊ	344

## PHẦN PHỤ LỤC

PHỤ LỤC A. VŨ KHÍ NGUYÊN TỬ VÀ CHIẾN TRANH HẠN CHẾ	385
PHỤ LỤC B. VỀ VIỆC TỬ BỎ TÍNH ĐỐI XỨNG TRONG LÝ THUYẾT TRÒ CHƠI	400
PHỤ LỤC C GIẢI THÍCH LẠI VỀ KHÁI NIÊM DÁP SỐ CHO CÁC TRÒ CHƠI "BẤT HỢP TÁC"	436

## **PHẦN I**

# **CÁC YẾU TỐ CỦA MỘT LÝ THUYẾT VỀ CHIẾN LƯỢC**

# 1

## MÔN KHOA HỌC CHẬM PHÁT TRIỂN VỀ CHIẾN LƯỢC QUỐC TẾ

Trong rất nhiều lý thuyết về xung đột - tương ứng với rất nhiều cách hiểu khác nhau về từ "xung đột" - có sự phân chia ranh giới chủ yếu giữa một bên là các lý thuyết coi xung đột là một trạng thái bệnh hoạn và cần tìm căn nguyên và cách chữa trị, và bên kia là các lý thuyết coi xung đột là điều tất nhiên và cần nghiên cứu các hành vi liên quan đến nó. Các lý thuyết kể sau có thể phân chia thành các loại nhỏ hơn: các lý thuyết xem xét các thành phần tham gia vào cuộc xung đột với tất cả sự phức tạp của nó - về cả hành vi "hợp lý" lẫn hành vi "bất hợp lý," có ý thức và không có ý thức, và động cơ cũng như sự tính toán - và các lý thuyết chỉ tập trung vào hành vi hợp lý, có ý thức, và khôn khéo. Nói một cách đơn giản, các lý thuyết kể sau coi xung đột như một cuộc đua, trong đó các bên tham gia đều tìm cách "chiến thắng." Việc nghiên cứu hành vi xung đột có ý thức, khôn ngoan, tinh vi - tức là hành vi của người chiến thắng - sẽ giống như việc tìm các quy tắc hành xử "đúng đắn" theo nghĩa giúp giành được thắng lợi trong thi đấu.

Chúng ta có thể gọi lĩnh vực nghiên cứu này là *chiến lược* xung đột.<sup>1</sup> Chúng ta cần quan tâm nghiên cứu nó ít nhất vì ba lý do. Bản thân chúng ta có thể liên quan đến một cuộc xung đột. Trên thực tế, tất cả chúng ta đều tham gia vào xung đột quốc tế, và chúng ta ai cũng muốn “chiến thắng” theo một nghĩa thích hợp nào đó. Chúng ta có thể muốn biết xem những người tham gia thực sự có hành vi như thế nào trong các tình huống xung đột. Hiểu rõ được cách chơi “đúng” có thể cho chúng ta một chuẩn mực để nghiên cứu các hành vi thực tế. Chúng ta có thể muốn kiểm soát hay tạo ảnh hưởng lên hành vi của người khác trong xung đột, và do đó, chúng ta muốn biết các biến số do chúng ta kiểm soát có thể có ảnh hưởng như thế nào đến hành vi của họ.

Nếu chúng ta chuyên nghiên cứu lý thuyết về chiến lược, chúng ta cần tuân thủ giới hạn khắt khe của giả định về hành vi hợp lý – không chỉ là hành vi khôn ngoan, mà là hành vi được thúc đẩy bởi sự tính toán lợi ích một cách có ý thức, sự tính toán dựa trên một hệ thống giá trị rõ ràng và nhất quán nội tại. Như vậy, chúng ta sẽ giới hạn khả năng ứng dụng của mọi kết quả mà chúng ta tìm được. Nếu mối quan tâm của chúng ta là nghiên cứu hành vi thực tế, các kết quả chúng ta tìm được trong giới hạn này có thể là một ước lượng gần đúng thực tế hoặc chỉ là một bức tranh biếm họa. Mọi trừu tượng hóa đều gặp rủi ro này, và chúng ta cần chuẩn bị để biết phán xét tất cả kết quả tìm được.

---

<sup>1</sup> Thuật ngữ “chiến lược” ở đây lấy từ lý thuyết trò chơi, nhằm phân biệt các trò chơi đòi hỏi kỹ năng, trò chơi may rủi, và trò chơi chiến lược – tức là trò chơi trong đó nước đi tốt nhất của mỗi người chơi phụ thuộc vào nước đi của những người chơi khác. Thuật ngữ này được dùng để nhấn mạnh sự phụ thuộc lẫn nhau trong quyết định của các đối thủ và sự dự đoán hành vi của nhau. Đây không phải nghĩa dùng trong quân sự.

Lợi ích của việc bồi bổ mảnh đất “chiến lược” cho sự phát triển lý thuyết, dưới mọi cách tiếp cận có thể, không phải ở chỗ rõ ràng là nó gần gũi với chân lý nhất, mà ở chỗ giả định về hành vi hợp lý là một giả định hữu ích. Nó giúp nắm chắc vấn đề và điều này đặc biệt có ích cho sự phát triển lý thuyết. Nó cho phép chúng ta hòa lẫn các quá trình phân tích của mình vào những tính toán của những người tham gia xung đột giả định, và bằng cách đòi hỏi tính nhất quán ở một mức độ nhất định trong hành vi của những người tham gia cuộc xung đột giả định này, chúng ta có thể xem xét các biểu hiện của hành vi dựa theo việc nó có đáp ứng được các tiêu chuẩn nhất quán đó hay không. “Hành vi hợp lý” là một tiền đề vững vàng để xây dựng lý thuyết. Tôi muốn nhắc lại rằng, lý thuyết thu được có đem lại hiểu biết tốt hay tồi về hành vi thực tế phụ thuộc vào sự phán xét sau đó.

Tuy nhiên, trong khi coi xung đột là điều tất nhiên, và coi những người tham gia xung đột đều tìm cách “chiến thắng,” lý thuyết về chiến lược không phủ nhận rằng những người tham gia xung đột vừa có lợi ích chung vừa có lợi ích xung đột với nhau. Trong thực tế, sự phong phú của vấn đề xuất phát từ thực tế rằng, trong các mối quan hệ quốc tế, luôn có sự phụ thuộc lẫn nhau cũng như sự đối nghịch lẫn nhau. Sự xung đột thuần túy, trong đó Tội ác của hai bên đối kháng xung khắc hoàn toàn, là một trường hợp đặc biệt. Nó xuất hiện trong một cuộc chiến tranh hủy diệt hoàn toàn, vì nếu không thì chiến tranh sẽ không nổ ra. Vì lý do đó, “thắng lợi” trong một cuộc xung đột không có ý nghĩa cạnh tranh thuần túy. Nó không hẳn là thắng lợi đối với đối thủ của mình. Nó có ý nghĩa cải thiện đối với hệ thống giá trị của chính mình, và điều này có

thể đạt được thông qua mặc cả, nhân nhượng lẫn nhau, và tránh tàn phá lẫn nhau. Nếu chiến tranh triệt để là điều không thể tránh khỏi, thì sẽ không còn gì khác ngoài sự xung đột thuần túy. Tuy nhiên, nếu còn có khả năng nào đó để tránh một cuộc chiến tranh tàn phá lẫn nhau, hay tiến hành một cuộc chiến tranh sao cho gây ít tàn phá nhất, hay gây sức ép lên đối thủ bằng đe dọa chiến tranh hơn là tiến hành chiến tranh, thì khả năng nhân nhượng lẫn nhau cũng sẽ quan trọng và kịch tính chẳng kém gì các yếu tố xung đột. Các khái niệm như sự rắn đe, chiến tranh hạn chế, và giải trừ quân bị, cũng như thương lượng, có liên quan đến lợi ích chung và sự phụ thuộc lẫn nhau - những điều có thể tồn tại giữa các bên tham gia cuộc xung đột.

Như vậy, chiến lược - theo nghĩa tôi đang dùng ở đây - không liên quan đến việc sử dụng sức mạnh có hiệu quả mà là *việc khai thác sức mạnh tiềm năng*. Nó không chỉ liên quan đến các đối thủ thù địch nhau mà còn liên quan đến các đối tác không tin cậy hay bất đồng với nhau. Nó không chỉ liên quan đến việc phân chia cái được và cái mất của hai bên mà còn liên quan đến khả năng đưa đến những kết cục cụ thể tồi tệ hơn (hay tốt hơn) cho cả hai bên so với những kết cục khác. Theo thuật ngữ chuyên ngành của lý thuyết trò chơi, hầu hết các cuộc xung đột quốc tế đáng chú ý không phải là "các trò chơi có tổng lợi ích cố định" mà là "các trò chơi có tổng lợi ích thay đổi": tổng lợi ích của các bên tham gia không phải là cố định, và do đó cái được của một bên không nhất thiết là cái thiệt của bên kia. Có một lợi ích chung trong việc đạt được những kết cục mà các bên đều có lợi.

Việc nghiên cứu chiến lược xung đột đòi hỏi có quan điểm

rằng hầu hết các tình huống xung đột về cơ bản là các tình huống *mặc cả*. Đó là những tình huống trong đó khả năng một bên tham gia xung đột đạt được mục đích của mình phụ thuộc ở một mức độ quan trọng vào sự lựa chọn hay quyết định của bên kia. Việc mặc cả có thể diễn ra công khai, như khi một bên đưa ra sự nhượng, hay có thể được thực hiện bằng một động thái ngầm, như khi một bên chiếm giữ hoặc rút lui khỏi vùng lãnh thổ chiến lược. Nó có thể, giống như việc mặc cả thường ngày diễn ra trên các thương trường, coi *nguyên trạng* là điểm bắt đầu và tìm kiếm những dàn xếp nhằm đem lại lợi nhuận dương cho cả hai bên, hay nó có thể dùng đến các đe dọa gây tổn thất cho nhau, như trong bái công, tẩy chay, hay chiến tranh giá cả, hay tổng tiền.

Nhin nhận hành vi xung đột như một quá trình mặc cả giúp chúng ta không quá bận tâm hoặc vào sự xung đột hay vào lợi ích chung. Việc coi các thủ đoạn và hành động của chiến tranh hạn chế như là một quá trình mặc cả đồng nghĩa với việc nhấn mạnh rằng, bên cạnh sự khác biệt về lợi ích trong các vấn đề tranh chấp còn có lợi ích chung to lớn trong việc đi đến một kết cục không tàn phá các giá trị của hai bên. Một cuộc bái công "thành công" của công nhân không phải là một cuộc bái công làm người chủ kiệt quệ về mặt tài chính, mà thậm chí có thể là một cuộc bái công không bao giờ xảy ra. Điều tương tự như vậy cũng đúng đối với chiến tranh.

Ý tưởng về "sự răn đe" đã có những phát triển hữu ích cho mục đích của chúng ta. Hơn mươi năm đã trôi qua kể từ khi sự răn đe được minh định như là hòn đá tảng trong chiến lược quốc gia, và trong những năm đó khái niệm này đã được trau chuốt và hoàn thiện. Chúng ta biết rằng một sự đe dọa cần

phải đáng tin cậy thì mới có tác dụng, và sự đáng tin cậy này có thể phụ thuộc vào chi phí và rủi ro đi kèm với việc thực hiện cho bên đưa ra đe dọa. Chúng ta đã phát triển ý tưởng làm cho một sự đe dọa đáng tin cậy bằng cách cam kết thực hiện sự đe dọa đó, thông qua việc “giăng bẫy” ngang qua con đường tiến của kẻ thù, hay bằng cách làm cho việc thực hiện cam kết trở thành vấn đề danh dự và uy tín quốc gia – chẳng hạn như trường hợp Nghị quyết Formosa. Chúng ta nhận ra rằng việc sẵn sàng chiến đấu trong một cuộc chiến tranh hạn chế tại các khu vực cụ thể có thể làm giảm mối đe dọa của sự trả đũa lớn, bằng cách duy trì việc chọn giải pháp ít đen tối hơn nếu một sự kiện nào đó xảy ra. Chúng ta đã cân nhắc khả năng là sự đe dọa trả đũa có thể đáng tin hơn nếu các phương tiện thực hiện và trách nhiệm đối với việc trả đũa được đặt trong tay những con người có quyết tâm lớn nhất, như trong các gợi ý gần đây về “chia sẻ hạt nhân.” Chúng ta quan sát thấy tính hợp lý của đối phương ảnh hưởng đến hiệu lực của sự đe dọa, và những kẻ điên rồ, giống như những đứa trẻ con, thường không bị kiểm soát bởi sự đe dọa. Chúng ta nhận thấy rằng hiệu lực của sự đe dọa còn tùy thuộc vào việc có phương án nào khác cho kẻ thù tiềm tàng áp dụng không, vì cũng phải để lại cho họ một lối thoát có thể chấp nhận được nếu không muốn họ phản ứng như một con sư tử bị mắc bẫy. Chúng ta đi đến nhận thức rằng một sự đe dọa trả đũa tổng lực sẽ cho kẻ thù mọi sự khích lệ, trong trường hợp họ quyết định bỏ ngoài tai sự đe dọa, để khởi xướng hành động vượt quá giới hạn bằng một cuộc tấn công tổng lực vào chúng ta. Nó loại trừ các hành động nhẹ hơn và buộc họ phải lựa chọn giữa các thái cực. Chúng ta học được rằng việc đe dọa hủy

diệt hàng loạt chỉ có thể răn đe được kẻ thù nếu như có một lời hứa ngầm tương ứng về việc không tiến hành hủy diệt trong trường hợp họ tuân thủ, do đó chúng ta phải cân nhắc liệu khả năng tấn công bất ngờ của ta có quá lớn khiến họ có thể tấn công trước để tránh bị tước vũ khí bởi sự tấn công trước của chúng ta. Và gần đây, liên quan đến cái gọi là "các biện pháp tự vệ chống lại các cuộc tấn công bất ngờ," chúng ta đã bắt đầu xem xét khả năng cải thiện sự răn đe lẫn nhau thông qua việc kiểm soát vũ khí.

Điều gây ấn tượng không phải là ý tưởng về sự răn đe đã trở nên phức tạp như thế nào, và nó đã được gạn lọc một cách cẩn thận và phát triển ra sao, mà là quá trình này đã diễn ra chậm chạp như thế nào, các khái niệm vẫn còn mơ hồ ra sao và lý thuyết răn đe hiện thời còn thô sơ ra sao? Nói như vậy không phải để hạ thấp nỗ lực của những người đã vật lộn với khái niệm răn đe trong hơn 10 năm qua. Về các vấn đề chiến lược, trong đó răn đe là một ví dụ, những người hoạch định chính sách cho những vấn đề cấp bách đã không được hỗ trợ hoặc hỗ trợ rất ít từ *cơ sở lý thuyết sẵn có mà họ phải tự nghĩ ra trong khi tiến hành công việc*. Không có tài liệu khoa học nào về vấn đề răn đe có thể so sánh được với các tài liệu về các vấn đề như lạm phát, bệnh cúm châu Á, sách giáo khoa tiểu học, hay sương mù.

Hơn nữa, những người vật lộn với các ý tưởng như sự răn đe, chịu nhiều sức ép bởi các vấn đề trước mắt, không quan tâm nhiều đến quá trình phát triển của một hệ thống lý thuyết. Điều này đúng không chỉ đối với các nhà hoạch định chính sách và các nhà báo mà còn đối với các học giả nữa. Bất kể nó phản ánh sự quan tâm của các học giả hay các tổng biên

tập, các nghiên cứu về sự răn đe và các khái niệm liên quan phần lớn tập trung nói về việc giải quyết các vấn đề trước mắt hơn là phương pháp luận giải quyết vấn đề.<sup>2</sup> Chúng ta thậm chí còn không có một bộ thuật ngữ thích đáng. Các thuật ngữ thỉnh thoảng xuất hiện như sự răn đe "chủ động" và "bị động" chưa đủ để đáp ứng được nhu cầu.

Chúng ta giải thích như thế nào đối với tình trạng kém phát triển về lý thuyết? Tôi nghĩ có một thực tế quan trọng là quân sự, trái với hầu hết các ngành chuyên môn có quy mô và tầm quan trọng khác, không đi kèm với một ngành học thuật rõ ràng. Những nhà hoạch định chính sách trong các lĩnh vực kinh tế, y học, y tế, bảo quản đất trồng trọt, giáo dục, hay luật hình sự, có thể tìm được ngay đối tác trong thế giới học thuật. (Trong kinh tế học số lượng những người được đào tạo làm công tác nghiên cứu khoa học và viết sách có thể sánh ngang với số lượng người đang tham gia vào lĩnh vực hoạch định chính sách kinh tế hay quản lý). Vậy thì đối tác học thuật của ngành quân sự hiện nằm ở đâu?

Nó không nằm - với quy mô lớn - trong các học viện quân sự. Đây là những trường đào tạo ở mức đại học, chủ yếu dành

---

<sup>2</sup> Có một số ví dụ rất hay theo hướng ngược lại, như C. W. Sherwin, "Bảo đảm hòa bình thông qua công nghệ quân sự," *Bulletin of the Atomic Scientists* (tháng 5-1956). Và sự tham khảo của Sherwin đến một bài viết của Warren Amster nhắc nhở chúng ta rằng khi lý thuyết được thúc đẩy bởi các vấn đề quân sự như phần lớn đang diễn ra hiện nay, nó có thể không được xuất bản công khai. Ở đây chắc chắn có trở ngại ở khâu biên tập. Các tờ báo về quan hệ quốc tế chủ yếu thu hút độc giả phi lý thuyết, và các bài báo chứa đựng nhiều lý thuyết thường bị thanh lọc và tập trung vào các vấn đề trước mắt. Sự đóng góp gần đây của tờ *Conflict Resolution* dành cho tiểu luận tuyệt vời của Anatol Rapoport về "Lý thuyết toán học chiến tranh của Lewis F. Richardson" (tập I, số 3, tháng 9-1957) là một dấu hiệu đáng khích lệ theo hướng khác.

để giảng dạy hơn là nghiên cứu. Nó cũng không nằm - hay chưa đạt quy mô lớn - tại các trường quân sự và các học viện giáo dục cao cấp phi kỹ thuật khác thuộc ngành quân sự. Những nơi này chưa xây dựng đội ngũ giảng dạy chính quy, các định hướng nghiên cứu, và hệ thống giá trị cần thiết cho sự phát triển lý thuyết bền vững và có hệ thống.

Tại các trường đại học, chiến lược quân sự luôn là mối quan tâm của một số ít các nhà sử học và các nhà khoa học chính trị, ủng hộ một giả thuyết cho rằng việc ngăn chặn người Nga thôn tính châu Âu là điều quan trọng chẳng kém gì việc thực thi luật chống độc quyền. Nói như thế không phải để xem thường những thành tựu đã đạt được, mà để nhấn mạnh rằng trong các trường đại học thường không có khoa hay ngành nghiên cứu trực tiếp liên quan với ngành quân sự và vai trò của sức mạnh quân sự trong quan hệ ngoại giao. (Các chương trình ROTC gần đây đã trở thành một trường hợp ngoại lệ có giới hạn đối với vấn đề này, ít ra với mức độ là chúng thúc đẩy việc sắp xếp các môn học thích đáng về lịch sử và khoa học chính trị). Việc các chương trình và các viện nghiên cứu quốc phòng giờ đây đã có ở một số trường đại học, và sự quan tâm dành cho các vấn đề an ninh quốc tế của các cơ sở này là bước chuyển mới mẻ và quan trọng. Các việc nghiên cứu bán chính phủ như RAND Corporation và Viện Phân tích Quốc phòng đang hỗ trợ rất đáng kể để đáp ứng nhu cầu, nhưng so với mục tiêu của chúng ta, điều đó chỉ nên coi là chỉ dấu của nhu cầu mà thôi.

Người ta có thể đặt câu hỏi phải chăng bản thân ngành quân sự không có khả năng đưa ra một bộ lý thuyết nhằm soi sáng cho các ý tưởng như sự rắn đe hay chiến tranh hạn chế. Xét cho cùng, một lý thuyết không nhất thiết phải được phát

triển chỉ bởi các chuyên gia nằm cô lập trong các trường đại học. Nếu ngành quân sự cần được chuẩn bị một cách chu đáo để *sử dụng* sức mạnh quân sự có hiệu quả, có lẽ họ phải được trang bị để phát triển lý thuyết về việc đó. Nhưng ở đây, chúng ta cần phân biệt rõ giữa việc *sử dụng* sức mạnh và *việc đe dọa* (*sử dụng*) sức mạnh. Sự răn đe nói về việc kiai thắc sức mạnh tiềm năng. Nó quan tâm đến việc thay đổi một kẻ thù tiềm tàng rằng việc tránh một số hành động nhất định là vì lợi ích của chính hắn. Có một sự khác biệt lớn giữa các kỹ năng trí tuệ đòi hỏi để thực hiện một sứ mệnh quân sự và để sử dụng *tiềm năng* quân sự nhằm đạt được các mục tiêu của một quốc gia. Lý thuyết về sự răn đe nhìn chung phải là một lý thuyết về kỹ năng *không sử dụng* các sức mạnh của quân sự, và để đạt được mục đích này sự răn đe đòi hỏi điều gì đó rộng lớn hơn là các kỹ năng quân sự. Các ngành quái i sự có thể có các kỹ năng rộng hơn này, nhưng họ không tự i lộng có được kỹ năng đó nếu họ chỉ lo đáp ứng các trách nhiệm của chính của mình, mà những trách nhiệm chủ yếu này lại chỉ m i toàn bộ thời gian của họ.<sup>3</sup>

---

<sup>3</sup> Việc thiếu một truyền thống học thuật vững vàng trong lĩnh vực chiến lược quân sự được Bernard Brodie đề cập đến trong các chương đầu của cuốn sách của ông: *Chiến lược trong thời đại tên lửa* (Strategy in the Missile Age) (Princeton, 1959). Đáng chú ý nữa là lời nói đầu của Đại tá Joseph I. Greene, trong cuốn *Về chiến tranh* (On War) của Clausewitz do Modern Library xuất bản (New York, 1943): "Trong hầu hết những năm giữa các cuộc đại chiến, hai trường học lớn nhất của Quân đội chúng ta chỉ giới hạn trong việc dạy một khóa học kéo dài khoảng 10 tháng cho tất cả các sĩ quan được lựa chọn để tham dự... Có lẽ hai nơi này chẳng có mấy thời gian nghiên cứu sự phát triển lâu dài của các ý tưởng và lý thuyết quân sự... Nếu có được thời gian nhiều hơn cho việc đào tạo cao hơn trong quân đội - khoảng hai hay ba năm - thì các nhà tư tưởng quân sự vĩ đại nhất chắc hẳn sẽ là những đẽ tài đáng để nghiên cứu".

15 năm trước đây, một ngành nghiên cứu mới đã đem lại hy vọng đưa đến một lý thuyết về chiến lược, đó là lý thuyết trò chơi. Lý thuyết trò chơi nói về các tình huống - các trò chơi "chiến lược," khác với các trò chơi về kỹ năng hay các trò chơi may rủi - trong đó bước đi tối ưu của mỗi người tham gia phụ thuộc vào việc anh ta dự kiến các đối thủ khác sẽ có bước đi như thế nào. Sự đe dọa mang tính răn đe rất thích hợp với định nghĩa này. Sự đe dọa chỉ có kết quả khi đối thủ dự kiến chúng ta sẽ làm gì đó để phản ứng lại bước đi mà anh ta lựa chọn, và chúng ta chỉ có khả năng đưa ra đe dọa nếu dự kiến được hành động của chúng ta sẽ có ảnh hưởng đến sự lựa chọn của anh ta. Nhưng trong chiến lược quốc tế sự hứa hẹn của lý thuyết trò chơi cho đến nay vẫn chưa có kết quả. Lý thuyết trò chơi là cực kỳ hữu ích trong việc xác lập các vấn đề và làm rõ các khái niệm, nhưng thành công lớn nhất của nó lại nằm ở các lĩnh vực khác. Nhìn chung, lý thuyết này đã bị chôn chân ở mức độ trừu tượng có rất ít liên hệ với các yếu tố của một vấn đề như sự răn đe.<sup>4</sup>

Ý tưởng về sự răn đe xuất hiện rõ trong nhiều lĩnh vực xung đột ngoài lĩnh vực quan hệ quốc tế đến nỗi có một số người cho rằng đã có một lý thuyết được xây dựng vững chắc để có thể áp dụng vào lĩnh vực quan hệ quốc tế. Từ lâu,

<sup>4</sup>Bài viết của Jessie Bernard về "Lý thuyết trò chơi như một môn xã hội học hiện đại về xung đột," nêu ra sự đánh giá tương tư và nhấn mạnh rằng "chúng ta có thể hy vọng rằng toán học cần để áp dụng có kết quả lý thuyết trò chơi cho các hiện tượng xã hội sẽ xuất hiện trong tương lai không xa" (*The American Journal of Sociology* 59:418, tháng 3-1954). Quan điểm của tôi là các khuyết điểm hiện nay không phải ở toán học, và lý thuyết về chiến lược đã và đang gánh chịu thiệt thòi vì nhiệt tình quá lớn của các nhà khoa học xã hội muốn đổi xử với môn học này như thể nó đã là hoặc phải là một nhánh của toán học.

sự răn đe là một khái niệm quan trọng trong luật hình sự. Các nhà lập pháp, các chánh án, các luật sư, và các học giả pháp lý có thể đã nghiên cứu khái niệm này một cách sâu sắc và có hệ thống qua nhiều thế hệ. Quả vậy, sự răn đe không phải là điều được suy xét duy nhất trong luật hình sự, và không nhất thiết là điều quan trọng nhất. Tuy vậy, nó được nhắc đến nhiều đến nỗi người ta cho rằng đã có một lý thuyết tính đến hình thức và mức độ các hình phạt dành cho tội phạm bị kết án, hệ thống giá trị của tội phạm tiềm tàng, lợi ích thu được từ việc phạm tội, khả năng của hệ thống thực thi pháp luật để bắt giữ và kết án tội phạm, nhận thức của tội phạm về pháp luật và về khả năng bị bắt giữ và kết án, mức độ mà các loại tội phạm khác nhau bị thúc đẩy bởi sự tính toán theo lý trí, sự quyết tâm của xã hội không dè dặt và không mềm yếu trong việc áp dụng các hình phạt tối kém và còn nhiều ý kiến bất đồng và sự quyết tâm (hay thiếu quyết tâm) này được bọn tội phạm biết đến đâu, khả năng mắc lỗi của hệ thống, và khả năng các bên thứ ba lợi dụng hệ thống cho mục đích cá nhân, vai trò của trao đổi thông tin giữa xã hội có tổ chức và tội phạm, tổ chức của bọn tội phạm nhằm chống lại hệ thống pháp luật, v.v...

Tuy nhiên, không phải chỉ có bọn tội phạm, mà ngay cả con em của chúng ta cũng cần phải được răn đe. Một số khía cạnh của sự răn đe có vị trí nổi bật trong tính kỷ luật của đứa trẻ: tầm quan trọng của sự hợp lệ phải và kỷ luật tự giác đối với cá nhân bị răn đe, của khả năng hiểu được mối đe dọa khi nghe lời đe dọa đó và hiểu rõ và quyết tâm thực hiện lời đe dọa nếu cần thiết - và, quan trọng hơn là niềm tin của bên bị đe dọa về việc mối đe dọa sẽ được thực hiện. Có lẽ việc thực

hiện sự trùng phạt có khả năng làm đau đớn người đưa ra đe dọa nhiều hơn trong trường hợp kỷ luật đưa trẻ so với trường hợp răn đe tội phạm. Có một sự giống nhau giữa sự đe dọa của cha mẹ đối với một đứa trẻ và sự đe dọa của một mẫu quốc giàu có đối với một chính phủ yếu kém của một nước nghèo, chẳng hạn đưa ra viện trợ cho nước nghèo và đòi hỏi nước đó phải đáp lại bằng các chính sách kinh tế “đúng đắn” hay các chính sách hợp tác quân sự.

Và sự giống nhau này nhắc nhở chúng ta rằng, ngay cả trong quan hệ quốc tế, sự răn đe sẽ có tác dụng đối với các quan hệ giữa bạn bè cũng như đối với các quan hệ giữa kẻ thù tiềm tàng. (Việc đe dọa sẽ sử dụng “chiến lược ngoại vi” nếu Pháp không phê chuẩn Hiệp ước Phòng thủ chung châu Âu cũng gặp những hạn chế tương tự với tư cách một đe dọa trả đũa). Khái niệm răn đe đòi hỏi đồng thời phải có cả sự xung đột lẫn lợi ích chung giữa các bên liên quan. Nó sẽ không có tác dụng trong trường hợp có đối kháng lợi ích tuyệt đối cũng như trong trường hợp có sự thống nhất lợi ích tuyệt đối. Giữa các thái cực này, việc răn đe một đồng minh và việc răn đe một kẻ thù chỉ khác nhau về mức độ, và trong thực tế, chúng ta có thể cần phải xây dựng một lý thuyết mạch lạc hơn trước khi có thể phát biểu một cách có ý nghĩa rằng chúng ta có lợi ích chung nhiều hơn với Nga hay với Hy Lạp, so với những xung đột giữa chúng ta.<sup>5</sup>

---

<sup>5</sup>Có một điều quan trọng cần nhấn mạnh là, khi nói đến “lợi ích chung,” tôi không muốn nói phải có cái gì đó giống nhau mà thường được đề cập đến là sự tương đồng trong hệ thống giá trị của các bên. Có thể đơn giản chỉ là họ ở trên cùng một con thuyền: họ thậm chí cùng ở trên con thuyền đó chỉ vì một người trong số họ cho rằng sẽ có lợi thế chiến lược

Ý tưởng về sự răn đe cũng thường thấy trong cuộc sống hàng ngày. Những người lái xe ô-tô có một lợi ích chung hiển nhiên trong việc tránh không cho xe đụng đầu vào nhau và có xung đột lợi ích về việc ai sẽ là người đi trước và ai sẽ là người phải đạp phanh xe để nhường cho người khác vượt lên trước. Đụng xe có tính tương hỗ tối đa và thường là điều duy nhất người ta có thể dùng để đe dọa. Những động tác mà một lái xe dùng để chuyển tải mối đe dọa về hư hỏng cho cả hai bên khi một người khác đang lấn đường của mình là một ví dụ sinh động về một kiểu đe dọa được truyền đạt không phải bằng lời nói mà bằng hành động, và về một đe dọa trong đó quyết tâm thực hiện mối đe dọa đó không thể hiện bằng thông báo miệng mà bằng việc không thể làm khác đi.

Cuối cùng, còn có lĩnh vực quan trọng về thế giới ngầm. Cuộc chiến giữa các băng đảng và cuộc chiến quốc tế có rất nhiều điểm chung. Các quốc gia và các băng đảng sống ngoài vòng pháp luật đều thiếu một hệ thống pháp lý có hiệu lực cưỡng chế để giúp họ quản lý công việc. Cả hai đều sử dụng bạo lực đến cùng. Cả hai đều có lợi ích trong việc tránh sử dụng bạo lực, nhưng vẫn luôn dùng đến biện pháp đe dọa dùng bạo lực. Điều lý thú là những kẻ tống tiền cũng như các băng đảng nhỏ cũng tham gia vào chiến tranh hận ché, giải trừ vũ khí và ngưng bắn, tấn công bất ngờ, trả đũa và đe dọa

---

tốt hơn nếu ở vào vị trí đó - để nối kết lợi ích của họ trong việc không để cho con thuyền bị lật. Nếu như cùng bị lật nhào với nhau trên con thuyền là một kết cục tiêm tàng, xét đến mọi lựa chọn sẵn có của cả hai bên, thì họ có một "lợi ích chung" theo nghĩa đã nêu ra ở đây. "Lợi ích chung tiêm tàng" có vẻ như mô tả sát nghĩa hơn. Chẳng hạn, sự răn đe liên quan đến việc nối kết hành động của một bên với hành động của bên kia nhằm khai thác lợi ích chung tiêm tàng đó.

trả đũa, họ cũng tính đến “nhượng bộ” và sự mất thể diện, và họ cũng lập các liên minh và đi đến các thỏa thuận với cùng mức độ hạn chế mà các quốc gia gặp phải – sự không có khả năng khiếu nại lên tổ chức có quyền lực cao hơn để bảo đảm thực thi thỏa thuận.

Do đó chúng ta có thể nghiên cứu những lĩnh vực khác để có thêm hiểu biết về lĩnh vực mà chúng ta quan tâm, lĩnh vực quan hệ quốc tế. Thông thường thì một nguyên tắc trong lĩnh vực mà chúng ta quan tâm hay bị che khuất bởi nhiều chi tiết vụn vặt, hoặc cấu trúc của nó quá phức tạp, hoặc do chúng ta không thể nhìn thấy nó vì thiên kiến, nhưng lại dễ dàng được nhận ra ở các lĩnh vực khác là nơi có sự đơn giản và dễ thấy hoặc nơi mà chúng ta không bị che mắt bởi các thiên kiến của mình. Chúng ta sẽ có thể dễ dàng hiểu được sự khó khăn to lớn trong việc kiềm chế một người như Mossadeq thông qua đe dọa khi ta vừa mới trải qua thất bại trong nỗ lực dùng đe dọa để ngăn một đứa trẻ nhỏ đánh một con chó hay ngăn một con chó con cắn một đứa trẻ.

Không có lĩnh vực nào trong những lĩnh vực xung đột khác này được đúc kết thành lý thuyết vững chắc mà với sự bổ sung điều chỉnh có thể dùng để phân tích các vấn đề quốc tế. Các nhà xã hội học, kể cả những nhà nghiên cứu hành vi tội phạm trong các cuộc xung đột thuộc thế giới ngầm, theo truyền thống, thường không quan tâm nhiều đến *chiến lược* xung đột. Các tài liệu sách báo về luật và tội phạm học cũng không cho thấy có một lý thuyết khả dĩ nào về vấn đề này. Tôi không thể khẳng định chắc chắn rằng không hề có một cuốn sổ tay, sách giáo khoa, hay công trình nghiên cứu nào về lý thuyết tổng tiền được lưu hành trong thế giới ngầm. Nhưng chắc

chán không có một phiên bản nào đã được trải nghiệm, nói về phương pháp tổng tiền hay làm thế nào để chống lại nó, được đưa ra làm “Các phương pháp mới hướng dẫn trẻ em” mặc dù chúng ta cần đến nó.<sup>6</sup>

“Lý thuyết” trong lĩnh vực chiến lược này sẽ bao gồm những yếu tố gì? Nó sẽ trả lời cho những câu hỏi nào? Những ý tưởng nào sẽ được nó liên kết, làm rõ, hay truyền đạt một cách có hiệu quả hơn? Để bắt đầu, nó cần vạch ra những nét cơ bản của tình huống và hành vi cần xem xét. Sự răn đe - sự răn đe tiếp tục là khái niệm chiến lược điển hình - nói về việc gây ảnh hưởng đến sự lựa chọn hành vi của đối phương và để làm được điều đó cần gây tác động đến các dự đoán của anh ta về hành vi mà chúng ta sẽ thực hiện. Điều đó bao gồm việc đưa ra các bằng chứng làm cho anh ta tin rằng hành vi của chúng ta sẽ do hành vi của anh ta quyết định.

Nhưng cấu trúc các hệ thống giá trị của hai bên tham gia phải như thế nào - như hình thức “lợi ích” (payoffs), trong ngôn ngữ của lý thuyết trò chơi - mới làm cho sự răn đe đáng tin cậy? Làm cách nào chúng ta đo lường được sự kết hợp giữa xung đột và lợi ích chung, là những thứ cần thiết để hình thành một tình huống “răn đe”? Cần có kênh truyền đạt nào và làm cách nào để xác minh những bằng chứng được truyền đạt? Bên bị răn đe cần có tính “hợp lý” như thế nào - sự hiểu rõ hệ thống giá trị của anh ta, khả năng nhận thức về các giải pháp thay thế và khả năng tính toán các tình huống có thể

---

<sup>6</sup> Việc này đang có tiến bộ. Daniel Ellsberg đã đưa ra bài giảng về “Lý thuyết và Thực tiễn về Hỗn loạn,” và một bài khác về “Các ứng dụng chính trị của sự diễn rõ,” trong loạt bài của ông về “Nghệ thuật Cưỡng ép,” dưới sự tài trợ của Lowell Institute, Boston, tháng 3-1959.

xảy ra, khả năng thể hiện (hay không có khả năng che giấu) tính hợp lý của mình?

Điều gì cần thiết cho sự tin tưởng, hay khả năng thực thi các lời hứa? Nói một cách cụ thể hơn, bên cạnh việc đưa ra đe dọa gây tổn thất, chúng ta còn cần phải bảo đảm rằng việc gây tổn thất sẽ được rút lại nếu có sự hứa hẹn tuân thủ. Nói một cách khác có phải điều này phụ thuộc vào cấu trúc của "lợi ích" không? Chúng ta cần có "hệ thống pháp lý nào," hệ thống truyền đạt gì, hay cơ cấu thông tin nào để làm những lời hứa quan trọng có tính thực thi?

Liệu ai đó có thể đưa ra sự đe dọa rằng anh ta "có thể" sẽ thực hiện lời đe dọa, hoặc anh ta buộc phải đe dọa rằng anh ta chắc chắn sẽ làm như vậy? Ý nghĩa của sự đe dọa là gì, khi người đưa ra sự đe dọa nói "có thể" sẽ thực hiện trong khi rõ ràng là nếu anh ta còn giữ lại cho mình một sự lựa chọn, anh ta sẽ không có động cơ để thực hiện sự đe dọa đó? Nói rộng hơn, có cách nào làm cho một người phải thực hiện cam kết mà lê ra anh ta có thể sẽ không thực hiện, trong tình huống nếu một cam kết đưa ra đủ sức tin cậy làm cho lời đe dọa có tác dụng mà không cần phải thực hiện cam kết đó nữa. Đâu là sự khác nhau, nếu có, giữa một đe dọa ngăn chặn được hành động và một đe dọa thúc đẩy có hành động, hay là một đe dọa nhằm ngăn chặn bên thứ hai trước lỗi lầm của chính anh ta? Có sự khác nhau nào không về mặt logic giữa các đe dọa có tính răn đe, kỷ luật, và cưỡng ép?

Tình huống sẽ bị ảnh hưởng như thế nào khi có bên thứ ba, là người có các xung đột và lợi ích chung với các bên đã có mặt, và là người tiếp cận hay kiểm soát hệ thống trao đổi

thông tin, mà hành vi của người này có thể là hợp lý hay không hợp lý theo một nghĩa nào đó, và người này giành được niềm tin hay một số phương tiện thực thi thỏa thuận với một trong hai bên đối thủ chính? Những câu hỏi này sẽ bị ảnh hưởng ra sao khi có một hệ thống pháp lý cho phép thực hiện hoặc nghiêm cấm một số hành động nhất định, áp đặt hình phạt khi không thực hiện hợp đồng, hay đòi hỏi các bên phải cung cấp thông tin xác thực. Chúng ta có thể lý giải đến mức độ nào các khái niệm như "uy tín," "thể diện," hay "sự tin cậy" dưới góc độ của một hệ thống pháp lý có thực hay giả định, dưới góc độ bổ sung sửa đổi các hệ thống giá trị của các bên tham gia, hay dưới góc độ của các mối quan hệ giữa những người tham gia chính và phụ, có thực hay giả định?

Những câu hỏi mang tính ví dụ ở trên nói lên rằng chúng ta có điều kiện để xây dựng "lý thuyết." Ở đây có cái gì đó trông giống như sự pha trộn của các lý thuyết như lý thuyết trò chơi, lý thuyết tổ chức, lý thuyết thông tin liên lạc, lý thuyết bằng chứng, lý thuyết lựa chọn, và lý thuyết quyết định tập thể. Sự pha trộn này phù hợp với định nghĩa "chiến lược" của chúng ta: nó coi xung đột là điều dĩ nhiên, nhưng nó cũng thừa nhận có lợi ích chung giữa các đối thủ, nó thừa nhận hành vi tối đa hóa giá trị một cách "hợp lý", và nó tập trung vào thực tế rằng quyết định hành động "tối ưu" của một bên phụ thuộc vào việc bên này dự đoán bên kia sẽ hành động như thế nào, và "hành vi chiến lược" bao gồm việc gây ảnh hưởng đến lựa chọn của đối thủ bằng cách tác động vào dự đoán của đối thủ về hành vi của mình sẽ phụ thuộc vào hành vi của đối thủ như thế nào.

Ở đây có hai điểm đáng quan tâm. Thứ nhất, mặc dù "chiến

lực xung đột” nghe có vẻ tàn bạo, nhưng lý thuyết này không liên quan gì đến việc áp dụng có hiệu quả bạo lực hay đại loại như thế. Về cơ bản nó không phải là một lý thuyết về sự xâm lược, kháng chiến, hay về chiến tranh. *Sự đe dọa* dùng chiến tranh thì có, cũng như mọi sự đe dọa khác. Nhưng nó nói về việc sử dụng các mối đe dọa, hay sự đe dọa và hứa hẹn, hay nói rộng hơn là việc điều chỉnh hành vi của một bên theo hành vi của những bên khác. Đó là đối tượng nghiên cứu của lý thuyết này.

Thứ hai, lý thuyết này không phân biệt giữa xung đột và lợi ích chung, cũng như giữa khả năng áp dụng đối với kẻ thù và đồng minh tiềm tàng. Lý thuyết này tiêu biến ở thái cực khi không có chỗ cho sự nhân nhượng lẫn nhau, không có lợi ích chung nào trong việc tránh thảm họa chung. Nó cũng sẽ tiêu biến ở thái cực kia khi không có xung đột nào cả và không có khó khăn gì trong việc nhận diện và đạt được các mục tiêu chung. Nằm giữa hai thái cực nói trên, lý thuyết này không áp đặt thiên kiến lên hồn hợp giữa xung đột và lợi ích chung. Chúng ta vừa có thể gọi đây là lý thuyết về sự liên kết mong manh hay lý thuyết về chủ nghĩa đối kháng không triệt để.<sup>7</sup> (Trong Chương 9, tôi có chỉ ra rằng một số khía cạnh chủ chốt của vấn đề tấn công bất ngờ trong các quan hệ quốc tế có cơ cấu đồng nhất với vấn đề về các đối tác nghi ngờ lẫn nhau.)

---

<sup>7</sup> Khi sử dụng từ “mối đe dọa”, tôi không nhất thiết hàm ý xâm lược hay thù địch. Trong quá trình thương lượng công khai giữa bạn bè hay sự hợp tác ngầm giữa họ, mối đe dọa (hay nguy cơ) về sự bất đồng hay giảm hợp tác, được nêu công khai hay ám chỉ là một hình phạt mà các bên sử dụng để cản trở đòi hỏi của mình, cũng như trong giao dịch thương mại sự chào hàng sẽ được củng cố bằng sự đe dọa “không bán.”

Cả hai điều này - sự trung lập của lý thuyết đối với mức độ sự xung đột diễn ra, và định nghĩa "chiến lược" có liên quan đến việc giới hạn đối phương thông qua dự tính của đối phương về những hậu quả do hành động của mình gây ra - gợi ra rằng chúng ta có thể gọi đề tài này là *lý thuyết về quyết định có tính phụ thuộc lẫn nhau* (theory of interdependent decision).

Ta có thể chia đe dọa và phản ứng lại đe dọa, trả đũa và chống trả đũa, chiến tranh giới hạn, chạy đua vũ trang, đe dọa gây chiến, tấn công bất ngờ, tin tưởng và lừa gạt thành những hành động nóng nảy hoặc bình tĩnh. Trong việc xây dựng lý thuyết, việc cho rằng sẽ là hợp lý nếu coi hành động xuất phát từ sự bình tĩnh không nhất thiết khẳng định rằng những hành động trong thực tế hoàn toàn dựa trên sự bình tĩnh. Điều đó chỉ khẳng định rằng giả định về hành vi hợp lý góp phần xây dựng lý thuyết có tính hệ thống. Nếu hành vi trong thực tế đều dựa trên sự bình tĩnh, chúng ta sẽ dễ dàng xây dựng lý thuyết đúng đắn và xác đáng hơn so với những gì chúng ta đang có. Nếu chúng ta coi các kết quả tìm được là cột mốc để tiếp tục xây dựng lý thuyết gần hơn với thực tiễn, chứ không phải xây dựng một lý thuyết hoàn bì, thì chúng ta có thể tránh được những kết cục tồi tệ nhất do một lý thuyết thiên lệch gây ra.

Thêm vào đó, một lý thuyết được dựa trên giả định rằng các bên tham gia tính toán lợi ích một cách bình tĩnh và "hợp lý" theo một hệ thống giá trị nhất quán sẽ buộc chúng ta phải suy nghĩ thấu đáo hơn về ý nghĩa của "sự bất hợp lý." Những người ra quyết định không chỉ đơn thuần nằm giữa hai thái cực hoàn toàn hợp lý và hoàn toàn vô lý. Tính hợp lý là tập

hợp của nhiều thuộc tính, và sự xa rời khỏi tính hợp lý triệt để có thể diễn ra theo nhiều hướng khác nhau. Tính bất hợp lý có thể bắt nguồn từ một hệ thống giá trị lộn xộn và thiếu nhất quán, sai lầm trong tính toán, khả năng tiếp nhận hay truyền đạt thông tin thiếu hiệu quả. Nó cũng có thể là biểu hiện của các ảnh hưởng ngẫu nhiên hoặc bất ngờ đến quá trình đạt được hay truyền động quyết định, hoặc đến việc tiếp nhận hoặc truyền đạt thông tin. Và đôi khi, nó chỉ là sự phản ánh tính tập thể trong quyết định do các cá nhân đưa ra, những người mà không có cùng một hệ thống giá trị, và cơ cấu tổ chức và hệ thống trao đổi thông tin của họ không làm cho các thành viên hành động như một thực thể duy nhất.

Trong thực tế, nhiều yếu tố quan trọng trong mô hình về hành vi hợp lý có thể được chia ra thành các hình thức cụ thể của tính hợp lý hoặc tính bất hợp lý. Chúng ta có thể coi hệ thống giá trị, hệ thống liên lạc, hệ thống thông tin, quá trình ra quyết định tập thể, hay một thông số thể hiện xác suất có sai sót hay mất kiểm soát là những đóng góp giúp hình thành ngành nghiên cứu về "tính bất hợp lý." Hitler, Quốc hội Pháp, người chỉ huy trên máy bay ném bom, các nhân viên điều khiển ra-đa tại Trân Châu Cảng, Khrushchev, và cử tri Mỹ đều có thể "bất hợp lý" theo một kiểu nào đó, nhưng tất nhiên không phải cùng một loại. Một số trong những người này có thể được tìm thấy trong lý thuyết về hành vi hợp lý. (Ngay cả một người mắc chứng loạn thần kinh chức năng, với các giá trị thiếu nhất quán không thể dung hòa, bị thôi thúc đè nén hơn là dung hòa các mục tiêu xung đột của anh ta, vì một lý do nào đó, có thể được xem là một *cấp* của các thực thể "hợp lý" với các hệ thống giá trị riêng biệt, đi đến các quyết định

tập thể thông qua một quá trình bỏ phiếu chịu ảnh hưởng của các yếu tố bất ngờ hay ngẫu nhiên, các hệ thống trao đổi thông tin bất đối xứng, v.v...)

Tính hạn chế hiển nhiên của giả định về hành vi “hợp lý” - bao gồm những tính toán và chiến lược ra quyết định nhằm tối đa hóa lợi ích - được giảm nhẹ bởi hai nhận xét bổ sung. Nhận xét thứ nhất, điều thứ yếu đối với tôi, là chúng ta thấy ngay cả những người mất thăng bằng về tinh cảm, trong số những người được chẩn đoán là “mất trí,” thường đề cao các nguyên tắc chiến lược, hay ít nhất là áp dụng chúng trong các trường hợp cụ thể. Tôi được người ta kể rằng bệnh nhân của các bệnh viện tâm thần thường nuôi dưỡng, có ý thức hay do bản năng, các hệ thống giá trị làm cho họ ít chịu đe dọa kỷ luật hơn và nâng cao khả năng tự chủ. Một thái độ bất cần và thậm chí tự hủy hoại bản thân - “Tôi sẽ cắt động mạch trên cổ tay của tôi nếu anh không cho phép tôi...” - có thể là một lợi thế chiến lược thực sự. Điều này cũng đúng đối với việc tỏ ra không nghe được hay nghe không hiểu, hay nói tiếng vì nhiều lần mất kiểm soát bản thân, khiến cho sự răn đe thông qua đe dọa trừng phạt mất đi tác dụng. (Điều này lại làm tôi nhớ đến các đứa con của mình.) Quả thực, một trong những ưu điểm của một lý thuyết tường minh về quyết định chiến lược “hợp lý” trong các tình huống hỗn hợp giữa xung đột và lợi ích chung là, bằng cách chỉ ra cơ sở chiến lược của các chiến thuật mâu thuẫn nào đó, nó có thể cho thấy một số chiến thuật vững vàng và hợp lý ra sao đến nỗi những kẻ dốt nát và nhu nhược cũng áp dụng. Không có gì là quá đáng khi nói rằng sự tinh tế của chúng ta đôi khi kìm hãm những trực giác đúng đắn, và một trong những hệ quả

của một lý thuyết tường minh có thể là phục hồi lại những nhận thức trực giác chỉ “vô lý” ở bì ngoài mà thôi.

Nhận xét thứ hai có liên quan đến nhận xét thứ nhất. Đó là một lý thuyết tường minh về quyết định “hợp lý,” và về các hậu quả chiến lược của các quyết định đó, cho thấy rất rõ là trong các tình huống xung đột việc luôn giữ vững tinh hợp lý trong các quyết định và động cơ không hẳn đã tốt. Nhiều thuộc tính của tinh hợp lý, như trong một vài minh họa nói trên, là những hạn chế chiến lược trong những tình huống xung đột nhất định. Có thể là hoàn toàn hợp lý khi ai đó mong cho mình không còn hợp lý một chút nào nữa, hay - nếu cách nói đó có thể bị bác bỏ về mặt triết lý - mong cho mình có sức mạnh để tạm ngưng một số năng lực hợp lý trong những tình huống cụ thể nào đó. Và người ta có thể đình hoãn hoặc hủy bỏ “tinh hợp lý” của chính mình, ít nhất là trong một mức độ hạn chế. Người ta có thể làm được điều này vì những thuộc tính tạo nên tinh hợp lý không phải là bất di bất dịch, mà lại mang tính cá nhân sâu sắc, là bộ phận cấu thành trong tâm hồn con người, nhưng nó lại bao gồm những thứ như máy trợ thính, độ tin cậy của bưu chính, hệ thống pháp lý, và tinh hợp lý của người đại diện và đối tác. Về nguyên tắc, một người có thể tránh bị tống tiền một cách dễ dàng bằng cách uống thuốc gây bất tỉnh, có lập bản thân về địa lý, để tài sản của mình sung công một cách hợp pháp, hay bẻ gãy cánh tay dùng để ký séc. Trong một lý thuyết về chiến lược, một số cách phòng thủ như vậy được đề cập đến như việc làm suy giảm tinh hợp lý nếu chúng ta muốn như vậy. Một lý thuyết dựa tinh hợp lý trở thành một định đè tường minh không chỉ có khả năng sửa đổi bổ sung định đè đó và xem xét ý nghĩa của nó mà còn làm

cho nó mất đi sự bí hiểm. Quả thực, vai trò nghịch lý của “tính hợp lý” trong các tình huống xung đột này là bằng chứng về sự hữu ích mà một lý thuyết có tính hệ thống có thể đem lại.

Và vì vậy, những kết quả tìm được nhờ phân tích lý thuyết hành vi chiến lược thường có phần nghịch lý. Chúng thường mâu thuẫn với lý lẽ thông thường và các quy tắc đã được thừa nhận. Quả là không đúng, như trong ví dụ về việc tống tiền, nếu nói rằng khi đối mặt với một sự đe dọa việc luôn luôn tỏ ra hợp lý là có lợi, đặc biệt khi ta không thể che giấu được việc mình hợp lý hay bất hợp lý. Khi đối mặt với một mối đe dọa, việc có một hệ thống liên lạc hoạt động tốt, có thông tin đầy đủ, hay hoàn toàn kiểm soát được hành động hay tài sản của mình không phải lúc nào cũng tốt. Mossadeq và những đứa con nhỏ của tôi là những ví dụ đã được nhắc tới. Nhưng chiến thuật tương tự như vậy cũng có thể được minh họa bởi việc một người đốt hết những cây cầu sau khi đi qua để thuyết phục đối thủ rằng không có gì có thể làm cho anh ta rút lui. Việc một đạo luật cổ của Anh coi việc *công nộp* cho bọn cướp biển ở ven biển là một trọng tội không nhất thiết là độc ác hay bất bình thường khi soi sáng bởi một lý thuyết về chiến lược. Điều lý thú là bản thân sự dân chủ trong chính trị lại dựa vào một hệ thống trao đổi thông tin đặc biệt ngăn cản việc truyền đạt các bằng chứng xác thực: hệ thống bỏ phiếu kín bắt buộc ngăn chặn cử tri bày tỏ họ bỏ phiếu cho bên nào. Bị tước bỏ khả năng bày tỏ việc bỏ phiếu cho ai, cử tri bị tước bỏ khả năng bị đe dọa. Không có cách nào chứng minh liệu mình có tuân theo sự đe dọa hay không, cử tri - và cả những người đe dọa họ - biết rằng mọi sự trừng phạt sẽ chẳng có liên quan gì đến cách bỏ phiếu của họ.

Nguyên tắc phổ biến cho rằng chúng ta nên chọn các nhà thương lượng giỏi để đại diện cho mình và trao cho họ sự linh hoạt và toàn quyền quyết định - một nguyên tắc thường do chính các nhà đàm phán nêu ra - hoàn toàn không phải là điều hiển nhiên như những người cổ vũ cho nó muốn nêu lên. Sức mạnh của một nhà thương lượng thường rõ ràng bị hạn chế do không có khả năng đưa ra nhân nhượng và đáp ứng các đòi hỏi của đối phương.<sup>8</sup> Tương tự như vậy, mặc dù sự khôn ngoan mách bảo rằng chúng ta nên để chừa một lối thoát khi đe dọa đối phương trong trường hợp có sự trả đũa gây tổn thất lẫn nhau, thì những lối thoát hiển hiện có thể làm cho sự đe dọa trở nên kém tin cậy hơn. Chính ý tưởng cho rằng việc cố tình loại bỏ một số lựa chọn có thể đem đến lợi thế chiến lược, hoặc thậm chí từ bỏ mọi sự kiểm soát đối với các hành động của mình trong tương lai và khiến mình phản ứng một cách máy móc, là một điều rất khó chấp nhận.

Nhiều ví dụ kể trên cho thấy có sự phủ định ở mức độ nhất định giá trị của kỹ năng, năng lực, tính hợp lý, kiến thức, sự kiểm soát, hay sự tự do lựa chọn. Tất cả những điều đó, về nguyên tắc, là đúng trong những hoàn cảnh nhất định. Tuy nhiên, việc tìm hiểu sự mới mẻ và logic đằng sau chúng sẽ dễ dàng hơn rất nhiều nếu chúng ta thiết lập thành bài toán, nghiên cứu nó ở mức trừu tượng, và nhận biết sự tương tự trong các tình huống khác khi mà ta dễ nắm bắt sự mới mẻ hơn.

---

<sup>8</sup> Việc quản lý viện trợ nước ngoài cho ta nhiều ví dụ. Chẳng hạn, xem "Viện trợ nước ngoài của Hoa Kỳ," T. C. Schelling, *World Politics* (tháng 7-1955), trang 614-615.

Một nguyên tắc khác trái ngược với ân tượng ban đầu có liên quan đến giá trị đạo đức tương đối của bom sạch và bom phóng xạ. Bernard Brodie đã chỉ ra rằng khi một người tính đến nhu cầu đặc biệt về sự răn đe, trái ngược với nhu cầu của một cuộc chiến tranh mà người đó dự tính sẽ tham gia, người đó có thể thấy sự hữu ích của bom siêu phóng xạ. Như nhận xét nêu ra trong Chương 10, kết luận như vậy không có gì lạ nếu chúng ta thừa nhận “cân bằng sức hủy diệt” (balance of terror) này đơn giản chỉ là một phiên bản hiện đại quy mô lớn của một thể chế cổ, sự trao đổi con tin.

Chúng ta có thể nhận thấy ở đây sự bất lợi đặc biệt đối với các sinh viên ngành quan hệ quốc tế trong thời văn minh hiện đại, ví dụ như nếu so sánh với Machiavelli hay người Trung Hoa cổ đại. Chúng ta thường có xu hướng nhận diện hòa bình, sự ổn định, và ngưng xung đột bằng những ý niệm như sự tin cậy, sự chân thành, và sự tôn trọng lẫn nhau. Chúng nào cách nhìn này khuyến khích niềm tin thực sự và sự tôn trọng thì đó còn là điều tốt. Nhưng ở những nơi mà sự tin cậy và chân thành không tồn tại và không thể được coi là đang tồn tại, thì khi đó chúng ta có thể học từ thế giới ngầm hay từ các chế độ chuyên chế xa xưa cách làm thế nào để thực thi các thỏa thuận nếu không có niềm tin và sự chân thành và không thể trông cậy vào sự can thiệp pháp lý khi có sự vi phạm hợp đồng. Thời cổ đại người ta trao đổi con tin, uống chung một ly rượu để chứng minh không có chất độc, gặp nhau tại các nơi công cộng để tránh bên này ám sát bên kia, và thậm chí còn có tình trao đổi gián điệp để tạo điều kiện cho việc trao đổi thông tin xác thực. Có lẽ một lý thuyết phát triển cao về chiến lược có thể đem lại ánh sáng về tính hiệu quả của

những biện pháp xa xưa đó, nêu ra những tình huống có thể áp dụng chúng, và tìm ra những phương thức tương tự cho thời hiện đại, dù có trái với khẩu vị của chúng ta, nhưng có thể vô cùng cần thiết cho việc điều chỉnh xung đột.

## 2

# MỘT TIẾU LUẬN VỀ MẶC CẢ

Chương này trình bày cách tiếp cận chiến thuật dùng phân tích sự mặc cả. Đối tượng nghiên cứu bao gồm cả sự mặc cả công khai và ngầm định trong đó các đối thủ theo dõi và lý giải hành vi của nhau, mỗi bên hiểu rõ rằng hành động của mình đang được đối phương lý giải và dự tính, mỗi bên hành động theo các dự đoán do mình lập nên. Trong lĩnh vực kinh tế, đối tượng nghiên cứu bao gồm các vấn đề thương lượng về tiền lương, thương lượng về thuế quan, cạnh tranh khi không có nhiều đối thủ cạnh tranh, giải quyết tranh chấp bên ngoài tòa án, người môi giới bất động sản và khách hàng của anh ta. Bên ngoài lĩnh vực kinh tế, đối tượng nghiên cứu trải rộng từ việc đe dọa trả đũa với quy mô lớn cho đến việc giành quyền ưu tiên của một chiếc taxi.

Mỗi quan tâm của chúng ta *không* phải là nghiên cứu bộ phận của lý thuyết mặc cả liên quan đến việc tìm những điều chỉnh để đôi bên cùng có lợi, điều có *ui* được gọi là khía cạnh “hiệu quả” của sự mặc cả. Ví dụ, *ui* một công ty bảo hiểm có thể vừa tiết kiệm được chi phí, vừa làm cho khách hàng vui

lòng hơn, bằng cách thanh toán tiền mặt thay vì sửa chữa xe cho khách hàng? Liệu một ông chủ có thể tiết kiệm được chi phí bằng cách tự tăng lương cho nhân viên nào đồng ý nhận phần lớn lương của mình bằng hàng hóa? Thay vào đó, chúng ta sẽ quan tâm đến cái có thể được gọi là khía cạnh "phân phối" của sự mặc cả: các tình huống trong đó sự mặc cả có lợi cho một bên đồng nghĩa với sự thiệt thòi cho bên kia. Khi doanh nghiệp cuối cùng được bán cho bên muốn mua, thì giá cả sẽ là bao nhiêu? Khi hai xe tải chở chất nổ gặp nhau trên một đoạn đường chỉ đủ rộng cho một xe, thì xe nào phải lùi lại?

Đó là những tình huống mà cuối cùng sẽ dẫn đến yêu tố mặc cả thuần túy - sự mặc cả trong đó mỗi bên bị chi phối chủ yếu bởi những dự tính của mình về những điều mà bên kia sẽ chấp nhận. Nhưng khi mỗi bên bị chi phối bởi các dự tính và đều biết rằng bên kia cũng như vậy, thì các kỳ vọng sẽ trộn lẫn vào nhau. Một cuộc mặc cả sẽ thành công khi một bên đưa ra nhân nhượng cuối cùng và đủ cho bên kia. Tại sao anh ta lại nhán nhượng? Bởi vì anh ta nghĩ rằng bên kia sẽ không nhán nhượng. "Tôi phải nhán nhượng vì anh ta sẽ không nhán nhượng. Anh ta không nhán nhượng vì anh ta nghĩ tôi sẽ nhán nhượng. Anh ta nghĩ tôi sẽ nhán nhượng vì anh ta nghĩ rằng tôi nghĩ anh ta nghĩ như vậy..." Có một khoảng kết cục nhất định mà tại mỗi điểm trong đó cả hai bên đều có lợi hơn so với việc không đạt được thỏa thuận gì cả. Việc thương lượng để đạt được một điểm trên khoảng đó có tính mặc cả thuần túy, vì người ta *thà* được ít đi còn hơn không đạt được gì cả, và vì người ta *luôn có thể* nhán nhượng, nếu sự lùi bước đó là cần thiết để đi đến thỏa thuận. Tuy

nhiên nếu cả hai bên đều biết các giới hạn của khoảng này, thì *bất cứ* kết cục nào cũng là một điểm mà từ đó ít nhất có một bên sẵn lòng nhân nhượng và bên kia biết điều đó! Ở đây không có điểm dừng.

Tuy nhiên, sẽ có một kết cục, và nếu chúng ta không tìm ra nó trong logic của tình huống, thì chúng ta có thể tìm ra nó trong các chiến thuật áp dụng. Mục tiêu của chương này là kêu gọi sự quan tâm đến một lớp chiến thuật quan trọng, một lớp đặc biệt thích hợp với logic của các tình huống bất định. Cốt lõi của các chiến thuật này là sự hy sinh quyền tự do lựa chọn một cách tự nguyện và không thể đảo ngược. Chúng dựa trên nghịch lý là khả năng giới hạn hành động của đối phương lại phụ thuộc vào khả năng ràng buộc bản thân, và trong mặc cả, điểm yếu thường là thế mạnh, sự tự do có thể là sự tự do đầu hàng, và việc đốt cháy những cây cầu sau lưng mình có thể đủ để vô hiệu hóa đối phương.

## SỨC MẠNH MẶC CẢ: KHẢ NĂNG RÀNG BUỘC BẢN THÂN

“Sức mạnh mặc cả,” “thế mạnh mặc cả,” “kỹ năng mặc cả” nói lên rằng lợi thế thuộc về kẻ có quyền lực, có thế mạnh, hay khôn khéo. Điều đó là tất nhiên, nếu như những phẩm chất này được định nghĩa chỉ để nói rằng các cuộc thương lượng thuộc về kẻ thắng. Tuy nhiên, nếu các thuật ngữ này hàm ý rằng ta sẽ có lợi thế nếu nhanh trí hay khôn khéo hơn trong tranh luận, hay có nguồn tài chính dồi dào hơn, có sức mạnh

cơ bắp hơn, tiềm lực quân sự lớn hơn, hay khả năng chịu đựng tổn thất nhiều hơn, thì thuật ngữ này sẽ phản tác dụng. Những phẩm chất này hoàn toàn không phải lúc nào cũng đem đến lợi thế trong các tình huống mặc cả. Chúng thường có tác dụng ngược lại.

Một nhà đàm phán tinh tế có thể sẽ thấy khó khăn để tỏ ra ngoan cố như một người thực sự ngoan cố. Nếu có một người đến gõ cửa và nói rằng anh ta sẽ tự đâm dao vào mình ngay tại ngưỡng cửa trừ khi cho anh ta 10 đô-la, thì anh ta sẽ dễ có được 10 đô-la hơn nếu như mắt anh ta đỏ ngầu như máu. Sự đe dọa tàn phá lẫn nhau không thể dùng để răn đe đối thủ là một kẻ dốt nát đến nỗi không hiểu được sự răn đe đó hoặc quá yếu thế để thực thi mong muốn của mình hơn là c Xia những người mà anh ta đại diện. Một chính phủ không thể kiểm soát được cán cân thanh toán, hay thu thuế, hay tập hợp sự đoàn kết chính trị để bảo vệ chính mình, có thể sẽ nhận được sự giúp đỡ mà lẽ ra sẽ bị từ chối nếu chính phủ đó có thể kiểm soát được các nguồn lực của mình. Và, như trong một ví dụ quen thuộc của lý thuyết kinh tế, "lãnh đạo giá" trong độc quyền nhóm có thể là một tước hiệu không lợi lộc gì mà các doanh nghiệp nhỏ lảng tránh, còn các công ty lớn lại coi là cần thiết.

Sức mạnh mặc cả còn được mô tả như là khả năng lừa dối và bịp bợm, tức là "khả năng đạt được cái giá tốt nhất cho mình và lừa người khác tin rằng đó là mức cao nhất mà bạn có thể đưa ra."<sup>1</sup> Lừa dối và bịp bợm chắc chắn luôn tồn tại,

---

<sup>1</sup> J. N. Morgan, "Độc Quyền Song Phương và Kết Cục Cảnh Tranh," tạp chí *Quarterly Journal of Economics*, 63:376 (tháng 8-1949).

nhưng có hai loại lừa dối. Một loại là nói trái với sự thật. Một người mua có thể nói dối về thu nhập thực của mình hay khai man số người trong gia đình của anh ta. Loại kia chỉ là thuần túy chiến thuật. Giả sử rằng mỗi bên đều biết mọi điều về đối thủ của mình, và mỗi bên cũng biết bên kia biết gì. Vậy thì các bên lừa dối được điều gì? Người mua có thể nói rằng, mặc dù anh ta thực sự có thể trả đến hai mươi và người bán cũng biết điều này, anh ta quyết tâm về mặt chiến thuật không nhích lên quá mươi sáu. Nếu như người bán chịu thua, thì có phải anh ta đã bị lừa gạt? Hay là anh ta bị sự thật khuất phục? Hay là người mua thực sự không biết sẽ làm gì nếu chiến thuật này thất bại? Nếu như người mua thực sự “cảm thấy” rất quyết tâm, và dựa quyết tâm này trên niềm tin rằng người bán sẽ phải chịu thua, và người bán làm đúng như vậy, thì người mua có thể sau đó nói rằng anh ta đã “không lừa dối.” Dù chuyện gì đã xảy ra thì cũng không thích hợp để cho rằng đây là sự lừa dối hay bịa bợm.

Muốn làm cho người khác tin vào điều gì đó thì người ta phải làm thế nào? Câu trả lời phụ thuộc chủ yếu vào câu hỏi xác minh, “Điều đó có đúng sự thật không?” Chứng minh tính chân thật của một điều có thật thì sẽ dễ hơn là chứng minh một điều sai sự thật. Để chứng minh tính chân thật về tình trạng sức khỏe của mình, chúng ta có thể nhờ một bác sĩ có uy tín. Để chứng minh tính chân thật về các chi phí hay thu nhập của mình, chúng ta có thể cho người khác xem sổ sách đã được kiểm toán bởi một công ty có uy tín hay Tổng cục Thuế. Nhưng để làm người khác tin một điều sai sự thật thì chúng ta không chắc đã có bằng chứng với sức thuyết phục như vậy.

Khi có ai đó muốn thuyết phục người khác rằng anh ta sẽ không trả giá cao hơn 16.000 đô-la cho một ngôi nhà mà theo anh ta nó thực sự đáng giá 20.000 đô-la, thì anh ta có thể làm gì để khai thác tính siêu tin cậy của sự thật so với việc đòi hỏi sai? Câu trả lời là: biến nó thành sự thật. Người mua làm thế nào để biến nó thành sự thật? Nếu anh ta thích ngôi nhà này vì nó gần chỗ kinh doanh, anh ta có thể dời đi doanh nghiệp của mình, rồi thuyết phục người bán nhà rằng ngôi nhà giờ đây đối với anh ta chỉ thực sự đáng giá 16.000 đô-la. Cách này có thể không lợi lộc gì, vì anh ta không có lợi so với trường hợp anh ta trả giá cao hơn.

Nhưng nếu giả sử người mua đánh cược với bên thứ ba, bằng một thỏa thuận hợp pháp và sẽ dứt khoát không thay đổi, được ghi nhận và chứng thực thích đáng, rằng anh ta sẽ không trả giá cho ngôi nhà quá 16.000 đô-la, nếu không sẽ mất 5.000 đô-la. Thế là người bán nhà chịu thua, và người mua chỉ cần đưa ra sự thật. Trừ phi người bán nổi giận và giữ lại ngôi nhà vì sự căm tức thuần túy, còn không thì tình hình đã bị thao túng gây bất lợi cho anh ta. Tình huống “khách quan” - động cơ thực sự của người mua - đã bị cố tình thay đổi hẳn và không thể đảo ngược được. Người bán có thể chấp nhận bán hoặc không bán nữa. Ví dụ này cho thấy rằng nếu người mua có thể chấp nhận một *cam kết* dứt khoát, mà người bán thấy rõ điều đó, thì anh ta có thể thu hẹp khoảng bất định về một điểm có lợi nhất đối với mình. Ví dụ này, qua tính giả định của nó, cũng gợi ý rằng chiến thuật đó có thể có hoặc có thể không tồn tại. Liệu người mua có thể tìm được một phương thức hữu hiệu để tự cam kết hay không có thể phụ

thuộc vào anh ta là người như thế nào, người bán là ai, họ sống ở đâu, và một số điều kiện pháp lý và thể chế (bao gồm, như trong ví dụ giả định của chúng ta, việc đánh cược có hợp pháp hay không).

Nếu như cả hai người này sống trong một nền văn hóa nơi mà “tôi xin thề” được công nhận phổ biến là có hiệu lực, thì điều mà người mua phải làm chỉ là tuyên bố rằng anh ta sẽ không trả quá 16.000 đô-la, sử dụng lời thề kèm hình phạt này, và anh ta được căn nhà - hoặc ít nhất anh ta thắng nếu người bán không bắt buộc anh ta bằng cách thết “19.000 đô-la, tôi xin thề đấy.” Nếu người mua là một đại diện được hội đồng quản trị ủy quyền mua nhà với giá 16.000 đô-la và không được quá một xu nào, và trong nhiều tháng tới hội đồng quản trị không thể họp lại, còn người mua không thể vượt quá thẩm quyền của mình, và nếu như tất cả những điều này được người bán biết rõ, thì người mua sẽ “thắng” - nếu, như đã nói, người bán không tự trói buộc vào một cam kết với mức giá 19.000 đô-la. Hoặc, nếu người mua có thể quyết rằng anh ta sẽ không trả quá 16.000 đô-la vì anh ta có thể bị mất mát lớn về uy tín cá nhân hay uy tín đàm phán nếu anh ta trả cao hơn, và nếu việc trả cao hơn đó chắc chắn sẽ lộ ra, và nếu người bán hiểu được tất cả những điều này, thì khi đó bản thân lời tuyên bố hùng hồn như vậy cũng đã thành cam kết. Tất nhiên, phương thức ở đây là từ bỏ sự linh hoạt một cách không cần thiết trừ phi nó quá rõ ràng và dễ hiểu đối với người bán.

Nhân đây cần nói thêm rằng, một số hình thức cam kết theo hợp đồng lại không có hiệu quả như người ta tưởng.

Trong ví dụ về trường hợp tự áp đặt hình phạt thông qua việc đánh cược, người bán vẫn có thể tìm bên thứ ba và gạ đưa một số tiền khiêm tốn để đổi lại việc người này giải thoát người mua khỏi việc đánh cược, điều dẫn đến đe dọa về việc bán ngôi nhà với giá 16.000 đô-la nếu việc giải thoát không diễn ra. Hiệu quả của việc đánh cược này - cũng giống như mọi cam kết theo hợp đồng khác - là làm thay đổi trong tâm và nhân sự của cuộc thương lượng, với hy vọng rằng bên thứ ba sẽ ít có khả năng tham gia cuộc thương lượng hoặc bớt động cơ để nhân nhượng. Nói cách khác, sự cam kết *theo hợp đồng* luôn dẫn đến việc gánh chịu “chi phí chuyển nhượng” có điều kiện, chứ không phải “chi phí thực”, và nếu tất cả các bên liên quan đều tham gia thương lượng thì khoảng bất định lại sẽ vẫn như cũ. Nhưng nếu bên thứ ba chỉ có mặt với chi phí đi lại lớn thì một cam kết thực sự không thể đảo ngược được có lẽ đã được chấp nhận. (Nếu việc đánh cược được thực hiện với nhiều người, thì “chi phí thực” để đưa họ đến cuộc thương lượng có thể là quá đắt.)<sup>2</sup>

---

<sup>2</sup> Có thể giải pháp “lý tưởng” cho vấn đề độc quyền song phương là như sau. Một trong hai đối thủ dịch chuyển đường chi phí biên của mình sao cho lợi nhuận chung bây giờ bằng không tại mức sản lượng mà lợi nhuận chung ban đầu có thể là tối đa. Anh ta làm điều này thông qua một thỏa thuận bán-cho thuê lại mà không thể hủy bỏ. Anh ta bán một hợp đồng tác quyền cho một bên thứ ba nào đó để thu một khoản cố định, phi bản quyền được gắn với mức sản lượng của anh ta sao cho chi phí chung vượt quá thu nhập chung tại tất cả các mức sản lượng khác. Giờ đây anh ta không thể sản xuất ở bất kỳ mức giá hay mức sản lượng nào ngoại trừ mức giá và mức sản lượng ở đó anh ta thu được toàn bộ lợi nhuận chung ban đầu. Bên thứ hai của độc quyền song phương thấy hợp đồng này, hiểu được tình huống, và chấp nhận lợi nhuận tối thiểu thực sự của mình. Người “chiến thắng” thực sự giành được toàn bộ lợi nhuận ban đầu nhờ việc bán tác quyền để lấy khoản cố định. Khoản lợi

Những phần đáng chú ý nhất trong chủ đề của chúng ta là xem xét các cam kết được hình thành hay không và bằng cách nào. Nhưng, chúng ta nên xem xét qua một mô hình không có các vấn đề của thực tế - một thế giới trong đó mọi cam kết tuyệt đối lúc nào cũng có. Ta hãy xem xét một nền văn hóa trong đó "tôi xin thề" được mọi người công nhận là ràng buộc tuyệt đối. Mọi đề nghị đi kèm với lời thề này đều là đề nghị cuối cùng, và mọi người chấp nhận như thế. Nếu mỗi bên biết được giá bảo lưu thực sự của bên kia, thì mục tiêu sẽ là trở thành người đầu tiên đưa ra một đề nghị kiên quyết. Sau đó, bên kia sẽ chịu trách nhiệm hoàn toàn về kết cục bằng cách chấp nhận hay không chấp nhận (và nếu anh ta chấp nhận). Vụ mặc cả kết thúc, và sự cam kết (tức là đề nghị được đưa ra đầu tiên) chiến thắng.

Hãy giả sử có khó khăn nào đó trong liên lạc. Các bên phải mặc cả với nhau qua thư từ. Cam kết có hiệu lực khi đã được ký tên nhưng bên kia vẫn chưa biết cho đến khi nhận được văn bản. Trong khi bên này viết một bức thư như thế thì bên kia có thể đã ký tên vào văn bản của mình, hoặc có thể còn chưa ký trước khi văn bản của bên kia gửi đến. Khi đó việc

---

nhuận này không ảnh hưởng động cơ của anh ta vì nó không phụ thuộc vào sản lượng. Bên thứ ba trả hết một lần (trừ đi một khoản khuyến khích nhỏ) vì anh ta biết rằng bên thứ hai sẽ chấp nhận và do đó anh ta sẽ nhận được khoản phí tác quyền trong trường hợp đó. Vấn đề ở đây là người mua tác quyền không được liên hệ với "bên thua", nếu không thì bên thua có thể sẽ buộc anh ta từ bỏ tác quyền bằng cách đe doa sẽ không chịu thỏa thuận, qua đó khôi phục lại tình trạng chi phí biên ban đầu. Nhưng chúng ta có thể hình dung ra được sự phát triển của các tổ chức chuyên mua bán tác quyền, mà sự thành công tối hậu của họ phụ thuộc vào danh tiếng không bao giờ thương lượng lại, và do đó động cơ của họ sẽ không bị lung lạc trong bất cứ một thương lượng nào.

mua bán không xảy ra. Cả hai bên bị trói buộc vào các vị trí không tương thích với nhau. Giờ đây mỗi bên đều nhận ra tình trạng bế tắc này và tính đến khả năng bên kia đã hoặc sẽ ký tên vào cam kết của họ.

Tính bất đối xứng trong liên lạc có thể có lợi nhiều cho bên nào không có khả năng nhận được thông điệp (và ~~được~~ biết như vậy), bởi vì anh ta là người không bị ngăn cản có cam kết riêng khi chưa nhận được văn bản của bên kia. (Mặt khác, nếu bên không liên lạc được giả vờ không biết điều đó, thì bên kia cũng có thể bị ngăn cản có cam kết của mình bởi sự lo ngại đối với cam kết không chủ tâm của đối phương.) Nếu việc cam kết không chỉ bằng lời nói mà còn dựa vào các hình thức hoặc nghi lễ khác, thì việc không biết đến các nghi lễ cam kết của bên kia có thể là một lợi thế, nếu việc không hay biết đó được chấp nhận hoàn toàn, vì điều đó làm cho bên kia hiểu rằng chỉ có sự kiềm chế của anh ta mới có thể tránh được bế tắc.

Giả sử chỉ có một bộ phận dân chúng thuộc nhóm tín ngưỡng coi “tôi xin thề” là (hoặc được tin là) ràng buộc tuyệt đối. Nếu mọi người đều biết (và được người khác biết là họ biết) về nhóm tín ngưỡng của mọi người khác, thì những người theo nhóm thề thốt này sẽ có lợi thế. Họ có thể tự ràng buộc mình, còn nhóm kia thì không. Nếu người mua nói “16.000 đô-la, tôi thề đấy” thì sự trả giá đó của anh ta là cuối cùng. Còn nếu người bán nói “19.000 đô-la” thì anh ta chỉ (và được biết là chỉ) đang “mặc cả” mà thôi.

Nếu mỗi bên đều không biết giá bảo lưu thực sự của bên kia thì sẽ có giai đoạn ban đầu trong đó các bên tìm cách phát hiện giá bảo lưu của đối phương và để lộ giá không đúng của mình, giống như trong các cuộc mặc cả thông thường. Nhưng

quá trình phát hiện và để lộ này nhanh chóng được kết hợp với quá trình hình thành và phát hiện các cam kết. Vì các lý do thiết thực, những cam kết này làm thay đổi vĩnh viễn các giá bảo lưu “thực sự.” Nếu một bên tin và bên kia không tin vào nghi lễ ràng buộc thì bên không tin sẽ theo đuổi thủ thuật mặc cả “thông thường” bằng cách *đòi hỏi* được giá bảo lưu, trong khi bên tin đang *tính toán giá* của mình.

Phản thảo luận ở trên nhằm nêu ra cả hai mặt hợp lý và logic của sự tự cam kết. Một số ví dụ có thể cho thấy sự đúng đắn của chiến thuật này, mặc dù ít khi người ta có thể phân biệt rõ ràng việc sử dụng một chiến thuật hữu hình là theo logic có ý thức, theo trực giác, hay không chủ ý. Thứ nhất, không phải là hiếm thấy trường hợp các cán bộ công đoàn gây sự kích động và lòng quyết tâm của các đoàn viên công đoàn khi hoặc trước khi diễn ra thương lượng về tiền công. Nếu công đoàn đòi mức lương 2 đô-la và dự kiến lãnh đạo nhà máy sẽ đáp lại với mức 1,60 đô-la, thì người ta sẽ thuyết phục các công đoàn viên rằng không những ban lãnh đạo nhà máy chịu trả 2 đô-la mà còn coi các nhà thương lượng - chính họ - là kém năng lực nếu không thương lượng được mức giá 2 đô-la. Mục đích của việc này - hay nói đúng hơn là mục đích hợp lý mà sự phân tích của chúng ta nêu ra - là làm rõ cho ban lãnh đạo nhà máy rằng các nhà thương lượng không thể chấp nhận mức thấp hơn 2 đô-la cho dù họ có muốn “i nữa bởi vì họ không còn kiểm soát các công đoàn viên được nữa hoặc bởi vì họ sẽ mất chức nếu có làm như vậy. Nói cách khác, các nhà thương lượng tự hạn chế phạm vi thẩm quyền của mình và đối đầu với ban lãnh đạo bằng sự đe dọa đình công mà bản thân công đoàn không thể ngăn cản, mặc dù

chính hành động của công đoàn đã làm mất đi khả năng ngăn cản đình công của mình.

Điều tương tự như vậy diễn ra khi Chính phủ Hoa Kỳ đàm phán với chính phủ các nước khác về những vấn đề như viện trợ nước ngoài sẽ dùng vào đâu, hay vấn đề cắt giảm thuế quan. Nếu cơ quan hành pháp này được tự do thương lượng để đi đến thỏa thuận tốt nhất dựa vào khả năng của mình, thì có thể nó không giữ được lập trường nào cả và kết thúc bằng sự nhân nhượng trên những điểm còn chưa ngã ngũ bởi vì đối phương biết hoặc tin tưởng một cách ngoan cố rằng Hoa Kỳ thà nhân nhượng còn hơn là hủy bỏ các cuộc thương lượng. Nhưng, nếu cơ quan hành pháp này đàm phán dưới thẩm quyền của cơ quan lập pháp, với lập trường bị ràng buộc bởi luật pháp, và rõ ràng là Quốc hội không thể được triệu tập để họp thay đổi luật trong khoảng thời gian cần thiết, thì cơ quan hành pháp sẽ có lập trường vững vàng rất rõ ràng đối với các đối tác thương lượng.

Khi những người đại diện cho quốc gia đi đàm phán tại các cuộc thương lượng quốc tế với sự hiểu biết rằng có một khoảng rộng rãi các thỏa thuận tiềm năng trong đó kết cục sẽ phụ thuộc vào việc mặc cả, họ thường đưa ra lập trường mặc cả bằng các tuyên bố công khai, những tuyên bố đã được tính toán để khơi lên luồng dư luận không cho phép nhân nhượng. Nếu luồng dư luận với tính ràng buộc có thể được nuôi dưỡng và thể hiện rõ cho đối phương, thì bằng cách đó, lập trường ban đầu sẽ có thể được nhìn nhận như là lập trường “tối hậu.”

Những ví dụ nói trên có một số đặc điểm chung. Thứ nhất, chúng không chỉ phụ thuộc một cách rõ ràng vào việc thực

hiện một cam kết mà còn vào việc truyền đạt cam kết đó một cách thuyết phục cho đối phương biết. Thứ hai, không phải hoàn toàn dễ dàng để xác lập một cam kết, và cũng không hoàn toàn rõ ràng đối với các bên liên quan là cam kết đó vững chắc đến đâu. Thứ ba, cả hai bên đều có thể có hành động tương tự. Thứ tư, mặc dù cả hai bên đều có khả năng đưa ra cam kết, khả năng đó không hoàn toàn phải là như nhau. Khả năng của một chính phủ dân chủ bị ràng buộc bởi dư luận có thể khác với khả năng của một chính phủ độc tài khi thực hiện cam kết. Thứ năm, tất cả đều chịu rủi ro trong việc đưa ra một lập trường bất di bất dịch nằm ngoài khả năng nhân nhượng của bên kia, và do đó có thể dẫn đến bế tắc và tan vỡ.

## *NHỮNG ĐẶC ĐIỂM THỂ CHẾ VÀ CẤU TRÚC TRONG ĐÀM PHÁN*

Một số đặc điểm thể chế và cấu trúc của các tình huống mặc cả có thể làm cho việc sử dụng chiến thuật cam kết dễ dàng hay khó khăn, hay làm cho bên này dễ áp dụng hơn bên kia, hay ảnh hưởng đến khả năng đưa ra cam kết đồng thời hoặc bế tắc.

*Sử Dụng Đại Diện Mặc Cả.* Việc sử dụng đại diện mặc cả có ảnh hưởng đến khả năng cam kết, ít nhất là theo hai cách. Thứ nhất, người đại diện có thể được giao những chỉ thị khó hoặc không thể thay đổi, mà những chỉ thị như vậy (và sự

cứng nhắc của nó) được đối phương thấy rõ. Nguyên lý này được áp dụng trong việc phân biệt giữa cơ quan lập pháp và cơ quan hành pháp, hay giữa ban điều hành và hội đồng quản trị, cũng như đối với đề nghị do sứ giả mang đi khi quá trình mặc cả có giới hạn về thời gian và người chủ đặt khoảng cách đáng kể giữa mình với sứ giả làm cho việc liên lạc rõ ràng là không thể kịp thời hạn.

Thứ hai, một người “đại diện” có thể được xem xét đầy đủ với tư cách của một người chủ, với một cấu trúc lợi ích riêng khác với cấu trúc lợi ích của người chủ. Cách làm này thường thấy trong ngành bảo hiểm ô-tô. Một công dân, khi giải quyết tranh chấp ngoài hệ thống tòa án, không thể đe dọa kiện ra tòa có hiệu quả bằng công ty bảo hiểm vì công ty bảo hiểm có động cơ rõ ràng hơn trong việc đưa ra những đe dọa như vậy để giữ uy tín của mình trong các vụ tai nạn sau này.<sup>3</sup>

*Bí Mật và Công Khai.* Một phương thức cam kết hiệu nghiêm, mà đôi khi là duy nhất, là lời hứa danh dự. Nếu những người đại diện cho quốc gia có thể bị buộc tội nhanh chóng với từng nhượng bộ nhỏ nhặt, họ sẽ tránh xa sự nhân nhượng. Nếu một công đoàn trong các xí nghiệp khác có thể thu xếp nhượng bộ một cách quá lỗ liễu, thì nó đặt uy tín mặc cả của mình vào

---

<sup>3</sup> Giải pháp chính thức cho vấn đề quyền ưu tiên trong giao thông xe có thể là bên thắng sẽ là người nào đó được bảo hiểm đầy đủ và công khai đối với mọi tai nạn xảy ra, vì anh ta không có động cơ để tránh tai nạn, nên người khác phải nhường và ho biết điều đó. (Họ không thể đối phó lại bằng cách giống như vậy. Không công ty nào bảo hiểm cho anh ta khi mà người đầu tiên đã được bảo hiểm). Nghiêm trọng hơn, việc gom lại quy định công giữa các công đoàn sẽ làm giảm đi động cơ tránh định công của mỗi công đoàn. Như giải pháp độc quyền tay đôi đã nói ở trên, việc chuyển nhượng lợi ích cho bên thứ ba sẽ làm thay đổi cấu trúc động cơ của mỗi bên.

thế nguy hiểm và do đó lộ rõ là không có khả năng dàn xếp nghiêm túc. (Thế nguy hiểm dễ mắc tương tự như vậy là cơ sở cho kiểu phòng thủ được áp dụng rộng rãi, “Nếu tôi làm điều đó cho anh, tôi sẽ phải làm điều đó cho mọi người khác.”) Nhưng để cam kết theo cách này thì cần phải có sự công khai hóa. Cả lời đề nghị đầu tiên lẫn kết quả cuối cùng phải được công bố, và nếu có bí mật xung quanh hai điều này, hoặc nếu kết cục vốn dĩ không thể nhìn thấy, thì phương thức này sẽ không áp dụng được. Nếu một bên có “dư luận công khai” còn bên kia không có, thì bên không có sẽ tìm cách xóa bỏ bất lợi của mình bằng cách gạt đi dư luận công khai đó, hoặc nếu cả hai bên đều không muốn đi đến bế tắc khi cùng sử dụng chiến thuật này, họ có thể tìm cách thực thi thỏa thuận một cách bí mật.

*Đàm Phán Tương Hỗ.* Nếu một tổ chức công đoàn tham gia đồng thời, hay sớm tham gia, vào nhiều cuộc đàm phán trong khi ban lãnh đạo công ty không có nhà máy nào khác hoặc không phải đối phó với các tổ chức công đoàn khác, thì ban lãnh đạo công ty không thể dùng uy tín đàm phán của mình một cách thuyết phục, trong khi tổ chức công đoàn lại có thể. Lợi thế sẽ thuộc về bên nào có thể chỉ ra một cách thuyết phục lập trường của mình sẽ bị tổn hại trong các cuộc đàm phán khác nếu như chịu nhân nhượng\* trong cuộc thương lượng này. (“Giá trị uy tín” của cuộc mặc cả này có thể ít liên quan đến kết cục thương lượng hơn là sự vững vàng mà đề nghị ban đầu sẽ dựa vào.) Việc chống lại chiến thuật này đòi hỏi nhiều yếu tố, trong đó có cả việc diễn đạt sai lập trường của đối phương và nỗ lực để đạt kết quả không tương ứng với những lập trường ban đầu. Nếu các chủ đề thương lượng có

thể được mở rộng trong quá trình đàm phán, hay con số tiền công được thay thế bằng các phúc lợi mà không thể quy về tiền công, thì “sự rút lui” sẽ đến bên nào đã tự cam kết, và sự tồn tại của “sự rút lui” này sẽ làm yếu đi bản thân sự cam kết, bất lợi cho bên có cam kết.

*Đàm Phán Liên Tục.* Một trường hợp đặc biệt của đàm phán tương hỗ xuất hiện khi cả hai bên sẽ phải thương lượng thêm về các chủ đề khác, ngay cùng một lúc hay trong tương lai. Logic của trường hợp này tinh tế hơn. Để thuyết phục đối phương rằng mình không có khả năng nhượng bộ, người ta thường nói đại loại như sau, “Nếu tôi nhân nhượng anh ở đây, anh có thể xem lại đánh giá của anh về tôi trong các cuộc thương lượng khác; để bảo vệ uy tín của mình với anh, tôi phải giữ vững lập trường.” Bên thứ hai cũng đồng thời là “bên thứ ba” mà uy tín thương lượng của một bên có thể được hướng tới. Tình huống này xuất hiện trong trường hợp có sự đe dọa kháng cự trước sự gây hấn địa phương. Bên đe dọa đạt được sự cam kết của mình, và qua đó nâng cao độ tin cậy cho đe dọa của mình, không phải bằng cách đe dọa đến kết quả nó sẽ có được do thực hiện lời đe dọa mà bằng cách chỉ ra giá trị lâu dài của một đe dọa đã được thực hiện trong việc nâng cao độ tin cậy của các đe dọa trong tương lai.

*Chương Trình Nghị Sự Hạn Chế.* Khi có hai mục tiêu đặt ra cho cuộc đàm phán, thì quyết định đàm phán đồng thời hay tách riêng tại các diễn đàn khác nhau hay tại thời các điểm khác nhau sẽ có ảnh hưởng đến kết quả đàm phán, đặc biệt khi có một mối đe dọa cưỡng ép tiềm ẩn mà chỉ có thể lợi dụng được nếu nó được gắn với một tình huống mặc cả bình thường và hợp pháp hơn. Việc chống lại đe dọa

cưỡng ép phụ thuộc vào việc từ chối đàm phán, không có mặt, hay không có khả năng đàm phán. Nhưng nếu đối tượng đe dọa cưỡng ép có thể được đưa vào chương trình nghị sự cùng với chủ đề còn lại, thì mối đe dọa tiềm ẩn đó sẽ trở nên có hiệu lực.

Lấy mặc cả về thuế quan làm ví dụ. Nếu các mức thuế quan hai chiều đối với pho-mát và ô-tô được đưa ra đàm phán, một bên có thể làm thay đổi kết quả đàm phán bằng cách đe dọa trừng phạt thuần túy thông qua việc thay đổi một thuế quan khác. Nhưng nếu đại diện thương lượng của bên bị đe dọa bị giới hạn bởi chương trình nghị sự pho-mát và ô-tô, và không được ủy quyền bàn đến các hàng hóa khác, hoặc có những quy định rõ ràng cấm không được đề cập đến các thuế quan khác khi chưa giải quyết xong vấn đề thuế quan của pho-mát và ô-tô, thì vũ khí cưỡng ép này dành phải chờ cơ hội khác. Nếu mối đe dọa có thể được đưa ra bàn thương lượng thuộc loại không thể công bố được, thì chính sự công khai sẽ ngăn cản sự truyền đạt hiệu quả của mối đe dọa đó.

*Khả Năng Đèn Bù.* Như Fellner đã chỉ ra, thỏa thuận có thể phụ thuộc vào một số phương thức phân phối lợi ích.<sup>4</sup> Ví dụ, nếu hai nhà độc quyền đối (duopolist) chia thị trường sao cho tối đa hóa tổng lợi nhuận của họ, thì sự phân chia lợi nhuận ban đầu sẽ được xác định. Mọi sự phân chia lại lợi nhuận đòi hỏi một bên có khả năng đèn bù cho bên kia. Nếu việc đèn bù là bằng chứng của sự cấu kết phi pháp, hoặc nếu các cổ đông hiểu sai động cơ đèn bù, hoặc nếu hai bên

---

<sup>4</sup> W. Fellner, *Cạnh Tranh Số Ít* (Competition Among the Few) (New York, 1949), trang 34-35, 191-197, 231-232, 234.

không tin cậy lẫn nhau thì lợi nhuận *chung* phải giảm so với mức tối ưu để bảo đảm cơ cấu lợi nhuận ban đầu của hai bên gần sát với sự phân chia lợi nhuận mà hai bên đã thỏa thuận.

Khi một thỏa thuận phải đạt được về một điều nào đó vốn là hành động của một bên, bất kỳ sự phân chia chi phí nào cũng sẽ phụ thuộc vào việc đèn bù. "Chương trình nghị sự" có vai trò đặc biệt quan trọng trong các trường hợp này, vì phương thức đèn bù chủ yếu là sự nhân nhượng về một đối tượng khác. Nếu hai cuộc đàm phán diễn ra đồng thời có thể được gắn kết với nhau, thì các bên đã có sẵn phương thức đèn bù. Nhưng nếu chúng bị tách riêng, thì mỗi cuộc thương lượng là một đối tượng không thể phân chia được.

Một bên có thể có lợi thế nếu tách riêng các cuộc mặc cả, còn bên kia sẽ có lợi thế nếu gắn cuộc mặc cả này vào cuộc mặc cả thứ hai. Nếu có hai dự án, mỗi dự án có chi phí là ba, và mỗi dự án có giá trị là hai cho bên A, và bốn cho bên B, và cả hai đều là dự án do "một bên" thực hiện, và nếu việc đèn bù không thể thực hiện được xét về thể chế, thì bên B sẽ buộc phải trả toàn bộ chi phí của mỗi dự án chừng nào mà hai dự án được tách riêng. Anh ta không thể dùng cách đe dọa không thực hiện vì bên A không có động lực để tự mình thực hiện cả hai dự án. Nhưng nếu bên B có thể liên kết các dự án lại với nhau, hứa hẹn sẽ thực hiện một dự án trong khi bên A thực hiện dự án kia, và có thể đe dọa có hiệu quả sẽ từ bỏ cả hai dự án trừ phi A thực hiện một dự án, A sẽ được lựa chọn phương án có thu nhập là bốn và chi phí là ba, còn bên B giảm được một nửa chi phí.

Một sự hạn chế quan trọng của các vấn đề kinh tế, với tư

cách là hình mẫu cho các tình huống mặc cả, nằm ở chỗ chúng thường liên quan một cách không cân đối đến các đối tượng phân chia được và các hoạt động có thể đèn bù. Nếu một rãnh thoát nước ở phía sau một ngôi nhà bảo vệ cả hai ngôi nhà, và nếu chi phí của nó là 1.000 đô-la và giá trị mang lại cho mỗi chủ nhà là 800 đô-la, thì chẳng nhà nào muốn xây nó riêng cả, mà chúng ta nghĩ rằng họ sẽ kết hợp với nhau triển khai dự án đem lại giá trị 1.600 đô-la cho họ. Nhưng nếu phải mất 10 giờ một tuần để trông người làm, và niỗi bên cho rằng nó xứng đáng 8 giờ lao động để thuê một người giám sát nhưng một trong hai người phải làm việc này, thì còn xa hai bên mới chắc chắn đạt được thỏa thuận theo đó một bên bỏ ra 10 giờ làm toàn bộ công việc và bên kia sẽ trả tiền hoặc bỏ ra 5 giờ để dọn vườn giúp đổi tác. Khi có hai chiếc ô-tô gặp nhau trên một con đường hẹp, thì sự bế tắc xảy ra sau đó trở nên nghiêm trọng hơn vì không ai có thói quen đấu giá trả cho quyền được ưu tiên. Những bế tắc ở quốc hội xuất hiện khi việc trao đổi lá phiếu không thực hiện được. Các biện pháp đòi hỏi sự nhất trí thường chỉ có thể bắt đầu nếu một số người gắn kết lại với nhau.<sup>5</sup>

*Cơ Chế Đàm Phán.* Một số đặc điểm khác đáng được nhắc đến, mặc dù chúng ta sẽ không đi sâu vào kết quả của chúng. Có một hình phạt nào đối với việc truyền đạt thông tin sai không? Có hình phạt nào cho việc chào bán giả dối không, tức là khi một người đưa ra chào bán và sau đó rút lại lời chào bán được chấp nhận? Có hình phạt nào trong việc thuê một

---

<sup>5</sup> Việc đưa thêm điều khoản về Saar trong "Các Hiệp định Paris" chấm dứt sự chiếm đóng Tây Đức có thể phản ánh tốt nguyên tắc này hoặc nguyên tắc ở đoạn trước.

người đại diện giả vờ quan tâm và đưa ra mời chào giả dối chỉ nhằm thử lập trường của bên kia hay không? Liệu ta có thể nhận biết được tất cả các bên liên quan hay không? Việc mặc cả có thời hạn hay không? Cuộc mặc cả có cấu trúc đặc biệt của một cuộc đấu giá thông thường, một cuộc đấu giá kiểu Hà Lan (nêu giá thấp xuống từng bước chớ đến khi có người mua - người dịch chú thích), một hệ thống đấu thầu kín, hay một giàn xếp chính thức? Liệu có *giải pháp nguyên trạng* (status quo) sao cho việc không tham gia thương lượng được sẽ đem lại thắng lợi *nguyên trạng* cho bên đàm phán nào muôn như vậy hơn? Liệu có thể đàm phán lại trong trường hợp gấp bế tắc? Cái giá phải trả khi gấp bế tắc là gì? Liệu việc tuân thủ thỏa thuận có giám sát được không? Nói chung, các phương tiện trao đổi thông tin là gì, và có phương tiện nào có thể dễ dàng bị bên này hay bên kia làm cho mất hiệu lực hay không? Nếu có nhiều đề mục được đưa ra đàm phán, thì các đề mục đó được đàm phán trong một cuộc thương lượng toàn diện, một cách riêng lẻ theo một thứ tự cụ thể sao cho từng đề mục được kết thúc rồi mới đến đề mục sau, hay tiến hành đồng thời thông qua các đại diện khác nhau hay theo các quy tắc khác nhau?

Tầm quan trọng của nhiều câu hỏi cấu trúc này trở nên rõ ràng khi ta nghĩ về kỹ thuật nghị viện. Quy tắc chỉ cho phép chủ tọa phủ quyết cả gói một dự luật phân bổ ngân sách, hay đòi hỏi mỗi sự bổ sung sửa đổi phải được bỏ phiếu trước khi bỏ phiếu bán thân dự luật ban đầu, hay một hệ thống ưu tiên áp dụng riêng cho từng loại kiến nghị, làm thay đổi đáng kể các động lực đằng sau từng hoạt động. Một người bị gây sức

ép buộc phải lựa chọn giải pháp tốt thứ nhì sẽ loại trừ được khả năng bị tổn thương nếu anh ta có thể bỏ phiếu trước đó để ngăn chặn khả năng này xảy ra, nhờ đó anh ta chỉ còn cần phải chọn giữa lựa chọn thứ nhất và thứ ba mà anh ta thích đến mức không thể bị đe dọa.

*Các Nguyên Tắc và Tiền Lệ.* Để có sức thuyết phục, các cam kết thường cần đến chất lượng hơn là số lượng, và phải dựa trên cơ sở hợp lý nào đó. Khó có thể quan niệm về một sự cam kết chắc chắn với cái giá 2,07 ½ đô-la. Tại sao không phải là 2,02 ¼ đô-la? Thước đo giá trị quá liên tục đến nỗi không cung cấp những điểm dừng tốt, ngoại trừ tại những điểm tròn số như 2,00 đô-la. Thế nhưng một cam kết dựa vào nguyên tắc “chia sẻ lợi nhuận,” “giá sinh hoạt tăng lên,” hay bất kỳ một căn cứ nào khác cho việc tính toán dẫn đến con số 2,07 ½ đô-la đều có thể là một cơ sở vững chắc cho một cam kết. Hơn nữa, người ta có thể tạo nên một yếu tố trong cam kết bằng cách đặt các nguyên tắc và tiền lệ trước nguy cơ bị vi phạm. Nếu như trong quá khứ một người đã từng thành công trong việc duy trì nguyên tắc nào đó, chẳng hạn như không công nhận chính phủ được dựng nên bằng bạo lực, và cương quyết đưa ra yêu cầu dựa theo nguyên tắc đó trong cuộc đàm phán hiện tại, thì anh ta không chỉ dùng tiền lệ đó làm căn cứ cho đòi hỏi của mình mà còn gây mạo hiểm do chính nguyên tắc đó. Khi đã cam kết theo nguyên tắc đó, anh ta muốn thuyết phục đối phương rằng anh ta thà chấp nhận bế tắc hơn là nhượng bộ và từ bỏ nguyên tắc.

*Lý Lê Ngụy Biện.* Nếu ai đó đi đến một điểm nêu nhân nhượng, thì anh ta cần phải cân nhắc hai hậu quả: việc nhân

nhưng sẽ đưa anh ta đến chỗ gần hơn với lập trường của đối phương, và đồng thời nó sẽ có tác động đến sự đánh giá của đối phương về sự giữ vững lập trường của anh ta. Nhân nhưng có thể được suy diễn không chỉ là sự đầu hàng, mà còn là dấu hiệu lừa dối trong cam kết được đưa ra trước đó, và làm cho đối phương nghi ngờ xem còn sự giả tạo nào nữa hay không. Do vậy, cần phải có "lý do" cho việc nhường nhịn đối thủ, tốt nhất là diễn giải lại một cách hợp lý cam kết ban đầu, có sức thuyết phục cả với đối phương.

Điều đáng chú ý hơn là việc sử dụng lý lẽ ngụy biện để giải phóng một đối thủ khỏi cam kết. Nếu ai đó có thể chứng minh cho đối phương thấy rằng anh ta không định cam kết, hay anh ta đã tính toán sai cam kết của mình, thì trong thực tế đã vô hiệu hóa hay làm thay đổi cam kết của đối phương. Hoặc ai đó có thể làm cho cam kết của đối phương rối tung lên khiến cho các cử tri, hay cấp trên, hay cử tọa của đối phương không thể nhận diện chính xác sự tuân thủ cam kết - cho thấy "năng suất lao động" chỉ là mơ hồ, hay "các đóng góp tương xứng" có nhiều cách hiểu khác nhau - thì có thể vô hiệu hóa hay hạ thấp giá trị trong cam kết của đối phương. Trong các trường hợp này, đối phương sẽ bất lợi khi cam kết của anh ta bị lý lẽ bác bỏ hoàn toàn. Nhưng khi đối phương quyết tâm đưa ra sự nhân nhưng vừa phải, thì ta có thể giúp đối phương bằng cách chứng minh rằng anh ta có thể đưa ra một nhưng bộ vừa phải nhất quán với lập trường trước đây, và rằng nếu anh ta làm như vậy thì không có cơ sở nào để tin rằng nó gây ra nghi ngờ về những nguyên tắc ban đầu. Nói cách khác, ta cần suy xét một cách nghiêm túc để tránh việc hưởng lợi quá

nhiều từ nhân nhượng của đối phương, nếu không đối phương sẽ không nhượng bộ.<sup>6</sup>

## SỰ ĐE DỌA

Khi ai đó đe dọa sẽ đánh lại nếu bị tấn công, hay dọa sẽ phá giá nếu đối thủ cạnh tranh hạ giá trước, thì lời đe dọa đó chẳng khác gì một sự truyền đạt về động cơ thúc đẩy của chính người đó nhằm khắc sâu vào đầu đối phương về các hậu quả tất yếu do hành động của họ gây nên. Nếu việc đe dọa đó mang lại kết quả thì cả hai bên đều có lợi.

Nhưng người ta còn cần đến những thứ khác ngoài việc truyền đạt khi đe dọa ngăn cản một hành động nào đó mà anh ta không có động cơ thực hiện mà chỉ nhằm mục đích răn

---

<sup>6</sup> Trong nhiều bài toán sách giáo khoa, như là độc quyền song phương giữa các công ty, hai điểm đầu của khoảng mắc cả là các điểm lợi nhuận bằng không của bên này hay bên kia, và việc đạt thỏa thuận với lợi nhuận tối thiểu cho một bên thì chẳng hơn gì việc không đạt được thỏa thuận. Nhưng, ngoài các tình huống mua và bán ra, thường có các giới hạn chung về khoảng các kết cục có thể chấp nhận được, và kết cục ít có lợi nhất mà một bên có thể chấp nhận có thể còn tốt hơn nhiều so với tình trạng bế tắc. Trong các trường hợp này mục đích chủ yếu của một bên có thể là ngăn ngừa cam kết lầm lẫn của bên kia. Nếu việc chứng minh sự thật dễ hơn chứng minh lập trường giả đối, thì nên đưa ra một lập trường bảo thủ ngay từ đầu, cũng như nếu rút lui khỏi trường "quá cao" ban đầu sẽ làm mất lòng tin đối với mọi cố gắng sau đó để nói sự thật. Thực vậy, dù không ai muốn bị trừng phạt do hành vi của mình, sự tồn tại của hình phạt cưỡng chế đối với sự giả đối có thể sẽ có ích. Ví dụ như nếu ai đó có thể chứng minh thực trạng chi phí hay thu nhập bằng cách đưa ra tờ hoàn thuế thu nhập, thì hình phạt về sự khai man có thể nâng cao giá trị của bằng chứng này.

đe thông qua lời hứa gây hại cho nhau. Việc đe dọa trả đũa quy mô lớn trước những lần chiếm nhỏ có đặc điểm này, cũng như dọa sẽ đâm vào chiếc xe nào không chịu nhường đường, hay kêu gọi một cuộc đình công tốn kém nếu tiền công không được tăng thêm một vài xu. Tính đặc thù của sự đe dọa này là ở chỗ người đe dọa không có động cơ thực hiện lời đe dọa của mình cả trước hay sau khi có hành động của đối phương. Anh ta sẽ có động cơ ràng buộc mình vào việc thực hiện lời đe dọa, nếu anh ta nghĩ rằng sự đe dọa sẽ thành công, bởi vì chính sự đe dọa chứ không phải việc thực hiện nó sẽ giúp đạt mục đích, và việc thực hiện là không cần thiết nếu đe dọa thành công. Việc sẽ thực hiện lời đe dọa càng chắc chắn bao nhiêu, thì khả năng phải thực hiện lời đe dọa càng ít bấy nhiêu. Nhưng tính hiệu quả của sự đe dọa phụ thuộc vào mức độ cả tin của bên kia, và sự đe dọa sẽ không có hiệu quả trừ phi người đe dọa có thể sắp xếp lại hoặc trình bày các động cơ của mình để tỏ rõ là anh ta sẽ, *về mặt hậu nghiệm* (*ex post*), có động cơ để thực hiện lời đe dọa.<sup>7</sup>

---

<sup>7</sup> Cần nói thêm rằng, sự đe dọa nhằm răn đe có một số đặc điểm đáng chú ý về mặt định lượng, phản ánh tính bất đối xứng thường thấy giữa phản thường và trường phạt. Chẳng hạn, không nhất thiết sự đe dọa hứa hẹn sẽ gây ra nhiều tổn thất hơn cho bên bị đe dọa so với bên đưa ra đe dọa. Việc đe dọa dùng chiếc xe mới đâm vào chiếc xe cũ có thể thành công nếu được tin như vậy, hay tiến hành kiện cáo tốn kém vì một số tổn thất nhỏ, hay khởi xướng một cuộc chiến tranh giá cả. Tương tự như vậy nếu xét về khía cạnh sức mạnh răn đe, thì không có sự đe dọa nào là “quá lớn”. Nếu sự đe dọa đó đủ sức mạnh để thành công thì đằng nào nó cũng sẽ không được thực hiện. Một sự đe dọa chỉ “quá lớn” nếu quy mô của nó ảnh hưởng đến độ tin cậy. Sự tàn phá bằng bom nguyên tử đối với một tội nhẹ, hay bỏ tù vì đỗ xe quá giờ sẽ là thừa nhưng không quá tổn kém trừ phi người bị đe dọa không tin và xem thường lời đe dọa đó.

Ngay cả trường hợp độc quyền song phương “thuần túy” cũng có tính chất này nếu việc mặc cả được tiến hành bởi các đại diện hay các nhân viên mà phần thưởng cho họ phụ thuộc nhiều vào việc có đạt được thỏa thuận hay không hơn là các điều khoản của thỏa thuận có lợi như thế nào.

Chúng ta quay trở lại vấn đề cam kết. Làm thế nào để một người có thể tự mình cam kết sẽ có hành động, mà trên thực tế anh ta không muốn thực hiện hành động đó khi sự kiện xảy ra, sao cho sự cam kết của mình có thể răn đe đối phương? Tất nhiên ai đó có thể dùng lời lừa gạt để thuyết phục người khác tin rằng các chi phí hay tổn thất đối với người đưa ra đe dọa là không đáng kể hoặc bằng âm. Thú vị hơn nữa, người đưa ra đe dọa có thể giả vờ hiểu sai rằng tổn thất đối với anh ta là nhỏ, và do đó vẫn tiếp tục đưa ra và thực hiện lời đe dọa. Hoặc có thể anh ta giả vờ có động cơ trả thù mạnh đến nỗi bất chấp khả năng gây tổn thất cho chính mình. Tuy nhiên trường hợp này có lẽ dễ xảy ra nhất đối với người muốn trả thù thực sự. Nếu không, anh ta sẽ phải tìm ra một cách để tự ràng buộc mình vào lời đe dọa.

Người ta có thể đặt uy tín của mình vào việc thực hiện lời đe dọa, theo một cách gây ấn tượng đối với người bị đe dọa. Thậm chí ai đó còn có thể đặt uy tín của mình vào chính người bị đe dọa, với cơ sở rằng chi phí và tổn thất là xứng đáng để dạy một bài học cho kẻ coi thường lời đe dọa. Hoặc người ta có thể tìm cách đưa ra sự cam kết pháp lý, như là thông qua việc ký kết hợp đồng với bên thứ ba.<sup>8</sup> Hoặc nếu ai

<sup>8</sup> Các hiệp định phòng thủ chung giữa các nước mạnh và nước yếu có thể được nhìn nhận tốt nhất theo cách này, tức là, chúng được thực hiện không phải để làm yên lòng các nước nhỏ hay để đổi lấy sự đền đáp, mà là một phương thức từ bỏ sự tự do lựa chọn gây khó xử.

đó có thể giao toàn bộ công việc cho người đại diện mà mức lương (hay uy tín kinh doanh) của người này phụ thuộc vào việc thực hiện sự đe dọa nhưng không phải chịu trách nhiệm về những chi phí khác, thì người đó có thể làm thay đổi động cơ của đối phương.

Vấn đề cam kết được minh họa rất rõ bởi học thuyết pháp lý “cơ hội rõ ràng cuối cùng” (last clear chance), tức là trong các sự kiện dẫn đến một tai nạn, có một thời điểm mà tai nạn trở nên không thể tránh khỏi vì những hành động trước đó, và rằng khả năng ngăn ngừa tai nạn của hai bên có thể không kết thúc cùng một thời điểm. Trong mặc cả, sự cam kết là một phương thức giành cơ hội rõ ràng cuối cùng quyết định kết quả cho bên kia theo cách mà anh ta hoàn toàn hiểu rõ. Nó khuyến khích các bước tiếp theo, sau khi đã gây ảnh hưởng đến động cơ, buộc đối phương phải lựa chọn có lợi cho mình. Nếu một người lái xe tăng tốc đến mức anh ta không thể dừng lại, và người lái xe khác nhận ra điều này, thì người sau phải nhường đường. Một điều luật đưa thêm vào cuối kỳ họp dành cho Tổng thống cơ hội rõ ràng cuối cùng để thông qua một dự luật. Học thuyết này giúp hiểu rõ một số trường hợp, trong đó “sức mạnh” mặc cả lại nằm trong các điểm được cho là yếu theo các tiêu chuẩn khác. Khi một người – hay một nước – đánh mất sức mạnh để tự cứu mình, hay để tránh tổn thất chung, thì bên kia không còn cách nào khác ngoài việc phải gánh chịu chi phí hay trách nhiệm. “Sự thâm thùng cưỡng ép” (coercive deficiency) là thuật ngữ mà Arthur Smithies dùng để mô tả chiến thuật cố tình làm cạn kiệt các khoản ngân sách được phân bổ hàng năm rất sớm nhằm làm cho nhu cầu tăng

ngân sách trở nên khẩn cấp không thể cưỡng lại được.<sup>9</sup>

Một chiến thuật có liên quan là dẫn dắt để đi đến một *nguyên trạng* mà người ta chỉ có thể thoát khỏi hiện trạng đó bằng một hành động công khai, một hành động sẽ gây ra tổn thất cho đôi bên bởi vì bên dẫn dắt đã xóa bỏ khả năng rút lui. Nếu một người công khai mang theo chất nổ bên mình, theo một cách rõ ràng là sẽ không tránh khỏi hủy hoại đối với bản thân và bắt cứ một người tấn công nào, thì anh ta có thể ngăn chặn được sự tấn công hơn nhiều so với việc nếu anh ta còn một ít khả năng kiểm soát lượng nổ đó. Nếu ai đó đưa một lực lượng quân đội tượng trưng vào chỗ không còn đường thoát thì sẽ làm tăng mức độ cam kết kháng cự tổng lực. Walter Lippmann đã từng lấy ví dụ dùng cửa sổ gắn kính dày để bảo vệ cửa hàng bán đồ trang sức: ai cũng có thể đập vỡ nó một cách khá dễ dàng, nhưng không thể tránh gây ra náo động.

Các kỹ thuật tương tự cũng có thể được bên bị đe dọa sử dụng. Tất nhiên, cách phòng thủ tốt nhất đối với anh ta là hành động trước khi lời đe dọa được đưa ra. Trong trường hợp đó động cơ hay cam kết trả đũa sẽ không còn nữa. Nếu anh ta không thể có hành động nhanh, anh ta có thể cam kết sẽ có hành động đó. Nếu người bị đe dọa đã cam kết rồi, thì người đe dọa sẽ không thể rắn đe bằng đe dọa được nữa, anh ta chỉ còn cách làm cho những hậu quả tai hại cho cả hai bên

---

<sup>9</sup> A. Smithies, *Quá Trình Phân Bổ Ngân Sách tại Hoa Kỳ* (The Budgetary Process in the United States) (New York, 1955), trang 40, 56. Một giải pháp là giới hạn chặt chẽ quá trình phân bổ. Ngoài ra, xem T. C. Schelling, "Viện Trợ Nước Ngoài của Mỹ," *World Politics*, 7:609-625 (tháng 7-1955), nói về nguyên tắc, tương tự trong việc phân bổ viện trợ nước ngoài.

chắc chắn sẽ xảy ra như anh ta dọa.<sup>10</sup> Nếu người bị đe dọa có thể thu xếp để chia sẻ rủi ro với người khác trước khi lời đe dọa được đưa ra (như đã được nêu trong ví dụ nói về giải pháp bảo hiểm cho vấn đề quyền được nhường đường), thì anh ta có thể cho bên kia thấy không còn sợ bị đe dọa nữa để làm nản lòng người đe dọa. Hoặc nếu bằng cách nào đó anh ta có thể thay đổi động cơ hay không để lộ đúng động cơ của mình để tỏ ra anh ta vẫn có lợi dù lời đe dọa được thực hiện (hay chỉ có lẽ chỉ cần anh ta nghĩ như vậy) thì người muốn đe dọa có thể từ bỏ ý định vì việc đó chỉ tốn kém và vô ích. Hoặc nếu ai đó có thể giả vờ không hiểu hay bướng bỉnh đến nỗi không để ý đến lời đe dọa, anh ta có thể vô hiệu hóa được sự đe dọa đó. Tốt nhất có thể là *thực sự* không hiểu, ngoan cố, hay đơn giản là không tin vào sự đe dọa, vì điều đó có thể sẽ dễ thuyết phục người đe dọa tiềm tàng hơn. Nhưng tất nhiên nếu cách này không thuyết phục được người đe dọa và anh ta vẫn cam kết thực hiện đe dọa, thì cả hai bên đều sẽ thiệt hại. Cuối cùng, cả sự đe dọa lẫn cam kết đều phải được truyền đạt. Nếu người bị đe dọa không có mặt để nhận thông điệp,

---

<sup>10</sup> Hệ thống cung cấp cho cảnh sát giao thông vé phạt có đánh số và không tẩy xóa được cho phép cảnh sát tránh được sự đe dọa của lái xe, bằng cách ghi số xe trước khi nói chuyện với lái xe. Một số xe tải trưng các bảng với nội dung “Hệ thống báo động và khóa không thuộc sự kiểm soát của lái xe.” Hệ thống khóa hầm két của ngân hàng được lập trình theo thời gian cũng có mục đích tương tự, cũng như việc bỏ phiếu bí mật bắt buộc trong bầu cử. Hay tương tự như vậy trong việc bắt đầu cuộc xâm chiếm bằng một lực lượng tiên phong nhỏ, tuy quá nhỏ và quá sớm để đạt mục tiêu, nhưng mang quá nhiều “thể diện” của quân đội khiến không thể rút lui: một lực lượng lớn hơn sau đó sẽ sẵn sàng tham chiến mà không sợ khơi ra một đe dọa ngán cản thuần túy. Tại nhiều trường đại học, người ta bảo vệ cán bộ giảng dạy bằng quy tắc không cho phép giảng viên được thay đổi điểm số môn học một khi đã được ghi vào sổ.

hay có thể phá hủy các kênh liên lạc, mặc dù anh ta công khai làm việc đó nhằm tránh bị đe dọa, anh ta có thể ngăn chặn được lời đe dọa.<sup>11</sup> Nhưng việc chứng tỏ sự không tin hay ngoan cố phải được thực hiện trước khi đối phương đưa ra đe dọa, tức là trước khi có sự cam kết của đối phương, chứ không phải chỉ trước khi sự đe dọa được thực hiện. Sẽ chẳng có ích gì nếu tỏ ra không tin, hay lánh mặt một khi lời đe dọa đã được chuyển đến.

Trong các tình huống có đe dọa, cũng như trong các tình huống mặc cả bình thường, các cam kết không phải đã rõ ràng hoàn toàn. Mỗi bên không thể ước lượng chính xác các chi phí và giá trị cho đối phương của hai hành động liên quan đến sự đe dọa. Quá trình cam kết có thể là một quá trình tăng dần, các cam kết thu được tính vững chắc thông qua một chuỗi các hành động. Việc truyền đạt thường không phải hoàn toàn không thể thực hiện được cũng không phải hoàn toàn đáng tin cậy. Trong khi một số bằng chứng của sự cam kết có thể được truyền đạt trực tiếp, những bằng chứng khác đòi hỏi phải được truyền đi qua báo chí hay tin đồn, hoặc được thể hiện qua hành động. Trong các trường hợp này, khả năng xấu vì cả hai hành động xảy ra, do sự cam kết đồng thời, sẽ tăng lên. Hơn nữa, việc nhận thấy khả năng có cam kết đồng thời

---

<sup>11</sup> Kẻ tống tiền không thể tống tiền nếu không tìm được nạn nhân. Kẻ bắt cóc không thể lấy được tiền chuộc nếu không liên lạc được với bạn bè hay thân nhân người bị bắt cóc. Do vậy, dù là một gợi ý không thực tế, nhưng nếu có luật đòi hỏi cách ly ngay lập tức tất cả bạn bè và thân nhân của người bị bắt cóc một khi có vụ bắt cóc xảy ra, thì luật như vậy có thể làm cho triển vọng lấy được tiền chuộc trở nên mờ mịt. Sự thay đổi luân phiên nhân viên bảo vệ và cảnh sát, hay phân công nhiệm vụ theo từng cặp một cách ngẫu nhiên, không chỉ giảm cơ hội nhận hối lộ mà còn bảo vệ họ khỏi mối đe dọa.

bản thân nó lại là sự ngăn chặn thực hiện cam kết.<sup>12</sup>

Trong trường hợp đe dọa được đưa ra và hành động không bị ngăn chặn, thì sẽ có một giai đoạn thứ hai trước khi việc thực hiện đe dọa trong đó cả hai bên đều có lợi ích trong việc gỡ bỏ cam kết. Mục đích của đe dọa không còn nữa, giá trị của sự răn đe là bằng không, và chỉ sự cam kết là còn tồn tại để thúc đẩy việc thực hiện. Tất nhiên, đặc điểm này cũng tương tự như sự bế tắc trong mặc cả thông thường, sự bế tắc xuất phát từ việc cả hai bên cam kết những lập trường không tương thích, hoặc một bên có lập trường làm lẩn đến nỗi bên kia không chấp nhận được. Nếu có một cơ hội để gỡ bỏ cam kết, cả hai bên sẽ có lợi ích để làm như vậy. Làm thế nào để gỡ bỏ là tùy thuộc vào lợi ích của hai bên xa nhau đến đâu, vì các cách gỡ bỏ khác nhau sẽ dẫn đến các kết quả khác nhau. Hơn nữa, "gỡ bỏ" không có nghĩa là bỏ qua cam kết bất chấp việc mất uy tín. Nếu cam kết về uy tín là thực, "gỡ bỏ" có nghĩa là tách sự đe dọa ra khỏi uy tín, có thể là uy tín của người đó đối với bản thân người bị đe dọa. Do vậy, đây là tình huống tê nhị và mong manh trong đó mặc dù cả hai bên đều có lợi khi gỡ bỏ cam kết, họ không thể hợp tác với nhau để làm điều đó.

---

<sup>12</sup> Một đặc điểm thể chế đáng kể là không có một phương pháp đơn giản phổ quát nào để các cá nhân hay quốc gia đưa ra những cam kết như chúng ta đang thảo luận. Có rất nhiều cách để họ thử nghiệm nhưng hầu hết là không rõ ràng, không chắc chắn, hay thỉnh thoảng mới có. Trong xã hội "tôi xin thế" trình bày ở trên, lý thuyết mặc cả chỉ còn là chiến lược trò chơi và cơ chế truyền đạt. Nhưng trong phần lớn thế giới đương thời, đề tài này chủ yếu có tính kinh nghiệm và thể chế về ai có thể cam kết được, làm cách nào, và với mức độ đảm bảo hiểu biết của đối phương như thế nào.

Chúng ta cần hết sức thận trọng trong việc xác định sự đe dọa, cả đối với hành động bị đe dọa chống lại lẫn hành động đáp trả được đưa ra để đe dọa. Khó khăn xuất hiện từ thực tế, như vừa nói, là một khi hành động trước đã được thực hiện thì sẽ không còn động cơ để thực hiện hành động sau. Tính tin cậy của đe dọa trước khi có hành động phụ thuộc vào việc bên bị đe dọa thấy rõ đến mức nào bên đưa ra đe dọa không thể biện luận cho sự rút khỏi cam kết của mình một khi nó không đạt được mục đích. Bất cứ kẽ hở nào của bên đe dọa, nếu để cho bên bị đe dọa nhìn thấy, sẽ làm suy yếu cam kết hữu hình và do đó làm giảm đi tính tin cậy của lời đe dọa. (Một ví dụ là cách xử lý không rõ ràng đối với quần đảo Kim Môn trong Nghị quyết và Hiệp định về Đài Loan).

Do vậy, để tăng tối đa sự tin cậy, điều cốt lõi là dành phạm vi phán xét hay lựa chọn khi thực hiện lời đe dọa càng hẹp càng tốt. Nếu một người cam đoan sẽ trừng phạt một hành vi khi nó vượt quá những giới hạn nào đó, nhưng những giới hạn này lại không được xác định một cách cẩn thận và khách quan, thì bên bị đe dọa sẽ nhận ra rằng khi đến lúc phải quyết định xem lời đe dọa đó có được thực thi hay không, lợi ích của anh ta và của bên đe dọa sẽ trùng hợp nhau trong nỗ lực tránh những hậu quả bất lợi cho cả hai bên.

Để làm cho lời đe dọa được chính xác với các yếu tố rõ ràng đối với cả bên bị đe dọa và các bên thứ ba mà phản ứng của họ đối với sự việc này đều có ảnh hưởng đến hai đối thủ, thì cần thiết phải đưa ra một số yếu tố tùy nghi. Việc đe dọa phải có những hành động công khai thay vì chỉ là ý định. Nó phải gắn liền với những hành động hữu hình, chứ không phải vô hình. Sự đe dọa có thể phải gắn với một số hành động phụ

không có ảnh hưởng gì đến bên đe dọa. Chẳng hạn, nó có thể đưa ra sự trừng phạt đối với việc mang theo vũ khí thay vì sử dụng vũ khí, đối với hành vi đáng nghi ngờ thay vì hành vi sai trái đã lộ rõ, đối với hành vi gần phạm tội thay vì bản hành vi phạm tội. Và cuối cùng, hành động trừng phạt phải là hành động mà hậu quả hay ảnh hưởng của nó được nhận biết một cách rõ ràng.<sup>13</sup>

Để có thể dùng uy tín đảm bảo cho lời đe dọa thì phải có tính liên tục giữa hiện tại và những vấn đề sẽ nảy sinh về sau. Sự cần thiết về tính liên tục này gợi ý một phương thức giúp cho lời đe dọa ban đầu có hiệu quả hơn. Nếu nó có thể được chia nhỏ thành một chuỗi các đe dọa nhỏ liên tiếp, thì sẽ có cơ hội để chứng minh qua một vài hành động vi phạm đầu tiên rằng sự đe dọa sẽ được thực hiện ở những lần sau. Thậm chí, một vài đe dọa đầu tiên trở nên đáng tin cậy hơn, vì động cơ thực hiện chúng với tư cách một “bài học” còn rõ ràng hơn.

Có lẽ nguyên tắc này phù hợp nhất cho những hành động vốn dĩ là vấn đề về mức độ. Trong các chương trình viện trợ nước ngoài, hành động chấm dứt viện trợ công khai có thể là điều đau đớn rõ ràng cho cả hai bên đến nỗi bên nhận viện trợ không nghĩ là có thể xảy ra. Nhưng nếu việc sử dụng sai từng khoản viện trợ sẽ đi kèm với một cắt giảm nhỏ trong viện

---

<sup>13</sup> Trong năm 1950, Cơ quan Hợp tác Kinh tế tuyên bố ý định thường những nước nào trong phạm vi Kế hoạch Marshall tuân thủ các chính sách lành mạnh đặc biệt, và trừng phạt những nước không làm như vậy, thông qua việc phân bổ các khoản viện trợ lớn hơn hay nhỏ hơn. Nhưng do các thước đo chuẩn không được xác định, và do việc xác định chúng cuối cùng dựa trên xét phán đoán chứ không phải tính toán, nên về sau không ai có cách nào biết được trong thực tế trường hợp nào được tăng thêm hay cắt giảm, và kế hoạch này gấp nhiều trở ngại do thiếu tính khả thi.

trợ, việc bỏ bên tiếp nhận rơi vào tình trạng vô vọng hay phá vỡ quan hệ ngoại giao sẽ không còn được coi là quá lớn lao nữa, thì ý định thực hiện nó sẽ đáng tin cậy hơn. Hoặc nếu lúc đầu nó không được tin cậy như vậy, thì một vài bài học sẽ có sức thuyết phục mà không gây ra tổn thất lớn.<sup>14</sup>

Tất nhiên bên đe dọa có thể không có khả năng chia hành động của mình thành nhiều bước khác nhau. (Cả hành động bị răn đe lẫn sự trừng phạt đều phải phân chia được.) Nhưng ít ra nguyên tắc này cho thấy sự thiếu khôn ngoan của việc vạch ra các hành động gây hấn hoặc vi phạm ở một mức độ hay số lượng tối hạn được coi như không thể chịu đựng được. Khi một hành động bị răn đe vốn dĩ là một chuỗi các bước mà hậu quả tích lũy của chúng mới là điều quan trọng, một sự đe dọa về các phần tăng dần có thể đáng tin cậy hơn so với sự đe dọa được thực hiện toàn bộ ngay lập tức hoặc không thực hiện gì cả khi đã đến một điểm nào đó. Thậm chí không thể xác định được một "điểm tối hạn" đủ rõ ràng để có sức thuyết phục.

Để có thể phân chia được các hành động đe dọa, bản thân các hành động này cần phải được điều chỉnh. Các phần hành động không thể chia nhỏ được có thể cần phải để ra ngoài. Những hành động phụ đi theo sự kiện, mặc dù bản thân chúng không có giá trị, lại có thể là đối tượng gắn vào đe dọa làm cho nó có hiệu quả hơn. Ví dụ, một số hành động phụ chỉ là sự chuẩn bị cho hành động chính, và bản thân nó không

<sup>14</sup> Có lẽ yêu cầu thường thấy đối với việc thanh toán dần các khoản vay qua nhiều khoảng thời gian ngắn, thay vì trả cùng một lúc vào cuối thời hạn vay, phản ánh một nguyên tắc tương tự, cũng như thông lệ kiểm tra thường xuyên trong một khóa cấp đại học để tránh việc thất bại của sinh viên hoàn toàn dựa trên một quyết định chấm thi sau khi khóa học đã kết thúc.

có hại gì, nhưng chúng có thể dễ dàng phân chia theo thời gian và do đó là đối tượng có ích cho sự đe dọa. Một người định đá vào con chó cần phải bị răn đe bởi sự trừng phạt khiêm tốn cho mỗi bước gần lại con chó, mặc dù bản thân cự ly gần của anh ta không có ý nghĩa gì.

Tương tự như việc chia nhỏ sự đe dọa thành một chuỗi là việc mở đầu sự đe dọa bằng một hành động trừng phạt có mức nặng nề tăng dần theo thời gian. Khi mà sự đe dọa chết chóc bằng bạo lực không làm cho người ta sợ hãi, thì việc cắt nguồn cung cấp lương thực có thể dẫn đến sự thu phục. Vì mục đích đạo đức hay quan hệ công chúng, biện pháp này có thể đem đến cho đối phương "cơ hội rõ ràng cuối cùng," mà sự kết liễu của anh ta bị gán cho tính bướng bỉnh nếu sự đe dọa thất bại. Nhưng trong mọi trường hợp người đe dọa thường rút lại hành động thái quá khi vẫn còn sớm và có mức độ nhỏ, thay vì để nó phát triển thành vật cản tối hậu, đáng sợ, và dễ thấy đối với quyết tâm của đối phương. Và nếu chỉ có bên chịu đựng là người duy nhất biết được, qua từng thời điểm, họ đã tiến gần đến thảm họa đến mức nào, thì hành động của anh ta là cơ hội rõ ràng cuối cùng với đúng nghĩa của nó. Hơn nữa, người đe dọa có thể bị bối rối bởi sự sụp đổ của đối thủ của mình chứ không phải bởi sự khó chịu của anh ta, và do đó biện pháp này có thể chuyển hóa một sự đe dọa nguy hiểm ngay một lúc thành sự đe dọa liên tục ít tổn kém hơn. Việc đe dọa trực xuất cưỡng bách người thuê nhà thì không hiệu quả bằng việc chỉ cần cắt điện nước.<sup>15</sup>

---

<sup>15</sup> Đây được xem là chiến thuật tránh được bùng nổ và khiêu khích các lực lượng của de Gaulle rút khỏi một tỉnh mà họ đã chiếm đóng ở phía bắc nước Ý vào tháng 6-1945, sau khi họ tuyên bố rằng mọi nỗ lực của đồng

Phương pháp chia nhỏ cũng có thể áp dụng cho bên bị đe dọa. Nếu anh ta không thể loại trừ mối đe dọa bằng cách có hành động tổng thể ngay lập tức, thì anh ta có thể nhanh chóng thực hiện một hành động nhỏ ban đầu qua đó khẳng định quyết tâm thực hiện đến cùng. Hoặc, nếu hành động của anh ta có thể chia nhỏ được trong khi sự trả đũa của người đe dọa chỉ có thể thực hiện được với quy mô kinh tế lớn, việc thực hiện hàng loạt những phần nhỏ đó có thể ngăn chặn bên đe dọa có hành động quá đáng kích thích sự phản ứng của anh ta.

### SỰ HỨA HẸN

Trong số những đặc quyền hợp pháp của công ty, có hai đặc quyền được nêu trong các sách giáo khoa là quyền đi kiện và “quyền” bị kiện. Ai muốn mình bị kiện chứ! Nhưng quyền bị kiện lại cho người ta khả năng để đưa ra hứa hẹn: để vay tiền, để tham gia giao kèo, để làm ăn với ai đó là người có thể bị tổn thất. Nếu kiện tụng phát sinh, “quyền” này có vẻ là nghĩa vụ khi nhìn lại, còn trước đó nó lại là điều kiện tiên quyết để làm ăn với nhau.

Nói tóm lại, quyền bị kiện cáo là sức mạnh để chấp nhận một sự cam kết. Trong các cam kết đã được thảo luận cho đến

---

minh nhằm đánh bật họ đi sẽ được coi là hành động thù địch. Xem Harry S. Truman, *Year of Decisions* (New York, 1955), trang 239-242 và Winston S. Churchill, *Triumph and Tragedy*, tập VI trong bộ *The Second World War* (Boston, 1953), trang 566-568.

nay, điều cốt lõi là đối phương của người nào đó (hay là “đối tác” tùy chúng ta gọi) không có khả năng giải thoát anh ta ra khỏi sự cam kết, mà thực sự ra đối với bên thứ ba nào đó sự cam kết có thể là thực sự hay tưởng tượng. Sự hứa hẹn là một cam kết đưa ra cho bên thứ hai trong cuộc mặc cả và là điều đòi hỏi phải có mỗi khi hành động cuối cùng của một bên hay của mỗi bên nằm ngoài sự kiểm soát của bên kia. Nó là điều đòi hỏi phải có mỗi khi bản thỏa thuận có yếu tố thúc đẩy sự lừa gạt.<sup>16</sup>

Sự cần thiết có những hứa hẹn này không phải là điều ngẫu nhiên. Bản thân nó có tầm quan trọng về mặt thể chế. Không phải luôn dễ dàng đưa ra một lời hứa hẹn có sức thuyết phục và tự ràng buộc. Cả bên bắt cóc muốn thả con tin, và con tin, đều rất cố gắng để tìm cách ràng buộc con tin sẽ không tố cáo kẻ bắt cóc, nhưng không tìm ra cách nào. Nếu nạn nhân đã có một hành động mà việc tiết lộ nó có thể dẫn đến việc tống tiền, anh ta có thể thú nhận về hành động đó. Nếu không anh ta có thể cam kết một lời hứa trước bên tống tiền, để tạo ra sự ràng buộc đảm bảo việc giữ anh ta im lặng. Nhưng các tình huống cực đoan này minh họa cho thấy việc đưa ra một lời hứa là khó khăn cũng như là quan trọng như thế nào. Nếu luật pháp không cưỡng chế các thỏa thuận về giá cả, hoặc nếu công đoàn không thể tự mình cam kết không có đình công, hoặc nếu một chủ thầu không có tài sản để đền bù

---

<sup>16</sup> Sự đe dọa có thể được xem như một lời hứa nếu lời đe dọa chỉ được hỗ trợ bởi uy tín của người đó đối với đối phương. Nhưng nó không phải là lời hứa mà bên thứ hai có thể đơn phương xóa bỏ cho người đe dọa, vì anh ta không thể tách rời ước đoán tương lai của mình về người đe dọa khỏi hành động của người này.

tổn thất khi anh ta thua kiện, và luật pháp không bắt giam các con nợ, hoặc nếu không có “khán giả” để ai đó có thể đặt uy tín của mình vào, thì không thể đi đến thỏa thuận mặc cả được, hoặc ít ra cuộc mặc cả này sẽ kết thúc theo cách khác.

Mặc cả còn có thể có liên quan đến hệ thống “khuyến khích” cũng như sự phân chia lợi ích. Các nhà độc quyền nhóm (oligopolist) có thể vận động hành lang cho một đạo luật về “thương mại bình đẳng”, hay trao đổi các cổ phiếu với nhau. Một thỏa thuận không xâm phạm thị trường của nhau có thể đòi hỏi phải có một thỏa thuận về thiết kế lại sản phẩm làm cho nó không thích hợp đối với thị trường của người khác. Hai quốc gia muốn có thỏa thuận không sử dụng một hòn đảo vào mục đích quân sự có thể phải phá hủy sự có ích của bản thân hòn đảo đó. (Trong thực tế, một “cam kết của bên thứ ba” cần phải có khi “sự cam kết của bên thứ hai” không thể tạo ra được.)<sup>17</sup>

Việc thực hiện lời hứa không phải lúc nào cũng quan sát được. Nếu một người bán lá phiếu của mình trong một cuộc bỏ phiếu kín, hay một chính phủ đồng ý sẽ đưa kiến nghị về một hành động ra quốc hội, hay một nhân viên hứa sẽ không lấy cắp hàng trong kho, hay một giáo viên hứa sẽ không đưa quan điểm chính trị của mình vào lớp học, hay một nước cam kết sẽ thúc đẩy xuất khẩu “càng nhiều càng tốt,” thì không có cách nào bảo đảm quan sát được hay đo lường được việc tuân thủ. Kết quả có thể quan sát được phụ thuộc vào nhiều yếu tố, trong đó chỉ có một yếu tố nằm trong thỏa thuận. Do đó việc mặc cả cần phải được thể hiện dưới dạng nào có thể quan

---

<sup>17</sup> Ở thời kỳ trước, người ta trao đổi các con tin cho nhau.

sát được, mặc dù cái có thể quan sát được này không phải là mục tiêu của cuộc mặc cả. Người ta có thể phải trả tiền cho một cử tri được đút lót do thắng cử, chứ không phải do cử tri đó bỏ phiếu như thế nào, trả hoa hồng cho nhân viên bán hàng theo doanh số, chứ không phải theo kỹ năng và nỗ lực bán hàng, trao thưởng cho cảnh sát dựa vào số liệu thống kê tội phạm chứ không phải dựa vào sự chăm chú làm việc, hay là trừng phạt tất cả nhân viên do lỗi của một người. Và, ở nơi mà việc thực thi công việc có mức độ khác nhau, sự mặc cả cần phải xác định rõ các giới hạn phân biệt giữa thực thi và không thực thi, một sự mảnh mai nhất định tại kho hàng được coi là bằng chứng của sự ăn cắp, mức tăng lên xác định trong xuất khẩu được coi là một nỗ lực “thỏa đáng”, những mẫu cụ thể nào từ công việc được lấy làm đại diện cho toàn bộ công việc.<sup>18</sup>

Chiến thuật chia nhỏ có thể áp dụng cho sự hứa hẹn cũng như sự đe dọa. Điều làm cho nhiều thỏa thuận được thực thi chỉ là việc nhận ra rằng các cơ hội thỏa thuận tương lai sẽ bị loại bỏ nếu sự tin cậy lẫn nhau không được tạo ra hay duy trì, và giá trị của nó vượt trội cái thu được nhất thời từ việc lừa gạt trong tình huống hiện tại. Mỗi bên cần phải tin tưởng rằng bên kia sẽ không làm hỏng các cơ hội tương lai bằng cách phá hoại niềm tin ngay từ đầu. Niềm tin này không phải lúc nào cũng tồn tại, và một trong những mục đích của các cuộc mặc

---

<sup>18</sup> Việc không có khả năng đưa ra lời hứa được đảm bảo, giống như việc không có khả năng thực hiện một hành động được yêu cầu, có thể bảo vệ một người khỏi sự đe dọa cưỡng ép. Việc bỏ phiếu kín bắt buộc là sự phiền muộn đối với cử tri nào muốn bán lá phiếu của mình, nhưng lại là sự bảo vệ đối với người lo sợ sự ép buộc.

cả từng phần là nuôi cấy những kỳ vọng chung cần thiết. Không bên nào sẵn lòng tin tưởng sự thận trọng của bên kia (hay tin tưởng sự tự tin của bên kia về sự thận trọng của mình, và cứ như vậy) về một vấn đề lớn. Nhưng, nếu nhiều cuộc mặc cả sơ bộ đạt được thỏa thuận ở quy mô nhỏ, thì mỗi bên có thể sẵn sàng mạo hiểm đầu tư một ít để tạo ra truyền thống tin tưởng. Mục đích là nhằm làm cho mỗi bên chứng tỏ rằng mình coi trọng sự cần thiết có sự tin tưởng và biết rằng bên kia cũng muốn như vậy. Do đó, nếu một vấn đề lớn cần phải được thương lượng, thì hai bên cần thiết tìm ra và tiến hành đàm phán một vài vấn đề nhỏ để "tập dợt," nhằm xác lập sự tin tưởng cần thiết trong nhận thức mỗi bên về giá trị lâu dài của thiện chí.

Ngay cả khi không có sự lặp lại trong tương lai, người ta vẫn có thể tạo ra sự liên tục bằng cách chia vấn đề mặc cả thành nhiều phần nhỏ liên tiếp. Nếu mỗi bên đồng ý gửi cho Hội Chữ thập đỏ một triệu đô-la với điều kiện bên kia cũng làm như vậy, thì mỗi bên có thể nảy ra ý muốn lừa gạt sau khi bên kia đã đóng góp, và dự đoán của bên này về khả năng lừa gạt của bên kia sẽ ngăn cản thỏa thuận. Nhưng nếu khoản đóng góp được chia thành nhiều phần nhỏ liên tiếp, thì mỗi bên có thể thử thiện chí của đối phương với một cái giá nhỏ. Hơn nữa, do mỗi bên có thể bám đuổi bên kia cho đến khi kết thúc, không bên nào phải mạo hiểm ngoài phần đóng góp nhỏ từng lúc. Cuối cùng, sự thay đổi trong bản thân cơ cấu khêu khích này sẽ loại bỏ rủi ro cho lần đóng góp đầu. Giá trị của sự tín nhiệm được xác lập sẽ trở nên rõ ràng cho cả hai bên.

Nhưng mặc cả sơ bộ còn phục vụ cho một mục đích khác. Sự mặc cả chỉ có thể diễn ra khi có ít nhất một bên khởi xướng

đưa ra việc mặc cả. Một trở ngại cho việc khởi xướng là thông tin mà nó gợi ra, hoặc có vẻ sẽ gợi ra, về sự sốt sắng của mình. Nhưng nếu mỗi bên có lý do rõ ràng để dự đoán rằng hai bên sẽ gặp nhau ở giữa, nhờ lịch sử mặc cả thành công, thì chính lịch sử đó sẽ giúp tránh suy diễn về sự quá sốt sắng.<sup>19</sup>

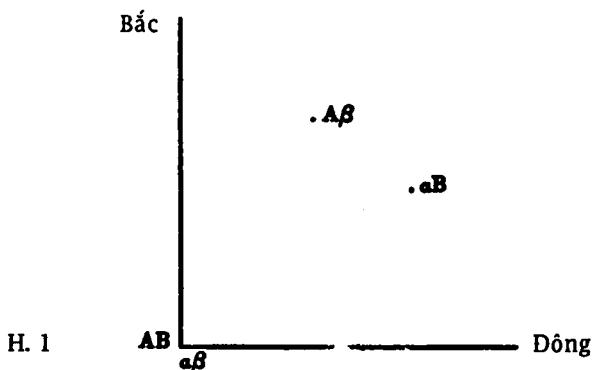
## MỘT TRÒ CHƠI MINH HỌA

Nhiều tình huống mặc cả liên quan đến sự cam kết, sự đe dọa, hứa hẹn, và các vấn đề truyền đạt thông tin, có thể được minh họa bởi các dạng khác nhau của một trò chơi trong đó mỗi người trong hai người có một cặp phương án để lựa chọn. Bắc chọn A hoặc  $\alpha$ , còn Đông chọn B hoặc  $\beta$ . Lợi ích thu được của mỗi bên phụ thuộc vào sự lựa chọn của cả hai bên. Mỗi lựa chọn kết hợp trong bốn khả năng có thể xảy ra,  $AB$ ,  $A\beta$ ,  $\alpha B$ , hay  $\alpha\beta$ , sẽ đem đến cái được hay mất cụ thể cho Bắc và cái được hay mất cụ thể cho Đông. Giữa Bắc và Đông không có việc đền bù cho nhau. Nói chung, sở thích của mỗi bên có thể phụ thuộc vào sự lựa chọn của bên kia.

---

<sup>19</sup> Có lẽ hai đối phương cùng hướng đến một giải pháp lớn qua thương lượng sẽ gặp thuận lợi nếu biết giữ lối thoát trong khi thương lượng những vấn đề nhỏ. Ví dụ, nếu như các lối ra cho tranh chấp Đông và Tây bị thu hẹp lại đến nỗi không còn gì để thương lượng trừ “vấn đề tối hậu” (một sự từ bỏ cuối cùng, vĩnh viễn về mọi lãnh thổ và vũ khí) thì thậm chí khả năng mở được cuộc thương lượng cho vấn đề lớn đó cũng rơi vào thế khó khăn. Hoặc nếu những vấn đề nhỏ không được loại bỏ, mà bị gắn quá chặt với vấn đề “lớn” đến nỗi việc sẵn sàng thương lượng về những vấn đề đó có thể được suy diễn là sự quá sốt sắng đối với giải pháp toàn bộ, thì khả năng dùng các cuộc mặc cả sơ bộ có thể sẽ biến mất.

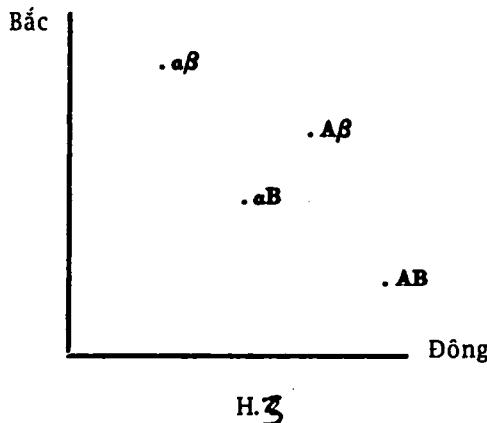
Mỗi một trò chơi như vậy có thể được biểu diễn định lượng bằng một đồ thị hai chiều, với cái được của Bắc đo theo chiều thẳng đứng và cái được của Đông đo theo chiều ngang, và các giá trị của bốn lựa chọn kết hợp được biểu thị bằng các điểm được đánh dấu  $AB$ ,  $A\beta$ ,  $\alpha\beta$  và  $\alpha B$ . Mặc dù trò chơi có cấu trúc đơn giản, nó thực ra lại có rất nhiều dạng khác nhau về chất lượng, không chỉ phụ thuộc vào vị trí tương đối của bốn điểm trên mặt phẳng mà còn vào "các quy tắc" về thứ tự bước đi, khả năng truyền đạt thông tin, các phương tiện cam kết sẵn có, khả năng thực thi các hứa hẹn, và liệu có kết hợp được từ hai trò chơi trở lên giữa hai người hay không. Số các dạng thức có thể được nhân lên hầu như không có giới hạn bằng cách lựa chọn các giả thuyết khác nhau về cái mà từng người chơi biết hay dự đoán về "các giá trị" của bốn kết cục cho đối thủ, và cái mà anh ta đoán bên kia sẽ đoán về mình. Để cho tiện, chúng ta giả định rằng tám "giá trị" ở đây là rõ ràng một cách tường minh cho cả hai bên. Và, cũng như chúng ta đã loại bỏ việc đèn bù cho nhau, chúng ta cũng loại bỏ những đe dọa hành động nằm ngoài phạm vi của trò chơi. Một ví dụ rất nhỏ của trò chơi như vậy được trình bày như sau.



Hình 1 biểu thị một tình huống mặc cả “thông thường” nếu chúng ta áp dụng quy tắc Bắc và Đông phải đạt được thỏa thuận dứt khoát trước khi họ lựa chọn.  $A\beta$  và  $\alpha B$  được coi là các thỏa thuận khác nhau mà hai bên có thể đạt được, trong khi  $AB$  và  $\alpha\beta$ , với giá trị bằng không cho cả hai bên, có thể được diễn giải tương đương với sự mặc cả “không bán.” Bất cứ bên nào có thể cam kết trước sẽ thắng. Nếu Bắc có thể tự cam kết  $A$  anh ta sẽ chắc chắn được  $A\beta$ , vì anh ta sẽ để cho Đông sự lựa chọn giữa  $A\beta$  và  $AB$  và rõ ràng phương án đầu sẽ là sự lựa chọn của Đông trong tình huống này. Nếu Đông có thể cam kết trước với phương án  $B$ , thì Bắc sẽ bị giới hạn lựa chọn giữa phương án  $\alpha B$  hoặc không có thỏa thuận (tức là hoặc  $\alpha B$  hoặc  $AB$ ) và sẽ phải chấp nhận  $\alpha B$ . Thực ra, sự cam kết trước là một kiểu “đi trước”, và trong trò chơi có cùng các con số như vậy nhưng với các nước đi lần lượt, đi trước sẽ là một lợi thế. Nếu do nhầm lẫn, cả hai bên đều cam kết, Bắc chọn  $A$  và Đông chọn  $B$ , họ sẽ tự khóa mình trong thế bế tắc tại  $AB$ .

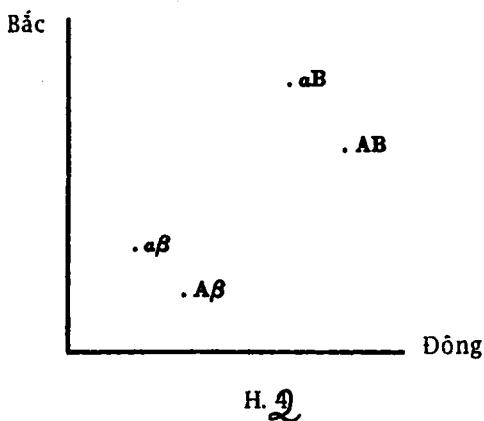
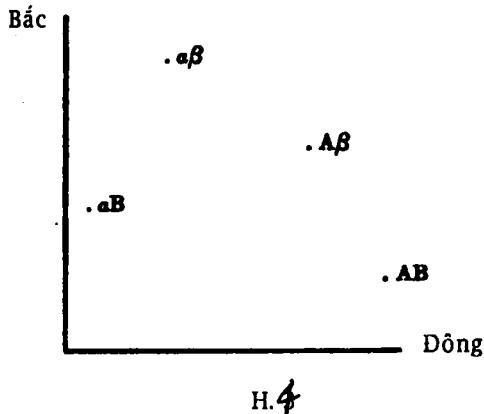
Hình 2 biểu thị một đe dọa có tính ngăn chặn, nếu chúng ta coi  $AB$  như là *nguyên trạng*, với việc Bắc muốn chuyển đến  $\alpha$  (dẫn đến  $\alpha B$ ) và Đông đe dọa chuyển đến  $\beta$  (dẫn đến  $\alpha\beta$ ) nếu Bắc làm như vậy. Nếu Bắc đi trước, Đông chỉ có thể bị thua nếu chuyển về  $\beta$ , và tương tự nếu Bắc có thể cam kết  $\alpha$  trước khi Đông có thể đưa ra đe dọa. Nhưng nếu Đông có thể đe dọa kết quả  $\alpha$  mà cả hai bên đều không muốn, thì Bắc chỉ còn một sự lựa chọn giữa  $\alpha$  hoặc  $AB$  và Bắc sẽ phải chọn phương án sau. Lưu ý rằng việc Đông cam kết *lựa chọn* của mình trước, như biểu thị trên Hình 1, là chưa đủ. Anh ta còn phải cam kết một sự lựa chọn *có điều kiện*,  $B$  hoặc  $\beta$  tùy thuộc vào việc Bắc chọn  $A$  hay  $\alpha$ . Nếu Đông cam kết lựa chọn của mình anh ta

sẽ chỉ có được lợi thế của “đi trước”, và trong trò chơi hiện tại này, nếu các nước đi là lần lượt, Bắc sẽ thắng tại B bất kể ai là người đi nước đầu tiên. (Đông sẽ chọn B thay vì  $\beta$ , để cho Bắc sự lựa chọn giữa  $\alpha B$  và  $AB$  thay vì giữa  $\alpha\beta$  hay  $A\beta$ , và Bắc sẽ chọn  $\alpha B$ . Còn Bắc, với quyền đi trước, sẽ chọn  $\alpha$  hơn là A, để lại cho Đông chọn  $\alpha\beta$  hay B hơn là A hay  $AB$ , Đông sẽ chọn  $\alpha B$ .)

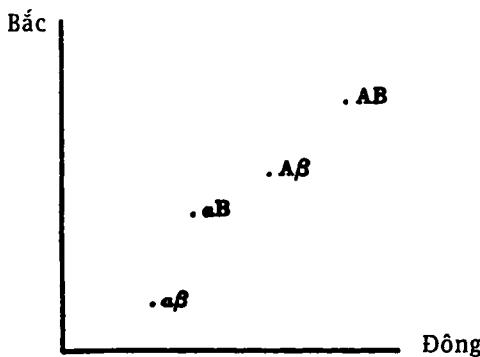


Hình 3 minh họa sự hứa hẹn. Bất cứ ai đi trước, hoặc thậm chí hai bên cùng đi một lúc, thì  $\alpha B$  là điểm “tối thiểu hóa tối đa” (minimax), ai cũng có thể tự mình đạt được, hay không bên nào có thể đe dọa bên kia bằng kết quả tồi tệ hơn. Tuy nhiên, cả hai đều muốn  $A\beta$  hơn là  $\alpha B$ , nhưng để đạt được  $A\beta$  hai bên phải tin cậy lẫn nhau hoặc có thể đưa ra hứa hẹn có tính đảm bảo thực thi. Bất kể ai đi trước, bên kia đều có động cơ gian lận. Nếu Bắc chọn A, Đông có thể chọn  $AB$ , và nếu Đông chọn  $\beta$  trước, Bắc có thể chọn  $\alpha\beta$ . Nếu các nước đi của hai bên diễn ra đồng thời, thì mỗi bên đều có động cơ gian lận, và mỗi bên có thể dự tính bên kia sẽ lừa gạt mình, mỗi

người sẽ có tình gian lận, hoặc tự vệ chống lại động cơ gian lận của bên kia, dẫn đến sự lựa chọn  $\alpha$  và  $B$ . Ít nhất một bên phải có khả năng cam kết không hành động, khi đó bên kia có thể đi trước. Nếu cả hai cùng phải đi cùng một lúc, thì cả hai đều chắc chắn có khả năng đưa ra hứa hẹn có tính đảm bảo thực thi.



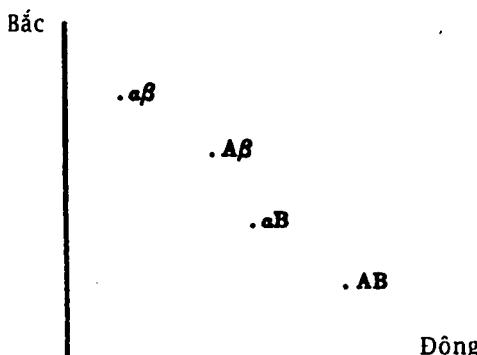
Hình 4 cũng giống như Hình 3, trừ việc  $\alpha\beta$  được dịch chuyển sang bên trái. Ở đây, với sự thiếu trao đổi thông tin, Bắc giành thắng lợi tại  $\alpha\beta$  bất chấp anh ta hay Đông đi trước hay các nước đi diễn ra đồng thời. Tuy nhiên, nếu Đông có thể truyền đạt một sự cam kết có điều kiện, anh ta có thể buộc Bắc phải lựa chọn phương án A và kết cục  $A\beta$ . Nhưng sự cam kết này còn hơn cả sự hứa hẹn hay sự đe dọa, nó vừa là sự hứa hẹn và đe dọa. Anh ta sẽ phải đe dọa  $\alpha\beta$  nếu như Bắc lựa chọn  $\alpha$ , và anh ta sẽ phải hứa hẹn "không AB" nếu Bắc chọn A. Chỉ có đe dọa không thôi sẽ không thúc đẩy Bắc tránh  $\alpha$ , đối với Bắc kết quả  $\alpha\beta$  là tốt hơn AB, và AB là kết quả anh ta sẽ có khi chọn A nếu Đông tự do lựa chọn B. Đông sẽ phải cam kết chọn, đổi lại với hoặc A, ngược lại với trường hợp nếu anh ta không cam kết: tránh chọn AB hoặc chịu hy sinh tại  $\alpha\beta$ .



H. 5

Cuối cùng, Hình 5 và 6 cho thấy hai trò chơi mà nếu tách riêng ra thì chẳng có gì đáng chú ý nhưng nếu kết hợp với nhau có thể tạo ra sự đe dọa cưỡng ép. Hình 5 có một kết quả

tối thiểu hóa tối đa tại  $\alpha\beta$ , mỗi bên đều có thể đạt được  $\alpha\beta$ , không ai có thể đảm bảo được một kết quả tốt hơn, không thể có sự hợp tác ở đây, không ai có thể đe dọa phương được. Hình 6, mặc dù tương phản với Hình 5 trong nhận dạng lợi ích giữa hai bên, cũng không có một nhu cầu bất kỳ nào về hợp tác, hay trao đổi thông tin, hay bắt cứ một sự đe dọa khả thi nào để mà khai thác. Dù có hay không có sự trao đổi thông tin, có hay không có thứ tự các nước đi, kết cục vẫn sẽ là  $AB$ .



H. 6

Nhưng giả sử hai trò chơi được đưa ra để quyết định cùng một lúc, và cả hai bên đều tham gia cả hai trò chơi. Nếu mỗi bên có thể cam kết thực hiện một đe dọa, anh ta có thể cải thiện được lợi ích của mình. Chẳng hạn, Đông có thể đe dọa lựa chọn  $\beta$  thay vì  $B$  trong trò chơi 6, trừ phi Bắc chọn  $A$  thay vì  $\alpha$  trong trò chơi 5. Hay theo một cách khác, Bắc có thể đe dọa  $\alpha$  trong trò chơi 6 trừ phi Đông đã chọn  $\beta$  trong trò chơi 5. Giả sử các khoảng cách là đủ lớn trong trò chơi 6, và lời đe dọa được cam kết và truyền đạt một cách thuyết phục, thì bên

đe dọa có lợi hơn trong trò chơi 5 mà không tốn kém gì ở trò chơi 6. Vì sự đe dọa của anh ta thành công nên anh ta không thực hiện nó. Do đó anh ta nhận được AB trong trò chơi 6 cũng như kết quả mong muốn trong trò chơi 5. Diễn đạt kết quả này theo một cách khác, trò chơi 6 cung cấp cái đã bị loại ra trước đó, tức là, sự đe dọa một hành động “nằm ngoài phạm vi trò chơi.” Theo quan điểm từ trò chơi 5, trò chơi 6 là một sự lạc đề, và Đông cũng có thể đe dọa sẽ đốt nhà của Bắc nếu anh ta không chọn A trong trò chơi 5. Nhưng một sự đe dọa cưỡng ép như vậy không phải lúc nào cũng dễ đưa ra. Nó thường đòi hỏi có cơ hội, đối tượng, và phương tiện liên lạc, và ngoài ra thường bị coi là bất hợp pháp, vô đạo đức, hoặc gặp phải sự phản kháng do sự ngoan cố. Việc kết hợp hai cuộc thương lượng trong cùng một chương trình nghị sự có thể thành công khi mà sự đe dọa vô cớ là không thể vận dụng được.

Nếu Bắc không thể cam kết sự đe dọa của mình, và do vậy chỉ muốn ngăn chặn sự đe dọa của Đông, thì việc không thể liên lạc được với nhau sẽ có lợi cho anh ta. Hoặc nếu việc liên lạc diễn ra, thì anh ta sẽ có lợi khi hai trò chơi không cùng nằm trong một chương trình nghị sự. Hoặc nếu anh ta không thể ngăn chặn Đông đe cập đến hai vấn đề cùng một lúc, thì anh ta sẽ có lợi nếu chuyển từng trò chơi cho một đại diện khác mà thù lao của người này chỉ phụ thuộc vào kết quả trong trò chơi của mình. Nếu Bắc có thể bắt buộc trò chơi 6 được chơi trước, và không có khả năng cam kết đáp lại một sự đe dọa, thì sự đe dọa sẽ được ngăn chặn. Nếu anh ta có thể cam kết lựa chọn của mình trong trò chơi 5 trước khi sự đe dọa được đưa ra, anh ta sẽ an toàn. Nhưng nếu anh ta có thể

cam kết trong trò chơi 5, và trò chơi 6 được chơi trước, Đông có thể đe dọa chọn  $\beta$  trong trò chơi 6 trừ phi Bắc có cam kết trước chọn  $A$  trong trò chơi 5. Trong trường hợp này khả năng tự cam kết của Bắc là một bất lợi, vì nó làm cho anh ta bị bắt buộc phải “chơi” trò chơi 5 trước trò chơi 6.

Cần nói thêm rằng, việc thả  $AB$  rơi theo chiều thẳng đứng trong Hình 2 xuống dưới mức  $\alpha\beta$  sẽ minh họa cho một nguyên tắc quan trọng, tức là, việc dịch chuyển một điểm theo hướng “bất lợi” cho Bắc có thể thực sự cải thiện kết quả cho anh ta. Sự đe dọa ngăn cản anh ta chiến thắng trong Hình 2 phụ thuộc vào sự hấp dẫn tương đối của  $AB$  so với  $\alpha\beta$  đối với Bắc, và nếu  $AB$  được dịch chuyển theo hướng bất lợi cho anh ta so với  $\alpha\beta$  anh ta sẽ không bị ảnh hưởng bởi lời đe dọa, mà khi đó không được đưa ra nữa, và anh ta sẽ thắng tại  $\alpha\beta$ . Đây là một ví dụ trừu tượng về nguyên tắc nói rằng, trong mặc cả, sự yếu đuối có thể là sức mạnh.

### 3

## MẶC CẢ, TRAO ĐỔI THÔNG TIN, VÀ CHIẾN TRANH HẠN CHẾ

Chiến tranh hạn chế đòi hỏi được dựa trên các giới hạn, các thủ đoạn chiến lược cũng vậy nếu chúng cần được giữ cân bằng để tránh gây ra chiến tranh. Nhưng các giới hạn đòi hỏi có sự thỏa thuận hoặc ít nhất cũng là việc công nhận và bằng lòng với nhau. Và thỏa thuận về các giới hạn là khó đạt được, không chỉ vì những yếu tố bất định và các khác biệt lợi ích sâu sắc mà còn vì việc đàm phán sẽ gặp những trở ngại nghiêm trọng cả trong thời gian chiến tranh lẫn thời gian trước khi chiến tranh bắt đầu vì trao đổi thông tin giữa các địch thủ trở nên khó khăn trong thời chiến. Hơn nữa có vẻ như một bên sẽ có lợi khi tránh có thỏa thuận về các giới hạn, nhằm tăng nỗi sợ hãi chiến tranh của bên kia. Hoặc một bên hoặc cả hai bên có thể lo sợ rằng ngay cả việc tỏ ý muốn đàm phán cũng sẽ được bị coi là sờ sắng quá mức.

Do vậy việc nghiên cứu mặc cả ngầm (tacit) - mặc cả trong đó trao đổi thông tin là không hoàn toàn hoặc không thể - có tầm quan trọng, khi có liên quan đến chiến tranh hạn chế,

hay, cũng vì lý do đó, cạnh tranh có giới hạn, các mưu mẹo pháp lý, việc xoay xở khi giao thông tắc nghẽn, hay chung sống với người láng giềng mà mình không nói chuyện. Vấn đề là cần xây dựng tam ước khi một bên hay cả hai bên hoặc không thể hoặc không muốn đàm phán công khai hoặc không bên nào tin cậy bên kia về bất cứ thỏa thuận nào có thể đạt được công khai. Chương này sẽ khảo sát một số khái niệm và nguyên tắc mà đường như là cơ sở cho mặc cả ngầm và sẽ cố gắng đưa ra một vài kết luận để minh họa vấn đề chiến tranh hạn chế hoặc các tình huống tương tự. Nó cũng sẽ gợi ý rằng chính những nguyên tắc này thường đem lại một căn cứ vững chắc để thậm chí hiểu rõ cả trường hợp khác biệt về logic về mặc cả công khai khi có trao đổi thông tin đầy đủ và sự đảm bảo thực thi.

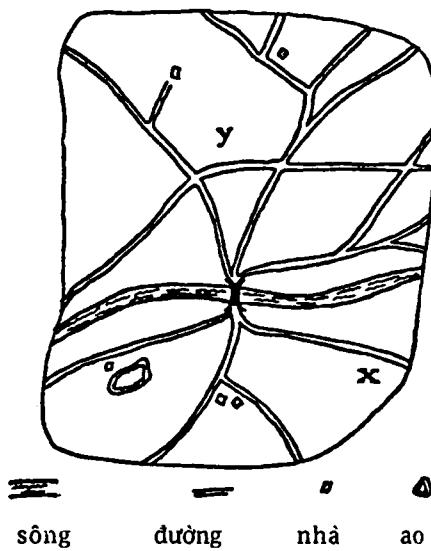
Các tình huống đáng chú ý nhất và quan trọng nhất là những tình huống trong đó có sự xung đột lợi ích giữa các bên liên quan. Nhưng sẽ dễ hiểu nếu ta bắt đầu bằng một trường hợp đơn giản hóa đặc biệt trong đó hai hoặc nhiều bên có những lợi ích đồng nhất và gặp phải vấn đề không phải về hòa giải các lợi ích mà chỉ về phối hợp các hành động vì lợi ích chung, khi mà việc trao đổi thông tin là không thể thực hiện được. Trường hợp đặc biệt này nêu ra một cách rõ ràng nguyên tắc sẽ được sử dụng để giải quyết vấn đề “mặc cả” ngầm về các sở thích xung đột.

## *SỰ PHỐI HỢP NGẦM (CÁC LỢI ÍCH CHUNG)*

Khi hai vợ chồng lạc nhau trong một cửa hàng bách hóa mà không có thỏa thuận trước nào về sẽ gặp nhau ở đâu nếu hai người bị lạc, họ có nhiều cơ hội để tìm ra nhau. Có thể mỗi người sẽ nghĩ đến một địa điểm rõ ràng dễ gặp nhau, rõ ràng đến mức mỗi người tin chắc rằng người kia cũng tin chắc chỗ đó là "rõ ràng" đối với cả hai người. Một trong hai người không chỉ đơn thuần dự đoán người thứ hai đi đâu, vì người kia sẽ đến chỗ mà người đó dự đoán người thứ nhất đi đến, tức là nơi mà người thứ nhất dự đoán người thứ hai dự đoán người thứ nhất sẽ đến đó, và cứ như vậy đến vô cùng. Không phải là "Mình sẽ làm gì nếu là cô ấy?" mà là "Mình sẽ làm gì nếu là cô ấy đang băn khoăn cô ấy sẽ làm gì nếu là mình đang băn khoăn mình sẽ làm gì nếu mình là cô ấy...?" Điều cần thiết là phối hợp được các dự đoán, để đọc được cùng một thông điệp trong tình huống chung, để nhận ra được một hướng hành động mà dự đoán về nhau hội tụ. Họ phải "cùng nhau nhận ra" dấu hiệu độc nhất phối hợp các dự đoán về nhau. Chúng ta không dám chắc là họ sẽ tìm ra nhau, cũng như không phải tất cả các cặp vợ chồng đều nhận ra một dấu hiệu giống nhau, nhưng họ chắc chắn có cơ hội tìm ra nhau hơn nhiều so với việc họ theo đuổi hướng tìm kiếm ngẫu nhiên.

Bạn đọc có thể tự mình thử giải bài toán này với bản đồ tiếp sau đây (Hình 7). Hai người bất ngờ nhảy dù xuống một

khu vực như trên hình vẽ, mỗi người đều có bản đồ và biết rằng người kia cũng có bản đồ này, nhưng không người nào biết chỗ người kia rơi xuống và đều không thể liên lạc trực tiếp với nhau được. Họ cần phải nhanh chóng tìm gặp nhau để được cứu giúp. Liệu họ có thể nghiên cứu bản đồ và “phối hợp” hành vi của họ? Liệu bản đồ này có gợi ý một địa điểm gặp nhau đặc biệt nào đó rõ ràng đến mức mỗi người đều tin rằng người kia cũng tự tin đọc được cùng một gợi ý?



Hình 7

Tác giả đã thử bài toán này và các bài toán khác tương tự trên một mẫu không có tính đại diện những người tham gia, và kết luận rút ra là người ta thường có thể phối hợp với nhau. Những câu đố trừu tượng dưới đây là ví dụ điển hình cho

những bài toán có thể “giải được” bởi một tỷ lệ lớn những người tham gia. Tất nhiên, các đáp số này có phần tùy nghi ở mức độ: mọi đáp số là “đúng” nếu có đủ số người nghĩ như vậy. Bạn đọc có thể muốn khẳng định khả năng của mình trong việc phôi hợp giải quyết các bài toán sau với những người mà kết quả của họ được nêu ra trong phần chú thích dưới đây.<sup>1</sup>

1. Nói “mặt ngửa” hay “mặt sấp.” Nếu bạn và đối tác của bạn nói giống nhau, cả hai đều trúng thưởng.

2. Khoanh tròn vào một trong các con số nằm trong hàng

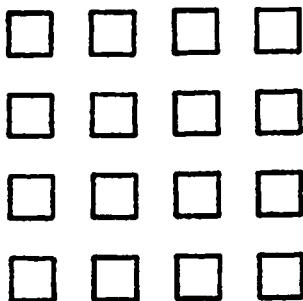
---

<sup>1</sup> Trong mẫu nghiên cứu của tác giả, có 36 người cùng chọn “mặt ngửa” trong bài toán 1, và chỉ có sáu người chọn “mặt sấp.” Trong bài toán 2, ba con số đầu tiên nhận được 37 phiếu trong tổng số 41 phiếu, số 7 dẫn số 100 với khoảng cách không lớn, số 13 đứng vị trí thứ ba. Góc phía trên bên trái trong bài toán 3 nhận được 24 phiếu trong tổng số 41, và tất cả trừ 3 số của phần còn lại được phân bổ trên cùng một đường chéo. Bài toán 4, mà có thể phản ánh địa điểm của mẫu điều tra tại New Haven, Connecticut, cho thấy một đa số tuyệt đối cố gắng gấp nhau tại Nhà ga Trung tâm Lớn (quầy thông tin), và hầu như tất cả thành công gấp nhau lúc 12 giờ trưa. Bài toán 6 cho thấy nhiều kiểu trả lời khác nhau, nhưng hai phần năm số người cùng chọn số 1. Và trong bài toán 7, trong số 41 người, có 12 người cùng chọn 1.000.000 đô-la, và chỉ có ba lựa chọn không phải là số lũy thừa của 10, trong đó hai lựa chọn là 64 đô-la, và theo kết quả gần đây hơn là 64.000 đô-la! Bài toán 8 không có khó khăn gì đối với 36 trong số 41 người, những người chia tổng số thành 50-50. Bài toán 9 có đa số 20 trong số 22 người chọn Robinson. Một cách thiết lập khác, trong đó Jones và Robinson hòa nhau trong lần bỏ phiếu thứ nhất với 28 phiếu, được tác giả cố tình sử dụng để chứng minh sự khó khăn trong phôi hợp khi có kết quả hòa, nhưng những người được hỏi ý kiến đã vượt qua khó khăn này và cho Jones 16 phiếu trong số 18 phiếu (rõ ràng là trên cơ sở vị trí đứng trước của Jones trên danh sách), chứng minh cho luận điểm chính nhưng lại áp đảo điểm phụ trong quá trình này. Trên bản đồ gần như hoàn toàn giống với bản đồ được vẽ lại ở đây (Hình 7), 7 trong số 8 người được hỏi ý kiến đều sẽ gấp tại cây cầu.

dưới đây. Các bạn sẽ thắng nếu tất cả cùng khoanh vào cùng một con số.

7      100      13      261      99      555

3. Đánh dấu vào một trong 16 ô sau. Các bạn sẽ thắng nếu tất cả cùng đánh dấu vào cùng một ô.



4. Bạn sẽ phải gặp ai đó tại thành phố New York. Bạn chưa được thông báo địa điểm gặp người đó, bạn chưa có thỏa thuận trước với người đó sẽ gặp ở đâu, và các bạn không thể liên lạc được với nhau. Bạn chỉ được thông báo là bạn sẽ phải đoán nơi sẽ gặp ở đâu và người đó cũng đã được thông báo như vậy và rằng các bạn sẽ phải cố làm sao cho hai sự phán đoán này khớp với nhau.

5. Bạn được thông báo ngày gặp chứ không phải giờ gặp trong bài toán 4. Bạn và người kia phải đoán chính xác đến phút để gặp nhau trong ngày hôm đó. Vậy bạn sẽ có mặt vào lúc nào tại địa điểm mà các bạn đã chọn trong bài toán 4?

6. Hãy viết ra một số dương. Nếu tất cả các bạn cùng viết ra một số giống nhau, các bạn sẽ thắng.

7. Nêu ra một số tiền. Nếu tất cả các bạn đều cùng nói ra một số tiền, các bạn sẽ được số tiền đó.

8. Bạn sẽ phải chia 100 đô-la thành hai tập, đánh dấu là A và B. Đối tác của bạn cũng sẽ chia số tiền 100 đô-la khác thành hai tập được đánh dấu là A và B. Nếu như bạn phân bổ cùng một số tiền vào A và B giống với cách phân bổ tương ứng của đối tác, thì mỗi người sẽ được 100 đô-la. Nếu cách chia của bạn khác với đối tác, cả hai sẽ chẳng được gì.

9. Trong cuộc bỏ phiếu thứ nhất, các ứng cử viên nhận được số phiếu như sau:

Smith.....	19	Robinson.....	29
Jones.....	28	White.....	9
Brown.....	15		

Cuộc bỏ phiếu thứ hai sắp diễn ra. Bạn không quan tâm gì đến kết quả bỏ phiếu, trừ việc bạn sẽ được thưởng nếu ai đó nhận được đa số phiếu trong lần bỏ phiếu thứ hai và bạn bỏ phiếu cho người đó. Tương tự như vậy, tất cả cử tri chỉ quan tâm đến việc bỏ phiếu theo đa số, và mọi người đều biết rằng đây là lợi ích của mọi người. Vậy bạn sẽ bỏ phiếu cho ai trong lần bỏ phiếu thứ hai?

Những bài toán này mang tính nhân tạo, nhưng chúng minh họa cho luận điểm đang bàn đến. Người ta có khả năng phối hợp các ý định hay dự đoán của mình với những người khác nếu mỗi người biết rằng người khác cũng đang cố gắng làm như vậy. Hầu hết các tình huống - có thể là tất cả tình huống đối với những người được thử trong trò chơi kiểu này - đều đưa ra điều nào đó để phối hợp hành vi, một tiêu điểm (focal point) cho sự dự đoán của mỗi người về cái mà người khác dự đoán anh ta sẽ dự đoán được người khác dự đoán sẽ làm gì. Việc tìm ra chìa khóa, hay *một* chìa khóa - tức

là bất cứ chìa khóa nào được thừa nhận chung là chìa khóa sẽ trở thành *chìa khóa* - phụ thuộc vào trí tưởng tượng hơn là logic. Nó có thể phụ thuộc vào sự tưởng tự, tiền lệ, cách xếp đặt ngẫu nhiên, tính chất đối xứng, cấu trúc theo thẩm mỹ hoặc hình học, sự lập luận ngụy biện, và các bên là ai và họ biết gì về nhau. Một ý nghĩ bất chợt có thể đưa người chồng và người vợ tìm đến “phòng khai báo thất lạc”, hay logic có thể khiến hai người ngẫm nghĩ và dự đoán người kia cũng ngẫm nghĩ nơi mà họ dự kiến gặp nhau nếu họ đã có thỏa thuận trước để đề phòng khi lạc nhau. Chúng ta đang không khẳng định rằng họ sẽ luôn luôn tìm ra một lời giải rõ ràng cho vấn đề này, nhưng qua cách làm này họ sẽ có cơ hội lớn hơn nhiều so với chỉ dựa vào logic thuần túy của phương pháp xác suất ngẫu nhiên trừu tượng.

Một đặc điểm chủ yếu của hầu hết các “lời giải” cho các vấn đề này, tức là, về các đầu mối hay các nhân tố phối hợp hay các tâm điểm, là ở tính nổi bật hoặc rõ ràng. Nhưng chính sự nổi bật lại phụ thuộc vào thời gian và địa điểm và những người đó là ai. Những người bình thường bị lạc trong một khu vực hình tròn bằng phẳng tất nhiên sẽ đi đến tâm hình tròn để tìm gặp nhau. Nhưng chỉ những người giỏi toán học mới có thể “tự nhiên” dự đoán sẽ gặp được người bạn của mình tại trọng tâm của một khu vực có hình thù bất thường. Điều cốt lõi không kém là tính duy nhất nào đó. Anh chồng và người vợ sẽ không gặp nhau được tại “phòng khai báo thất lạc” nếu cửa hàng bách hóa có nhiều phòng như vậy. Các thí nghiệm của tác giả với các bản đồ khác cho thấy rất rõ ràng một bản đồ có nhiều ngôi nhà và một ngã tư duy nhất thì sẽ đưa mọi người người đến ngã tư đó, trong khi một bản đồ có nhiều ngã

tư và một ngôi nhà độc nhất sẽ khiến hầu hết mọi người tìm đến ngôi nhà. Điều này một phần chỉ nói lên rằng tính duy nhất mang theo tính nổi bật, nhưng có thể quan trọng hơn rằng tính duy nhất tránh được sự mơ hồ. Các ngôi nhà vốn đã nổi bật hơn các thứ khác trên bản đồ. Nhưng nếu có đến ba ngôi nhà, thì không có cái nào nổi trội hơn những cái khác, nhưng cơ hội gặp nhau tại một ngôi nhà là một phần ba, và việc nhận ra thực tế này có thể dẫn đến việc bác bỏ lấy các ngôi nhà làm đầu mối.”<sup>2</sup>

Nhưng trong phân tích cuối cùng, chúng ta phải dùng đến trí tưởng tượng nhiều chẳng kém gì logic, và bản thân logic là một thứ khá nguy hiểm. Các nhà thơ có thể làm tốt hơn các nhà logic trong trò chơi này, mà có vẻ giống với chơi “đố chữ” hơn là đánh cờ. Logic giúp chúng ta – một số lớn người cùng viết số 1 trong bài toán 6 có vẻ như dựa vào logic – nhưng thường sau khi trí tưởng tượng đã chọn ra một đầu mối nào đó để tập trung vào đó trong số rất nhiều các chi tiết cụ thể của tình huống.

---

<sup>2</sup> Rằng lập luận này sẽ “đúng,” một cách tinh cù, được gợi ra bởi một trong các thí nghiệm về bản đồ của tác giả. Trên bản đồ với một ngôi nhà duy nhất và nhiều ngã tư đường, 11 người chọn ngôi nhà đều gặp nhau, trong khi 4 người chọn ngã tư đều chọn các ngã tư khác nhau và thậm chí không có ai gặp nhau.

## MẶC CẢ NGẦM (CÁC LỢI ÍCH KHÁC BIỆT)

Xung đột lợi ích xuất hiện trong bài toán của chúng ta nếu những người nhảy dù không thích đi bộ. Nếu có thông tin liên lạc, điều không được phép trong bài toán của chúng ta, có thể họ sẽ tranh cãi hoặc mặc cả về nơi sẽ gặp nhau, mỗi người đều thích điểm ở gần chỗ mình hoặc một điểm dừng chân mà anh ta đặc biệt ưa thích. Trong trường hợp không có thông tin liên lạc, mối quan tâm bao trùm của họ là phối hợp các ý nghĩ với nhau, và nếu một chỗ nào đó gây chú ý là điểm “rõ ràng” để gặp nhau, thì người thắng trong cuộc mặc cả chỉ đơn thuần là người tình cờ ở gần chỗ đó hơn mà thôi. Thậm chí nếu người ở cách xa tiêu điểm nhất biết anh ta ở xa như vậy, anh ta không thể không bằng lòng và đòi sự phân chia công bằng hơn về đoạn đường đi bộ. “Đè xuất” cho cuộc mặc cả mà do chính bản thân bản đồ đưa ra – nếu, như trong thực tế, nó đưa ra một đè xuất như vậy – chỉ là lời đè xuất tồn tại duy nhất, và khi không có thông tin liên lạc, sẽ không có phản đè nghị nào có thể được đưa ra. Xung đột đã được hòa giải – hoặc chúng ta có thể nói là đã được bỏ qua – như là một sản phẩm phụ của nhu cầu nổi trội về phối hợp.

Nói “thắng” hay “thua” có thể là không được hoàn toàn chính xác, vì cả hai có thể thua so với cái mà họ đã có thể thỏa thuận với nhau thông qua liên lạc. Nếu hai người thực sự ở gần nhau và cách xa ngôi nhà duy nhất trên bản đồ, họ đã có thể loại bỏ việc đi bộ xa đến ngôi nhà nếu như họ nhận biết

được các vị trí của họ và phối hợp với nhau công khai để xác định địa điểm gặp nhau nằm giữa họ. Hoặc có thể là một người “thắng” trong khi người kia thua nhiều hơn người thứ nhất thắng: nếu cả hai người cùng ở về một phía của ngôi nhà và đi bộ đến ngôi nhà đó, họ đi bộ với khoảng cách tổng cộng lớn hơn cần thiết, nhưng người ở gần ngôi nhà hơn vẫn có thể có lợi so với trường hợp nếu anh ta phải thỏa thuận với người kia.

Ví dụ sau cùng này minh họa một điều là một bên có thể có lợi nếu không thể trao đổi được. Ở đây xuất hiện động cơ phá hủy trao đổi thông tin hay từ chối việc phối hợp từ trước về cách thức gặp gỡ nếu như người đó biết được lợi thế của mình và tin vào “đáp số” mà người đó nhìn thấy trước. Trong một biến thể do tác giả thử nghiệm, A biết B ở đâu, nhưng B lại không biết A ở đâu (và mỗi bên đều biết bên kia biết đến đâu). Hầu hết những người nhận được bảng hỏi loại B thì ngồi yên một cách ung dung, tận hưởng việc mình không biết gì, còn gần như tất cả những người tham gia với bảng hỏi loại A thì chắc chắn nhận ra việc mình không thể tránh và đi đến tận chỗ của B. Tốt hơn nữa là có khả năng gửi đi thông điệp nhưng không tiếp nhận được thông điệp: nếu một bên thông báo về địa điểm của mình và cho biết máy truyền tin của anh ta hoạt động tốt nhưng máy nhận tin không hoạt động được, và nói rằng anh ta sẽ ngồi chờ tại vị trí của mình cho đến khi người kia tìm đến, thì người sau sẽ không còn lựa chọn nào khác. Anh ta không thể đưa ra phản đòn nghị có hiệu quả, vì không có phản đòn nghị nào đến được tai bên kia.<sup>3</sup>

---

<sup>3</sup>Đây là một ví dụ về một nghịch lý chung, đã được minh họa chi tiết tại Chương 2, tức là sự bất lực theo các tiêu chuẩn thông thường thì, trong mặc cả, có thể là một nguồn “sức mạnh.”

Tác giả đã thử nghiệm một mẫu các trò chơi xung đột lợi ích đối với một số người, bao gồm các trò chơi thiên vị bên này hay bên kia, và về tổng thể, kết quả gợi ra cùng một kết luận như trong các trò chơi hợp tác thuần túy. Tất cả các trò chơi này đòi hỏi sự phối hợp. Tuy nhiên, chúng cũng đưa ra một vài lựa chọn khác mà trên đó lợi ích của hai bên khác biệt. Ấy thế mà, trong số tất cả các kết quả, một số kết quả luôn có vẻ là tiêu điểm cho lựa chọn phối hợp, và bên mà kết quả này tương đối không có lợi thường xuyên chấp nhận nó đơn giản là vì anh ta biết rằng bên kia kỳ vọng anh ta sẽ làm như thế. Những lựa chọn mà không thể phối hợp được các dự đoán không phải thực sự "có sẵn" nếu không có trao đổi thông tin. Đặc điểm khác thường của tất cả các trò chơi này là không đối thủ nào có thể giành được thắng lợi bằng sự khôn ngoan hơn bên kia. Bên nào cũng sẽ thua trừ phi anh ta làm đúng như bên kia kỳ vọng anh ta phải làm. Mỗi bên đều vừa là con tin vừa là bên hưởng lợi của các kỳ vọng chung của họ. Không bên nào có thể chối bỏ dự đoán của mình về cái mà bên kia kỳ vọng anh ta sẽ kỳ vọng được kỳ vọng sẽ làm. Sự cần thiết có thỏa thuận sẽ chỉ phối sự bất đồng tiềm tàng, và mỗi bên cần phải phối hợp với bên kia nếu không cả hai đều thua. Một số trò chơi sau đây được tạo ra bằng cách thay đổi một ít các bài toán đã đề cập, như chúng ta đã làm đối với bài toán bản đồ bằng cách giả định rằng việc đi bộ là một khó khăn.

1. A và B sẽ phải chọn "mặt ngửa" hay "mặt sấp" mà không có sự liên lạc với nhau. Nếu cả hai chọn "mặt ngửa," A sẽ được 3 đô-la và B sẽ được 2 đô-la. Nếu cả hai chọn "mặt sấp," A được 2 đô-la và B được 3 đô-la. Nếu họ chọn khác nhau, cả hai bên sẽ không được gì. Nếu bạn là A (hoặc B), bạn sẽ chọn

cái nào? (Lưu ý rằng nếu cả hai chọn một cách ngẫu nhiên, thì cơ hội trùng hợp chỉ là 50-50 và giá trị kỳ vọng là 1,25 đô-la cho mỗi người - ít hơn so với 3 đô-la hoặc 2 đô-la).

2. Bạn và đối tác (hoặc đối thủ) mỗi người có một trong các chữ cái A, B, và C. Mỗi người sẽ phải viết ra ba chữ cái A, B, và C theo một thứ tự bất kỳ. Nếu như thứ tự trong cả 3 danh sách của các bạn giống nhau, các bạn sẽ được thưởng tổng cộng 6 đô-la, trong đó 3 đô-la dành cho người mà chữ cái của người đó xếp đầu tiên trong cả ba danh sách, 2 đô-la cho người mà chữ cái của người đó xếp thứ hai, và 1 đô-la cho người có chữ cái của mình xếp thứ ba. Nếu như các chữ cái không có thứ tự giống nhau trong cả ba danh sách thì không ai trong số các bạn được thưởng gì cả. Giả sử chữ cái của bạn là A (hoặc B, hay C), hãy viết ra đây ba chữ cái theo thứ tự mà bạn chọn:

---

3. Bạn và đối tác của bạn (hay đối thủ) mỗi người được phát một tờ giấy, một tờ giấy trắng và tờ kia có chữ "X" được viết trên đó. Người nhận được tờ giấy có chữ "X" được lựa chọn hoặc để nguyên hoặc xóa nó đi, còn người nhận được tờ giấy trắng được lựa chọn để nguyên giấy trắng hoặc viết chữ "X" lên tờ giấy đó. Nếu, sau khi các bạn đã lựa chọn xong mà không có trao đổi với nhau, và chỉ có một tờ giấy có chữ "X", thì người cầm tờ giấy có chữ "X" được 3 đô-la và người cầm tờ giấy trắng được 2 đô-la. Nếu cả hai tờ giấy đều có chữ "X" hoặc cả hai tờ giấy đều trắng, thì không ai được gì. Giả sử tờ giấy của bạn ban đầu có chữ "X" trên đó, bạn sẽ để nguyên như thế hay xóa nó đi? (Hoặc ngược lại: tờ giấy của bạn là trắng, bạn sẽ để nguyên như vậy hay viết lên đó chữ "X"?).

4. Bạn và đối tác của bạn (hoặc đối thủ) sẽ được tặng 100 đô-la nếu hai người đồng ý cách chia nó mà không trao đổi với nhau. Mỗi người sẽ viết ra số tiền mình muốn trên một tờ giấy. Và nếu hai đối hỏi cộng lại không quá 100 đô-la, thì mỗi người sẽ được nhận đúng số tiền mình viết ra. Nếu hai con số cộng lại quá 100 đô-la, không ai nhận được gì. Vậy bạn sẽ đòi bao nhiêu tiền? \_\_\_\_\_ đô-la

5. Bạn và đối tác của mình mỗi người sẽ chọn ra một trong năm chữ cái, K, G, W, L, hoặc R. Nếu hai người chọn cùng một chữ cái, hai người sẽ được thưởng. Nếu các bạn chọn hai chữ cái khác nhau, không ai được thưởng. Giải thưởng nhận được phụ thuộc vào chữ cái mà cả hai chọn, nhưng giải thưởng sẽ không giống nhau cho hai người, và chữ cái cho bạn giải thưởng lớn nhất không nhất thiết phải là chữ cái đem lại giải thưởng lớn nhất cho đối tác của bạn. Đối với bạn, các giải thưởng sẽ là như sau:

K.....\$4

L.....\$2

G.....\$3

R.....\$5

W.....\$1

Bạn không biết cấu trúc giải thưởng của đối tác của bạn là như thế nào. Bạn sẽ bắt đầu bằng cách gợi ý anh ta chọn chữ cái R, là chữ sẽ cho bạn giải thưởng lớn nhất. Trước khi anh ta trả lời, thì người chủ trò can thiệp và nhắc nhở bạn không được phép trao đổi với nhau và nếu tiếp tục trao đổi thì cả hai sẽ không được chơi nữa. Bạn đơn giản là phải viết ra một trong các chữ cái đó, hy vọng rằng bên kia cũng sẽ chọn cùng một chữ đó. Bạn sẽ chọn chữ cái nào? (Cách xác lập khác về nửa thứ hai cho thấy biểu phần thưởng K-\$3, G-\$1, W-\$4, L-\$5, R-\$2,

và để “bên kia” đưa ra đề xuất đầu tiên với chữ cái R trước khi việc liên lạc bị cắt).

6. Hai lực lượng chống đối nhau có mặt tại các điểm đánh dấu X và Y trên một bản đồ tương tự như hình 7. Chỉ huy của hai bên đều muốn chiếm được trận địa càng nhiều càng tốt và biết rằng bên kia cũng muốn như vậy. Nhưng cả hai chỉ huy đều không muốn có đúng độ quân sự và biết rằng bên kia cũng muốn như vậy. Cả hai bên ra lệnh cho quân của mình chiếm giữ đến đường ranh giới định trước và sẽ đánh đổi phương nếu bị cản lại. Một khi quân lính đã được phái đi, kết cục sẽ chỉ phụ thuộc vào các đường ranh giới mà hai bên chỉ huy đã ra lệnh cho đội quân của mình chiếm giữ. Nếu hai đường ranh giới đó chồng lên nhau, các đội quân sẽ gặp và chiến đấu với nhau, gây tổn thất cho cả hai bên. Nếu quân lính hai bên chiếm lĩnh các vị trí mà để lại bất kỳ một khoảng trống chưa có ai chiếm giữ đáng kể nào giữa họ, thì tình hình sẽ trở nên “bất ổn” và sự xung đột là không tránh khỏi. Chỉ trong trường hợp khi các đội quân được lệnh chiếm giữ đến một đường ranh giới chung hay giữa các đường ranh giới gần như không còn khoảng trống chưa bị chiếm giữ nào thì mới tránh được xung đột. Trong trường hợp đó, mỗi bên đạt được một cách thành công vùng đất do mình chiếm giữ, lợi thế sẽ thuộc về bên nào chiếm giữ vùng đất có giá trị nhất đứng về phương diện đất đai và các công trình nằm trên đó. Giả sử bạn là chỉ huy đội quân đóng tại điểm đánh dấu X (hay Y). Bạn hãy vẽ đường ranh giới trên bản đồ mà bạn sẽ cử quân đội đến chiếm lĩnh.

7. A và B có mức thu nhập tương ứng là 100 đô-la và 150 đô-la một năm. Họ được thông báo về mức thu nhập của người kia và được báo cho biết họ phải bắt đầu đóng thuế

tổng cộng là 25 đô-la một năm. Nếu họ có thể đạt được thỏa thuận về cách phân chia khoản thuế tổng cộng này, thì họ sẽ chỉ phải đóng thuế hàng năm theo mức thỏa thuận đó. Nhưng họ phải đạt được thỏa thuận đó mà không được trao đổi với nhau. Mỗi người sẽ viết ra phần thuế mà mình đề nghị đóng, và nếu các phần đó cộng lại bằng 25 đô-la hoặc lớn hơn, thì mỗi bên sẽ đóng thuế chính xác theo mức mình đề nghị. Tuy nhiên, nếu các phần thuế hai người đề nghị cộng lại không tới 25 đô-la, thì mỗi người buộc phải đóng toàn bộ 25 đô-la, và cơ quan thu thuế sẽ giữ lại phần phụ trội. Giả sử bạn là A (hoặc B), bạn sẽ đề nghị đóng thuế bao nhiêu? \_\_\_\_\_ đô-la

8. A đánh mất một số tiền, và B nhận được. Theo quy định chung của nhà tập thể, A không được nhận lại số tiền trừ phi đồng ý thưởng một khoản thích đáng cho người nhận được, và B cũng không được giữ số tiền nhận được nếu không được A đồng ý. Nếu không đạt được thỏa thuận, số tiền đó sẽ thuộc về nhà tập thể. Số tiền đó là 16 đô-la, và A đề nghị thưởng 2 đô-la. B từ chối, đòi một nửa số tiền đó. Một cuộc tranh luận nổ ra, và nhà tập thể can thiệp, yêu cầu mỗi người viết ra đề nghị của mình, chỉ một lần dứt điểm, mà không được trao đổi thêm. Nếu đề nghị của hai người cộng lại phù hợp với 16 đô-la, mỗi người sẽ nhận được đúng số tiền mình đề nghị. Nhưng nếu cộng lại quá 16 đô-la, số tiền đó sẽ bị nhà tập thể tịch thu. Trong khi họ đang ngồi băn khoăn không biết nên viết ra bao nhiêu, thì một nhà trung gian có uy tín bước vào và đề nghị giúp đỡ dàn xếp. Ông ta nói rằng ông không thể tham gia vào bất kỳ cuộc mặc cả nào, nhưng ông ta có thể đưa ra một đề xuất “công bằng.” Ông ta bước đến bên A và nói “Tôi nghĩ sự phân chia hợp lý trong trường hợp này là chia theo tỷ lệ 2:1, chủ sở hữu ban đầu nhận hai phần ba và người nhận

được nhận một phần ba, có thể làm tròn tương ứng thành 11 đô-la và 5 đô-la. Tôi cũng sẽ đề nghị với anh ta như vậy.” Không cần đợi câu trả lời, ông ta bước tới người nhặt được tiền và cũng đưa ra đề nghị tương tự, và nói rằng ông ta đã đưa ra đề xuất đó với chủ sở hữu ban đầu. Một lần nữa không đợi câu trả lời, ông ta rời đi. Giả sử bạn là A (hoặc B). Số tiền bạn định viết ra là bao nhiêu?

Kết quả trong các mẫu nghiên cứu không chính thức của tác giả được ghi lại trong phần chú thích.<sup>4</sup> Trong các bài toán đó nơi xuất hiện sự bất đối xứng giữa “bạn” và “anh ta,” tức

---

<sup>4</sup> Trong bài toán đầu tiên, 16 trong số 22 người của nhóm A và 15 trong số 22 người của nhóm B chọn mặt ngửa. Với cái mà A đã chọn, mặt ngửa là câu trả lời tốt nhất cho B. Với cái mà B đã chọn, mặt ngửa là câu trả lời tốt nhất cho A. Cộng lại cả hai sẽ đạt được kết quả tốt hơn nhiều so với việc làm một cách ngẫu nhiên, và tất nhiên, nếu mỗi bên đều cố để thắng 3 đô-la, thì cả hai sẽ nhận được con số 0. Bài toán 2, tương tự về logic như bài toán 1, nhưng có cấu trúc hấp dẫn hơn, cho thấy 9 trong số 12 người nhóm A, 10 trong số 12 người nhóm B, và 14 trong số 16 người nhóm C, phối hợp thành công chọn thứ tự ABC. (Trong số 7 trường hợp còn lại, một cách ngẫu nhiên, 5 trường hợp gây ra khác biệt bằng cách không theo thứ tự ABC, tất cả đều thua bại.) Bài toán 3, về cấu trúc giống bài toán 1, cho thấy 18 trong số 22 người nhóm A phối hợp thành công với 14 trong số 19 người nhóm B, cho A được thưởng 3 đô-la. Trong bài toán 4, 36 trong số 40 chọn 50 đô-la. (Hai trong số còn lại chọn 49 đô-la và 49,99 đô-la.) Trong bài toán 5, chữ cái R giành được 5 trong số 8 phiếu của những người đưa ra đề nghị, và 8 trong số 9 phiếu từ những người ở phía bên kia. Trong bài toán 6, 14 trong số 22 người của bên X và 14 trong số 23 của bên Y chọn đường ranh giới của mình dọc theo con sông. Sự “đúng đắn” của kết quả này được thể hiện qua thực tế là 15 người còn lại, những người không chọn con sông, vẽ ra 14 đường ranh giới khác nhau. Trong số 8x7 các cặp có thể có, có 55 cặp thất bại và một cặp thành công. Bài toán 7 cho thấy 5 trong số 6 người với thu nhập 150 đô-la và 7 trong 10 người với thu nhập 100 đô-la phối hợp để có tỷ lệ phân chia thuế 15-10. Trong bài toán 8 cả hai bên những người đánh mất tiền và người nhặt được tiền, lần lượt là 8 và 7 người, nhặt trúng số tiền với nhau theo gợi ý của người trung gian với mức thưởng 5 đô-la chẵn.

là giữa A và B, các phát biểu của A được so với các phát biểu của B để đi đến “kết quả.” Kết luận chung, như đã được nêu cụ thể hơn trong chủ thích, là các bên tham gia có thể “giải” bài toán của họ trong phần lớn các trường hợp. Họ thực sự có thể giải quyết vấn đề tốt hơn một cách rõ rệt so với kết quả mà các phương pháp ngẫu nhiên cho phép, và thậm chí bên bị bắt lợi trong các trò chơi thiên vị vẫn tuân thủ luật chơi đòi hỏi có sự phối hợp.

Các “đầu mối” trong những trò chơi này rất đa dạng. Kết quả chọn mặt ngửa áp đảo mặt sấp một cách rõ ràng vì một quy ước nào đó về ưu tiên, tương tự như quy ước chỉ ra A, B, C, mặc dù nó không rõ bằng. Số trường hợp giữ chữ X ban đầu trội hơn việc có tờ giấy trắng, rõ ràng là vì việc giữ “nguyên trạng” là điều hiển nhiên hơn là sự thay đổi. Chữ cái R chiến thắng vì không có đề xuất nào vượt lên đề xuất đầu tiên. Việc lựa chọn các con đường, về nguyên tắc, cũng hợp lý chẳng kém việc chơi các con sông, đặc biệt khi sự đa dạng của chúng cho phép ít tùy nghi hơn trong lựa chọn. Nhưng, chính vì sự đa dạng của chúng, mà bản đồ không thể chỉ rõ được con đường nào, do đó các con đường đã không được lựa chọn mà được thay thế bằng một con sông duy nhất, và không còn sự mơ hồ. (Có lẽ trong một bản đồ đối xứng với địa hình đồng nhất, kết cục sẽ gần với cách phân chia 50-50 trong ví dụ về 100 đô-la - có thể là sự chia đôi theo đường chéo - nhưng một bản đồ với nhiều điểm bất thường đã loại trừ một đáp án hình học.)

Bài toán về thuế minh họa rõ ràng sức mạnh gợi ý trong các con số thu nhập. Logic trừu tượng của bài toán này tương tự như cách phân chia 100 đô-la. Thực ra, nó có thể được mô

tả lại như sau: mỗi bên sẽ đóng thuế 25 đô-la, và khoản hoàn thuế 25 đô-la sẽ được phân chia cho hai người nếu họ đạt được thỏa thuận về cách chia. Cách xác lập bài toán này cũng tương đương về mặt logic với bài toán 7, và như vậy nó chỉ khác với bài toán 4 về khoản tiền 25 đô-la thay vì 100 đô-la. Tuy nhiên, việc đưa vào các con số về thu nhập, chỉ bằng cách *gợi ý* sự can hệ của chúng và làm cho chúng trở nên nổi bật trong bài toán, làm dịch chuyển hẳn tiêu điểm về cách phân chia 10-15 thay vì 12,5-12,5. Và tại sao, khi thu nhập có liên quan, thì mức thuế theo tỷ lệ hoàn hảo lại hiển nhiên như vậy, khi mà có lẽ có cơ sở về các mức thuế thu nhập lũy tiến? Câu trả lời phải là không có cách tính mức thuế theo lũy tiến *cụ thể nào* là hiển nhiên đến mức có thể chấp nhận mà không cần trao đổi. Và nếu không thể trao đổi được, thì nghiêm nhiên nguyên tắc đơn giản độc nhất và có thể được công nhận về tính thuế theo tỷ lệ phải được áp dụng. Trước tiên các con số về mức thu nhập làm cho việc phân chia thuế theo tỷ lệ 50-50 không còn thích hợp. Sau đó tính đơn giản của việc phân chia theo tỷ lệ làm cho việc phân chia 10-15 là phương án duy nhất được ngầm công nhận. Nguyên tắc tương tự cũng được trình bày bởi một thử nghiệm trong đó bài toán 7 được phức tạp hóa bởi các dữ kiện *bổ sung* - quy mô gia đình, các thói quen chi tiêu, v.v... Ở đây, sự hấp dẫn đặc biệt của cách phân chia theo tỷ lệ thu nhập trở nên bị pha loãng đến nỗi ý kiến áp đảo từ những người có cả thu nhập cao và thu nhập thấp đơn giản là chia 50-50 khoản đóng thuế. Tín hiệu tinh tế của cách phân chia theo tỷ lệ thu nhập bị chìm đi vì "nhiều," và tín hiệu thô hơn về sự bình đẳng là cái duy nhất được nhận ra.

Cuối cùng, bài toán 8 cũng lại giống bài toán 4 xét về logic, khoản tiền 16 đô-la sẽ được chia cho hai người nếu họ có thể viết ra hai con số cộng lại không quá số tiền đó. Nhưng sắp xếp về mặt thể chế lại có sự phân biệt. Người nhặt được tiền và người đánh mất tiền không có sự bình đẳng có tính thuyết phục về cả đạo lý lẫn pháp lý, do đó việc chia theo tỷ lệ 50-50 không hẳn đã là đương nhiên. Đề xuất của nhà trung gian cung cấp tín hiệu duy nhất khác mà có thể thấy được. Sức mạnh của đề xuất này với tư cách là sự phối hợp được thấy thậm chí trong việc làm tròn thành 11 đô-la và 5 đô-la, giải pháp mà cả hai người chấp nhận.

Trong từng tình huống này, kết cục được quyết định bởi những yếu tố khá tùy nghi. Nó không phải là một kết cục đặc biệt “công bằng,” từ góc nhìn của người quan sát lẫn những người trong cuộc. Thậm chí cách phân chia 50-50 cũng là sự tùy tiện dựa trên một dạng thuần túy toán học dễ thấy, và nếu đó là “công bằng,” thì cũng chỉ vì chúng ta không có dữ liệu cụ thể dựa vào đó để phán xét tính không công bằng của nó, như là nguồn gốc của số tiền, nhu cầu tương đối của các bên đưa ra đòi hỏi, hay bất kỳ cơ sở tiềm tàng nào để đòi hỏi dựa trên cơ sở đạo lý hay pháp lý. Chia đôi sự khác biệt trong một cuộc tranh cãi về khoản tiền thuộc không phải là đặc biệt “công bằng,” nhưng nó lại có những giá trị về mặt toán học của bài toán 4.

Nếu chúng ta đặt câu hỏi cái gì quyết định kết cục trong các tình huống nói trên, thì câu trả lời một lần nữa sẽ là vấn đề phối hợp. Mỗi một trong các bài toán này đều đòi hỏi sự phối hợp vì lợi ích chung, mặc dù có sự cạnh tranh giữa các phương án thực hiện hành động chung đó. Nhưng trong nhiều

lựa chọn khác nhau thường xuyên có một hoặc chỉ một vài lựa chọn có thể được dùng làm phương tiện điều phối. Ta hãy xem xét trường hợp về đề xuất đầu tiên trong bài toán 5. Lập luận có sức mạnh nhất cho R chỉ là một câu hỏi tu từ, "Nếu không phải R, thì sẽ là cái gì?". Không có câu trả lời nào rõ ràng hơn là phối hợp một cách ngẫu nhiên, ngay cả khi hai bên đều muốn tránh chữ cái R sau khi đề nghị đầu tiên được đưa ra. Để minh họa cho sức mạnh của luận điểm này, giả sử rằng người chủ trò trong bài toán đó coi đề nghị đầu tiên là đủ để làm hỏng trò chơi và cho rằng anh ta có thể làm cho các đối thủ nhầm lẫn bằng cách tuyên bố đảo ngược các biểu phần thưởng. A sẽ nhận được bất cứ phần thưởng nào mà B sẽ nhận được, và B sẽ nhận được phần thưởng đã nêu ra trong biểu phần thưởng của A trong bài toán 5. Liệu người đưa ra đề nghị ban đầu về R có lý do nào để thay đổi sự lựa chọn của mình hay không? Hoặc, giả sử chủ trò tuyên bố rằng các phần thưởng sẽ là như nhau, bất kể chữ cái nào được chọn, miễn là cả hai bên đều lấy cùng chữ cái. Họ sẽ vẫn hướng về chữ cái R như là phương tiện duy nhất cho sự phối hợp lựa chọn. Nếu chúng ta quay trở lại phần đầu của trò chơi và giả sử rằng không có việc đưa ra đề nghị đầu tiên chọn R, chúng ta có thể tưởng tượng ra ruột bảng hiệu trên tường nói rằng "Trong trường hợp còn nghi ngờ hãy luôn lựa chọn R; bảng hiệu này được mọi người chơi nhìn thấy và là một phương tiện để phối hợp lựa chọn." Ở đây chúng ta quay trở lại ví dụ về hai vợ chồng tại cửa hàng bách hóa, các vấn đề của họ sẽ kết thúc khi họ nhìn thấy một bảng hiệu dễ thấy với dòng chữ "Ban quản lý đề nghị quý khách bị lạc nhau hãy đến quầy thông tin ở giữa tầng trệt." Người ta không thể đòi hỏi cái mà

họ có về nguồn tín hiệu, hoặc về tính hấp dẫn của nó so với những tín hiệu khác mà họ chỉ có thể ước chừng cũng rõ ràng như thế.

Sự trớ trêu sē là hoàn toàn nếu, trong trò chơi 5, đối thủ của bạn biết được biểu phần thưởng của bạn và bạn không biết biểu của anh ta (cũng giống như trường hợp một biến thể của bài toán 5 được dùng trong một số bảng hỏi). Do bạn không có căn cứ nào để phỏng đoán sở thích của anh ta, và thậm chí không thể giúp anh ta, hay đưa ra sự nhượng bộ “vừa phải” nếu bạn muốn, thì cơ sở duy nhất cho sự phối hợp là xem xét thông điệp mà cả hai cùng đọc được trong biểu phần thưởng của mình. Chữ cái mà bạn ưa thích xem ra là sự lựa chọn rõ ràng. Thật khó có thể hiểu lý do tại sao phải chọn chữ cái khác hay sē chọn chữ cái khác nào, bởi vì bạn không có cơ sở để biết chữ cái nào khác sē tốt cho anh ta hơn bản thân chữ R. Hiểu biết của anh ta về sở thích của bạn, kết hợp với việc bạn không biết sở thích của anh ta và việc không có cơ sở nào khác cho sự phối hợp, sē đặt lên vai anh ta trách nhiệm đơn thuần là chọn theo ý muốn của bạn. (Thực tế, đây là kết quả chủ yếu trong mẫu nhỏ đã được thử nghiệm.) Nó có cùng một tình huống như khi chỉ có một người trong hai người nhảy dù biết được người kia ở đâu.<sup>5</sup>

---

<sup>5</sup> Và đây là thêm một ví dụ nữa về sức mạnh nằm trong “điểm yếu” đã được nêu ở một chú thích trước.

## *MẶC CẢ CÔNG KHAI*

Khái niệm “phối hợp” được phát triển ở đây áp dụng cho măc cả ngầm (tacit) có vẻ không trực tiếp thích hợp cho măc cả công khai (explicit). Khi có thể giao tiếp bằng lời nói, thì rõ ràng không cần thiết phải dùng đến liên hệ bằng trực giác, và những manh mối ngẫu nhiên giúp phối hợp các ý nghĩ và gây ảnh hưởng đến kết cục trong trường hợp măc cả ngầm sẽ lại là các tình tiết phụ.

Tuy nhiên có nhiều bằng chứng cho thấy một số ảnh hưởng như vậy có tác động mạnh mẽ ngay cả trong măc cả công khai. Chẳng hạn, trong các cuộc măc cả có liên quan đến các đại lượng số, thì tính hấp dẫn của sự đơn giản toán học lại có sức mạnh. Một ví dụ nhỏ là xu hướng biểu thị các kết quả bằng “các số làm tròn”. Một người bán hàng tính toán số học giá bán “thấp nhất” cho chiếc xe 2.507,63 đô-la sẽ sẵn lòng bớt đi 7,63 đô-la. Việc thường xuyên đạt nhanh được các thỏa thuận cuối cùng nhờ đưa ra đề nghị “chia đôi sự khác biệt” minh họa thêm cho điểm này, và sự khác biệt được chia đôi đó không phải lúc nào cũng nhỏ. Ấn tượng hơn có lẽ là sự xuất hiện khá thường xuyên của các cuộc đàm phán kéo dài về những công thức định lượng phức tạp hay sự chia phần *đặc biệt* về chi phí hay lợi ích cuối cùng quy về những cái rất đơn giản như là chia đều, chia phần tỷ lệ với một số đại lượng chung (tổng sản phẩm quốc gia, dân số, thâm hụt ngoại hối,

v.v...), hay chia phần được thỏa thuận trong các cuộc đàm phán trước nhưng không có liên quan về logic.<sup>6</sup>

Tiền lệ hình như có ảnh hưởng lớn hơn tầm quan trọng về logic hay sức mạnh pháp lý. Giải pháp cho cuộc đình công hay giải pháp cho các khoản nợ quốc tế thường tạo ra một "khuôn mẫu" mà sẽ được tuân theo gần như một cách nghiêm nhiên trong các cuộc đàm phán sau đó. Thêm vào đó, thỉnh thoảng người ta lại cần một thước đo về sự đồng nhất, và đôi khi có sự tương tự cần thiết trong nhiều trường hợp để giải thích cho các kết cục giống nhau vậy. Nhưng ngày càng có vẻ rằng việc đàm phán không còn tính hăng hái nếu như nó bị che phủ bởi những tiền lệ đầy kịch tính và dễ thấy.<sup>7</sup> Theo cách tương tự, các nhà trung gian hòa giải thường chứng tỏ khả năng của mình trong việc đẩy nhanh thỏa thuận và góp ý các điều khoản của thỏa thuận. Các đề xuất của họ dường như thường được chấp nhận ít vì tính công bằng hiển nhiên hay hợp lý của chúng hơn là một cách nhân nhượng nào đó của hai bên tham gia. Các bản báo cáo "Điều tra thực tế" cũng thường dẫn dắt các kỳ vọng về một tiêu điểm, bằng cách đưa ra gợi ý nhằm lấp chân không của sự bất định lẽ ra đã tồn tại: không phải là các thực tế, mà là việc tạo ra một gợi ý cụ thể là cái dường như gây ra ảnh hưởng.

<sup>6</sup> Từ hàng loạt công thức được đề xuất để đóng góp cho UNRRA, thì công thức được chọn là chẵn 1% tổng sản phẩm quốc dân - công thức đơn giản nhất có thể hiểu được và con số được làm tròn dễ được chấp nhận nhất. Chắc chắn rằng, công thức này chính là lập trường ưu tiên của Hoa Kỳ trong cuộc thảo luận. Nhưng thực tế đó có lẽ chứng minh thêm cho ví dụ dã nêu hơn là làm giảm ý nghĩa của nó.

<sup>7</sup> Đoạn này và đoạn trước được minh họa bởi tốc độ đạt được các thỏa thuận về thu nhập từ khai thác dầu mỏ tại Trung Đông nhờ quy về công thức 50-50 một vài năm sau Thế chiến thứ hai.

Theo cách tương tự như vậy, người ta bị thu hút bởi *nguyên trạng* *tiên nghiêm* cũng như các biên giới tự nhiên. Thậm chí các đường vĩ tuyến gần đây cũng cho thấy vai trò lâu đời của chúng trong việc làm tiêu điểm cho các thỏa thuận. Chắc chắn có một số lý do về tính thuận tiện trong việc sử dụng các con sông hay các đường biên lâu đời làm địa điểm tập kết cho quân lính, bất kể sự thích hợp của chúng như thế nào. Nhưng thường thì các đặc điểm địa hình này có vẻ ít có ý nghĩa cho sự tiện lợi thực tế hơn là khả năng đạt được thỏa thuận.

Những nhận xét này sẽ chẳng có gì là quan trọng nếu chúng chỉ mang ý nghĩa rằng các kết quả mặc cả được *diễn đạt* bằng ngôn ngữ đơn giản và định tính, hoặc rằng những nhượng bộ nhỏ được đưa ra để làm tròn một vài xu, một vài dặm, hay một vài người cuối cùng. Nhưng nó thường được coi như thể tiêu điểm cuối cùng cho thỏa thuận không chỉ phản ánh cán cân sức mạnh thương lượng mà còn đem lại thế mạnh đàm phán cho bên này hoặc bên kia. Đường như kẻ đa nghi thường dự đoán kết quả dựa trên cơ sở một tiêu điểm “rõ ràng” nào đó của thỏa thuận, một gợi ý mạnh mẽ nằm trong bản thân tình huống, không liên quan nhiều đến lợi ích của vụ việc, các lập luận được đưa ra, hay các sức ép được áp dụng trong quá trình mặc cả. Cái điểm “rõ ràng” để thỏa hiệp thường đạt được bằng một sự nghiêm nhiên nào đó, như thể là không còn lý do nào xác đáng để giải quyết tại một điểm khác. Hoặc, nếu kết quả “tự nhiên” được chấp nhận để phản ánh kỹ năng tương đối của các bên tham gia mặc cả, thì điều quan trọng là việc nhận ra kỹ năng đó như là năng lực dàn dựng sân khấu làm nổi bật một kết cục nhất định có lợi cho mình. Kết cục đó có thể không đến mức quá công bằng hay

tương xứng một cách rõ ràng với sức mạnh mặc cả dự kiến khi so với sự "dễ thấy" đơn thuần.

Kết luận này dường như thu hẹp phạm vi kỹ năng đàm phán, nếu kết cục đã được xác định bởi cách xác lập bản thân bài toán và vị trí của tiêu điểm. Nhưng điều mà kết luận này muốn rút ra có lẽ là sự thay đổi quỹ tích nơi mà kỹ năng phát huy được tác dụng. Kết quả "dễ thấy" phụ thuộc rất nhiều vào việc bài toán được xác lập như thế nào, vào sự tương tự hay tiền lệ nào mà định nghĩa về vấn đề mặc cả sẽ áp dụng, vào loại dữ liệu nào có thể có để sử dụng cho vấn đề tranh chấp. Khi ủy ban bắt đầu tranh luận làm cách nào để phân chia chi phí, nó đã bị giới hạn bởi điều khoản tham chiếu về hoặc chia sẻ "nợ đáo hạn" hoặc đóng "các khoản thuê," bởi vấn đề liệu cần có một ủy ban phụ chuẩn bị trước các số liệu về thu nhập quốc gia hay về cán cân thanh toán để dùng cho cuộc họp, bởi vấn đề liệu thành viên của ủy ban có nên bật một số tiền lệ nhờ đích thân tham gia vào các cuộc đàm phán trước đó, bởi vấn đề liệu đưa hai đề tài khác biệt vào trong cùng một chương trình nghị sự có làm nổi bật và hợp lý hóa những đặc điểm cụ thể chung hay không. Phần lớn kỹ năng đàm phán đã được áp dụng khi cuộc thương lượng chính thức bắt đầu.<sup>8</sup>

Nếu tất cả điều này là đúng, mà hình như tác giả thường xuyên thấy như vậy, thì sự phân tích của chúng ta về sự mặc

---

<sup>8</sup> Có lẽ một vai trò khác của kỹ xảo nằm trong cách tiếp cận tổng quát này. Nếu ai đó không thành công trong việc xác lập vấn đề sao cho kết quả "hiển nhiên" gần với lập trường mong muốn của mình, anh ta có thể tìm cách làm rối chủ đề. Dưa ra nhiều định nghĩa cho tất cả các thuật ngữ và thêm "nhiều" vào để lấn át tín hiệu mạnh trong cách xác lập ban đầu. Thủ thuật này có thể không thành công, nhưng nó lại thành công trong bài toán biến thể về thuế-thu nhập như đã nói ở trên.

cả ngầm có thể giúp hiểu rõ những nhân tố có ảnh hưởng, và có lẽ logic của mặc cả ngầm thậm chí còn tạo ra cơ sở để tin điều đó là đúng. Vấn đề cơ bản trong mặc cả ngầm là *sự phối hợp*. Từ đó, chúng ta cần biết phải phối hợp những gì trong mặc cả công khai. Câu trả lời có thể là mặc cả công khai đòi hỏi một sự phối hợp kỳ vọng của các bên tham gia nhằm đạt được thỏa thuận cuối cùng. Định đề này có thể được trình bày như sau.

Hầu hết các tình huống mặc cả xét cho cùng bao gồm một loạt các kết cục có thể xảy ra, trong đó mỗi bên thà nhận nhượng còn hơn là không đạt được thỏa thuận nào cả. Trong tình huống đó, bất cứ một kết cục tiềm tàng nào đều là kết cục mà có ít nhất một bên, và có thể cả hai bên, sẽ sẵn sàng nhận nhượng nhằm đi đến thỏa thuận, và thường thì bên kia cũng biết được điều đó. Do đó, mọi kết cục tiềm tàng là kết cục mà mỗi bên đều có thể cải thiện lợi ích của mình bằng cách đòi hỏi thêm. Tuy nhiên anh ta lại có thể không có cơ sở để làm việc đó, vì đối phương biết hoặc nghi ngờ rằng anh ta thà nhận nhượng còn hơn không có thỏa thuận. Chiến lược của mỗi bên được chi phối chủ yếu bởi điều mà anh ta kỳ vọng bên kia sẽ chấp nhận hay kháng khăng đòi hỏi, tuy nhiên mỗi bên đều biết rằng bên kia cũng bị chi phối bởi suy nghĩ qua lại như thế. Kết cục cuối cùng sẽ phải là điểm mà không bên nào kỳ vọng bên kia nhận nhượng tiếp. Tuy nhiên, nhân tố chính của sự kỳ vọng này là điều mà mỗi bên nghĩ bên kia kỳ vọng người đó kỳ vọng gì, và cứ tiếp tục như thế. Bằng cách nào đó, trong tình huống dễ thay đổi và bất định này, mà xem ra không có căn cứ logic để ai đó có thể kỳ vọng vào cái gì khác ngoại trừ điều mà anh ta kỳ vọng bên kia kỳ vọng anh

ta sẽ kỳ vọng, người ta vẫn đi đến được quyết định. Những kỳ vọng phản thân vô tận này bằng cách nào đó phải hội tụ tại một điểm duy nhất, tại đó mỗi bên kỳ vọng bên kia không kỳ vọng được kỳ vọng là sẽ nhân nhượng.

Nếu khi đó chúng ta đặt câu hỏi cái gì sẽ đưa sự kỳ vọng của họ về chỗ hội tụ và đưa cuộc thương lượng đến chỗ kết thúc, chúng ta có thể đề xuất rằng cái đó chính là sức hút tự thân của một số kết cục cụ thể, đặc biệt là những kết cục có tính nổi trội, tính độc nhất, tính đơn giản, có tiền lệ, hay một số lý do hợp lý khác làm cho chúng khác biệt về chất so với một khoảng liên tục các kết quả. Chúng ta có thể lập luận rằng các kỳ vọng thường không hội tụ về một số kết cục chỉ khác về mức độ so với các kết cục khác, mà rằng người ta phải tự chôn chân mình để tỏ lòng quyết tâm. Mỗi bên cần có lý do xác đáng để giữ vững lập trường của mình, và dọc theo khoảng liên tục các lập trường không khác nhau về chất, người ta không tìm ra được một lý do xác đáng nào. Lý do này có thể không đủ mạnh tại "tiêu điểm" tùy ý đó, nhưng ít ra nó cũng có thể tự bảo vệ mình bằng lập luận "Nếu không phải ở đây, thì ở đâu?"

Có lẽ còn cần thêm một ít để có được điểm dừng mà cả hai bên cùng nhận ra. Nếu một bên muốn đưa ra nhân nhượng, anh ta cần phải kiểm soát được các kỳ vọng của đối phương. Anh ta cần có một giới hạn dễ nhận thấy về sự xuống thang của bản thân mình. Nếu một bên muốn đưa ra sự nhân nhượng có giới hạn mà không bị coi là đầu hàng, anh ta cần có một điểm dừng rõ ràng. Một người trung gian có thể đề xuất điểm dừng đó, hay bất kỳ yếu tố nào khác chỉ ra sự khác biệt về chất của lập trường mới này so với các lập trường gần nó. Nếu

một bên đã đưa ra đòi hỏi 60% và rút xuống còn 50%, anh ta có thể đứng vững ở đó. Nếu anh ta rút xuống 49%, thì đối phương sẽ nghĩ rằng anh ta đã rời bỏ vị trí quá nhanh và sẽ còn trượt xuống nữa.

Nếu một vài binh lính rút lui về con sông nằm trên bờ đồi của chúng ta, họ có thể kỳ vọng bên kia nghĩ rằng họ đã xác định một lập trường vững chắc. Đây là điểm họ có thể rút về mà sẽ không làm cho bên kia nghĩ rằng họ sẽ tiếp tục rút lui, trong khi đó, nếu họ rút lui thêm nữa, thì sẽ không còn chỗ nào có thể làm cho đối phương tin rằng họ đã có một lập trường cương quyết. Tương tự như vậy, bên tấn công cũng có thể kỳ vọng ép được bên kia rút lui về chỗ con sông mà không bị đối phương cho rằng sự tiến lên của mình là quá tham lam vì muốn ép đối phương rút lui vô hạn. Như vậy tại bờ sông có một sự ổn định - và có lẽ không còn chỗ nào khác giống như vậy.

Định đề này có vẻ hợp lý theo trực giác. Nó có ý nghĩa như vậy đối với tác giả, và bất cứ sự kiện nào cần đến việc giải thích cho khuynh hướng đi đến thỏa thuận tại các tiêu điểm. Nhưng định đề này vẫn còn mơ hồ và có phần huyền bí, nếu không áp dụng cho logic có phần rõ ràng hơn của mặc cả ngầm. Logic này không chỉ đem đến một sự tương tự mà còn là bằng chứng cho thấy hiện tượng tâm lý cần thiết - sự phối hợp ngầm giữa các kỳ vọng - là một khả năng có thực, và trong một số bối cảnh, là một khả năng rất đáng tin cậy. "Sự phối hợp" các kỳ vọng cũng tương tự như "sự phối hợp" các hành vi khi thông tin liên lạc bị cắt đứt, và, thực tế là cả hai đều đòi hỏi, không hơn không kém, các kỳ vọng lẫn nhau được nhận thức một cách trực giác. Do đó, các kết quả có thể

kiểm nghiệm của một số trò chơi mặc cả ngầm, cũng như vai trò logic hơn của các kỳ vọng được phối hợp với nhau trong những trường hợp đó, chứng minh rằng các kỳ vọng có thể được phối hợp với nhau và một số chi tiết khách quan của tình huống có thể có ảnh hưởng chi phối khi sự phối hợp các kỳ vọng là thiết yếu. Khi không có thông tin liên lạc, *điều gì đó* vẫn được cả hai bên nhận thức. Điều đó vẫn phải nhận thức được, dù không nghi ngờ nữa là với sức mạnh kém hơn, trong trường hợp có thể trao đổi thông tin. Khả năng trao đổi thông tin liên lạc không làm cho giải pháp 50-50 bớt đi tính chất đối xứng, hay làm cho con sông bớt đi tính duy nhất, hay A B C bớt đi tính chất có thứ tự tự nhiên của các chữ cái đó.

Nếu chúng ta lập luận chỉ dựa trên logic của mặc cả ngầm, thì nó chỉ là một phỏng đoán và có lẽ là một phỏng đoán hoang tưởng cho rằng cùng một loại thu hút tâm lý giống như vậy áp dụng được cho mặc cả công khai. Và nếu chúng ta chỉ dựa vào quan sát các kết cục "hợp lý" khác thường của các cuộc mặc cả diễn ra trong thực tế để khái quát hóa, chúng ta có thể sẽ không muốn thừa nhận sức mạnh của các chi tiết mang tính ngẫu nhiên. Nhưng cả hai dòng căn cứ này bổ sung cho nhau chặt chẽ đến mức làm cho sự tương đồng giữa mặc cả ngầm và mặc cả công khai có sức thuyết phục mạnh mẽ.

Để minh họa, ta xem xét bài toán về sự thỏa thuận công khai việc phân chia 100 đô-la: công thức 50-50 hình như là cách phân chia hợp lý, nhưng nó có vẻ như thế vì nhiều lý do khác nhau. Vì cách chia này có vẻ như "công bằng", nó có vẻ làm cân bằng sức mạnh mặc cả, hoặc có thể, như cuốn sách

đã gợi ý, nó đơn giản là có khả năng truyền đạt tính không thể làm thế nào khác của bản thân nó cho hai bên theo cách mà mỗi bên hiểu rằng cả hai cùng biết điều đó. Phân tích của chúng ta về mặc cả ngầm cung cấp bằng chứng cho quan điểm sau. Bằng chứng đó đơn giản là *nếu* hai bên phải phân chia 100 đô-la mà không có sự liên hệ thông tin với nhau, họ sẽ cùng chọn phương án 50:50. Thay vì dựa vào trực giác, chúng ta có thể dựa vào thực tế là trong một bối cảnh không khác biệt lắm - bối cảnh mặc cả ngầm - lập luận của chúng ta có một cách diễn đạt có thể chứng minh được một cách khách quan.

Để minh họa một lần nữa: khả năng nhận thức của hai người chỉ huy trong một bài toán đã nêu của chúng ta về khả năng ổn định hóa của con sông – hoặc ngược lại, sự bất lực của họ không nhận ra được điều đó – được chứng minh là rất quan trọng bởi một thực tế là nếu sự sống còn của họ phụ thuộc vào thỏa thuận về vị trí nhằm ổn định các phòng tuyến và việc trao đổi thông tin là không được phép, thì họ có khả năng nhận ra và hiểu được các phẩm chất của con sông với tư cách là tiêu điểm cho thỏa thuận ngầm giữa họ. Do đó, sự tương đồng ngầm ít nhất cũng chứng tỏ rằng ý tưởng “phối hợp các kỳ vọng” là điều có ý nghĩa hơn là sự bí hiểm.

Có lẽ chúng ta còn có thể đẩy sâu hơn lập luận này. Ngay cả đối với các trường hợp trong đó đặc tính nổi bật duy nhất của một kết quả mặc cả là “sự công bằng” rõ nét của nó, theo tiêu chuẩn được các bên tham gia tỏ ra công nhận, chúng ta vẫn có thể cho rằng sức mạnh đạo lý của sự công bằng sẽ được tăng cường rất nhiều nhờ khả năng quy tụ được sự chú ý của một kết quả “công bằng”, nếu như nó lấp được chán

không của sự bất định lẽ ra có thể có. Tương tự như vậy, khi sức ép của công luận có vẻ như ép buộc các bên đi tới giải pháp “công bằng” hay “hợp lý” rõ rệt, chúng ta vẫn có thể phóng đại “sức ép,” hoặc ít ra là hiểu sai tác động của nó đối với các bên tham gia trừ phi chúng ta ghi nhận sức mạnh của nó trong việc phối hợp các kỳ vọng của các bên. Nói cách khác, chính sức mạnh *gợi ý*, thông qua một cơ chế đã được trình bày trong phần này, làm cho công luận, hay tiền lệ, hay các tiêu chuẩn đạo đức trở nên có hiệu quả như vậy. Như đã nói, để có bằng chứng cho quan điểm này, chúng ta chỉ cần giả thiết rằng các bên tham gia cần phải đạt được thỏa thuận cuối cùng mà không có liên lạc với nhau và nhìn nhận ý kiến công luận hay các tiêu chuẩn đạo đức nổi bật nào đó như là nguồn cung cấp các gợi ý mạnh mẽ, tương tự như những gợi ý được nêu ra trong các ví dụ nói trên. Nhà trung gian hòa giải trong bài toán 7 là một ví dụ gần gũi. Cuối cùng, mặc dù đúng là trách nhiệm đạo lý hay sự nhạy cảm với ý kiến công luận có sức mạnh chi phối các bên tham gia, chứ không phải “tin hiệu” mà họ nhận được, chúng ta vẫn phải nhìn vào nguồn gốc của bản thân công luận, và ở đó, theo ý kiến của tác giả, nhu cầu về một cơ sở hợp lý đơn giản thường phản ánh cơ chế được thảo luận trong phần này.

Nhưng nếu lập luận chung này là đúng, thì bất kỳ sự phân tích nào về mặc cả công khai cũng phải chú ý đến điều mà chúng ta gọi là “trao đổi thông tin” là cái vốn có trong các tình huống mặc cả, những tín hiệu mà các bên tham gia đọc được trong các chi tiết thiêuron động của vụ việc đó. Và điều đó có nghĩa rằng mặc cả ngầm và mặc cả công khai không phải là các khái niệm hoàn toàn tách biệt nhau, mà sự tăng dần về

mức độ, từ sự mặc cả ngầm ngầm lên các dạng trao đổi không hoàn toàn, có khiếm khuyết, hay bị hạn chế đến chổ thông tin liên lạc hoàn toàn, tất cả đều cho thấy sự phụ thuộc ở mức độ nào đó vào sự cần thiết phải kết hợp các kỳ vọng. Do vậy, tất cả cho thấy các bên tham gia đều phụ thuộc đến mức nào đó vào việc họ không có khả năng đồng thời rời mắt khỏi một số kết cục nhất định.

Đây không hẳn là lập luận cho việc trông đợi các kết cục công khai như là một quy tắc làm chổ hướng tới cho các kết quả sẽ phải xuất hiện khi trao đổi thông tin không thể diễn ra. Các tiêu điểm có thể sẽ khác đi khi sự liên hệ bằng lời nói được phép thực hiện, trừ một số trường hợp hư cấu mà chúng ta đã sử dụng để minh họa. Tuy nhiên những nguyên tắc *chính* dùng trong mặc cả ngầm ít nhất cũng là *một* trong các nguyên tắc quan trọng trong phân tích về mặc cả công khai. Hơn nữa, khi mà thậm chí cái hay được gọi là mặc cả "công khai" lại bao gồm thủ đoạn, trao đổi gián tiếp, lập mưu giành ưu thế, tung tin đồn, hay bị làm rối bởi một số lượng người tham gia đông đảo với lợi ích quá phân tán, thì nhu cầu về các kỳ vọng hội tụ và vai trò của các tín hiệu có khả năng phối hợp các kỳ vọng là rất quan trọng.

Có lẽ nhiều hình thức ổn định xã hội và việc hình thành các nhóm lợi ích phản ánh sự phụ thuộc tương tự vào các nhân tố phối hợp mà địa thế và tình huống có thể cung cấp: đoàn nhạc diễu hành tại các hội nghị chính trị thường có thể biến một dấu hiệu đa số yếu ớt nhất thành một đa số áp đảo, sức mạnh của tính hợp hiến được dùng để giành sự ủng hộ của dân chúng trong những thời kỳ bất ổn chính trị và vô chính phủ, sức mạnh huyền thoại của một thủ lĩnh bằng đảng đã

lớn tuổi có thể đem lại trật tự trong thế giới ngầm, đơn giản là vì sự phục tùng phụ thuộc vào kỳ vọng cho rằng người ta sẽ tuân thủ trong việc trừng phạt những kẻ không phục tùng. Khái niệm hay được thể hiện về một “điểm tập hợp quần chúng” trong các hành động xã hội dường như cũng phản ánh khái niệm tương tự. Trong kinh tế học các hiện tượng về lãnh đạo giá, nhiều hình thức cạnh tranh khác không dựa trên giá cả, và thậm chí bản thân sự ổn định giá cả có vẻ như tuân theo một phân tích nhấn mạnh tầm quan trọng của trao đổi ngầm và sự phụ thuộc của nó vào các tín hiệu có thể nhận diện về chất và khá rõ ràng mà có thể đọc được trong bản thân tình huống. Cuộc chính biến “tự phát” có thể phản ánh các nguyên tắc tương tự: khi các nhà lãnh đạo có thể dễ dàng bị lật đổ, người dân đòi hỏi phải có một tín hiệu nào đó cho sự phối hợp, một tín hiệu mà được hiểu một cách không thể nhầm lẫn và có sức mạnh kêu gọi hành động mà mọi người tin chắc rằng mọi người khác cũng nhận được tín hiệu như vậy với đầy đủ sự tự tin để hành động theo đó, nhờ đó cung cấp cho nhau sự miễn trách nhiệm về hành động của đám đông. (Thậm chí còn có khả năng một tín hiệu như vậy là do bên ngoài đưa vào, ngay cả bởi một cá nhân mà khả năng nắm vai trò lãnh đạo duy nhất là khả năng đưa ra các tín hiệu chỉ thị cần có cho một hành động phối hợp.)

## ĐÀM PHÁN NGẦM VÀ CHIẾN TRANH HẠN CHẾ

Dòng phân tích này đem lại cách nhìn hữu ích gì cho các vấn đề thực tiễn về mặc cả ngầm mà chúng ta thường gặp, đặc biệt là các vấn đề về âm mưu chiến lược và chiến tranh hạn chế? Nó chắc chắn gợi ý rằng có thể tìm ra các giới hạn cho chiến tranh - chiến tranh thực sự, chiến tranh pháp lý, hay bất kể chiến tranh nào - mà không có đàm phán công khai. Nhưng nó không hề cho chúng ta một nhận thức mạnh mẽ mới mẻ nào về xác suất. Chiến tranh đã được giới hạn tại Triều Tiên, và khi độc đã không được sử dụng trong Thế chiến thứ hai. Đứng về khả năng chiến tranh hạn chế, hai thực tế này có sức thuyết phục hơn mọi gợi ý được nêu trong phần bàn luận trước. Nếu sự phân tích này đem lại cái gì đó, thì đó không phải là sự phán xét về xác suất đạt được thỏa thuận ngầm mà là sự hiểu biết tốt hơn về chỗ tìm kiếm các điều khoản cho thỏa thuận.

Nếu chúng ta cần rút ra một số kết luận quan trọng thì đó có thể là: (1) Các thỏa thuận ngầm hay các thỏa thuận đạt được thông qua cuộc đàm phán từng phần hay đàm phán tùy nghi đòi hỏi những điều khoản có sự khác biệt về chất so với các phương án khác và không chỉ khác biệt về mức độ. (2) Khi buộc phải đạt được thỏa thuận trong điều kiện liên lạc không đầy đủ, các bên tham gia phải sẵn sàng để cho bản thân tình huống đặt các giới hạn chặt chẽ lên kết cục. Cụ thể hơn, một

đáp án mà có thể bất lợi cho bên này hoặc bên kia hoặc thậm chí gây ra những khó chịu “không cần thiết” cho cả hai bên có thể là giải pháp duy nhất mà có thể phối hợp được các kỳ vọng của họ.

Khí độc đã không được sử dụng trong Thế chiến thứ hai. Thỏa thuận này, mặc dù không phải không có tiền lệ, phần lớn là một thỏa thuận ngầm. Sẽ là điều lý thú khi phỏng đoán liệu có đạt được một thỏa thuận khác về khí độc mà không có sự liên lạc chính thức hay không (hoặc thậm chí, cả trong trường hợp có sự liên lạc chính thức đi nữa). “Một số khí độc” sẽ đặt ra hàng loạt câu hỏi phức tạp về bao nhiêu, ở đâu, trong điều kiện như thế nào. “Không dùng khí độc” là đơn giản và rõ ràng. Khí độc chỉ dùng cho nhân viên quân sự, khí độc chỉ dùng cho lực lượng phòng thủ, khí độc chỉ được vận chuyển bằng xe cơ giới hay dưới dạng đầu đạn, không được dùng khí độc nếu không có cảnh báo. Hàng loạt giới hạn có thể được nêu ra, một số có thể có ý nghĩa, và nhiều giới hạn không hoàn toàn trung tính cho kết cục của cuộc chiến. Nhưng tính chất gián dị của giải pháp “không dùng khí độc” làm cho nó hầu như là tiêu điểm duy nhất cho thỏa thuận khi mà mỗi bên chỉ có thể phỏng đoán bên kia sẽ đề xuất quy tắc nào và khi mà sự thất bại của sự phối hợp đầu tiên này có thể làm hỏng các cơ hội thỏa thuận về bất kỳ giới hạn nào khác.

Hình thể địa lý của Triều Tiên có thể đã giúp cho việc xác định giới hạn chiến tranh và làm cho các giới hạn địa lý thực hiện được. Khu vực này được bao quanh bởi biển, và đường biên chính trị chủ yếu phía bắc được đánh dấu nổi bật và không thể nhầm lẫn bởi một con sông. Vĩ tuyến 38 đường như

là một tiêu điểm có sức mạnh gây ra tình trạng bế tắc, và giải pháp chủ yếu, cái "eo," là một ứng cử viên có nhiều triển vọng không chỉ vì nó cho một tuyến phòng thủ ngắn hơn mà còn vì lý do có thể cả hai bên đều hiểu rõ việc tiến quân về eo đất này không hẳn là tín hiệu về một quyết tâm muốn tiến xa hơn, và việc rút lui về cái eo đất này cũng không chuyển di tín hiệu về mọi ý định rút lui thêm nữa.

Eo biển Đài Loan giúp cho việc ổn định biên giới giữa hai phía đối lập ở Trung Quốc, không chỉ vì biển có lợi cho bên phòng thủ và ngăn cản tấn công, mà còn vì hòn đảo là một đơn vị hoàn chỉnh và biển là biên giới rõ ràng. Sự hy sinh bất kỳ phần nào trên hòn đảo sẽ làm cho đường biên giới trở nên không ổn định, và việc giữ lại bất kỳ mảnh đất nào trên lục địa cũng sẽ gây ra sự bất ổn định tương tự. Ngoại trừ tại bờ biển, mọi di chuyển chỉ khác nhau về mức độ, và một cuộc tấn công qua eo biển là sự tuyên bố rằng "thỏa thuận" đã bị hủy bỏ.

Tại Triều Tiên, vũ khí bị giới hạn bởi sự phân biệt về chất giữa vũ khí nguyên tử và tất cả các vũ khí khác còn lại. Chắc chắn sẽ khó khăn hơn rất nhiều nếu muốn giữ ổn định cho một thỏa thuận ngầm về giới hạn nào đó đối với quy mô vũ khí nguyên tử hay sự lựa chọn các mục tiêu đánh phá.<sup>9</sup> Không có định nghĩa về quy mô hay mục tiêu nào rõ ràng và tự nhiên mà không cần phải nói ra, trừ khi nói rằng "không có quy mô nào, không vào bất kỳ mục tiêu nào." Viện trợ của Mỹ cho quân đội của Pháp tại Đông Dương được giới hạn rõ ràng là với thiết bị chứ không phải con người, và nó được hiểu rằng một sự mở rộng, chẳng hạn sự tham gia của không quân có

<sup>9</sup> Điểm này sẽ được trình bày kỹ hơn ở Phụ lục A.

thể được coi là giới hạn trong phạm vi trên không, trong khi không thể thiết lập một *số lượng* giới hạn tham gia của không lực hay bộ binh. Những ý định của ai đó tránh sự can thiệp trên bộ có thể được truyền đạt bằng việc không đưa một lực lượng bộ binh nào vào tham chiến. Người ta không thể dễ dàng cam kết một số lực lượng *nào đó* và tuyên bố một giới hạn có tính thuyết phục về *số lượng* mà người đó có ý định cam kết.

Chiến lược trả đũa chịu ảnh hưởng bởi sự cần thiết phải có trao đổi hay phối hợp về các giới hạn. Cuộc xâm lược lãnh thổ sẽ xác định vị trí. Với may mắn và các biên giới tự nhiên, có thể có sự chấp nhận ngầm về các giới hạn địa lý hay giới hạn về các loại mục tiêu. Một bên hoặc cả hai bên có thể sẵn sàng chấp nhận sự thất bại có giới hạn hơn là tự ý phá vỡ quy tắc, và có hành động để khẳng định với bên kia về ý định đó của mình. “Những quy tắc” có thể được tôn trọng bởi vì nếu chúng bị phá vỡ một lần sẽ không có gì để đảm bảo sẽ tìm được những quy tắc mới và sẽ được các bên cùng công nhận kịp thời để ngăn chặn sự lan rộng của cuộc xung đột. Nhưng nếu sự trả đũa hoàn toàn do bên trả đũa lựa chọn về phương pháp và địa điểm tiến hành, thì có thể sẽ khó khăn hơn rất nhiều trong việc chuyển đến bên bị trả đũa các giới hạn mình đề nghị, để anh ta có cơ hội để chấp nhận điều đó qua sự phản công của anh ta. Thực tế là, việc tránh trả đũa ngay từ đầu khỏi địa điểm đã kích thích nó có thể là một kiểu tuyên bố độc lập không giúp cho việc hình thành những kỳ vọng chung ổn định. Do đó vấn đề tìm kiếm các giới hạn chiến tranh bị nghi ngờ là khó thực hiện nếu như định nghĩa hàm chứa trong hành động của bên xâm lược không được chấp nhận.

Tóm lại, vấn đề giới hạn chiến tranh không liên quan đến một khoảng liên tục các khả năng có thể có từ khả năng có lợi nhất cho đến bất lợi nhất cho mỗi bên. Đó là một thế giới không đồng nhất và rời rạc, mà dễ nhận ra sự khác biệt về định tính hơn là về định lượng, một thế giới rắc rối vì có quá nhiều lựa chọn, và một thế giới buộc hai bên phải chấp nhận một số sự áp đặt bởi các nhân tố của nó. Tác giả cho rằng điều này cũng đúng với mọi cuộc cạnh tranh có giới hạn trong mọi lĩnh vực.

## *NHỮNG DÀN XẾP TRƯỚC*

Trong khi ý chính của phần này nói rằng mặc cả ngầm là điều có thể thực hiện được và có thể phân tích hệ thống, thì không có gì bảo đảm rằng nó sẽ thành công trong các trường hợp cụ thể, hoặc nếu thành công thì chưa chắc nó sẽ đem lại kết quả có lợi cho các bên tham gia so với các phương án khác khi mà sự thông tin liên lạc được phép thực hiện đầy đủ. Không có gì bảo đảm rằng cuộc đại chiến mới, nếu xảy ra, sẽ tìm được các giới hạn được các bên chấp nhận một cách kịp thời và có khả năng bảo vệ được nó, nếu thương lượng công khai không được thực hiện. Do đó người ta có lý do để xem xét các bước có thể được thực hiện trước khi mặc cả ngầm xảy ra, để nâng cao khả năng có một kết quả thành công.

Việc duy trì các kênh trao đổi thông tin có vẻ như là một cách khá rõ ràng. (Về tối thiểu, đây có thể là sự bảo đảm rằng

một đề nghị đầu hàng có thể đến tai đối phương và được đáp lại.) Khía cạnh kỹ thuật của nguyên tắc này sẽ là việc xác định ai là người gửi và nhận thông điệp, dựa vào thẩm quyền nào, bằng phương tiện gì, sử dụng trung gian nào nếu cần đến trung gian, ai sẽ đứng trong hàng ngũ để thực thi nhiệm vụ theo cách thức nào nếu các bên và phương tiện theo chỉ định bị triệt tiêu. Trong trường hợp nỗ lực tiến hành một cuộc chiến tranh hạt nhân có giới hạn, có thể chỉ có một khoảnh khắc ngắn ngủi và bận rộn trong đó mỗi bên phải xét xem liệu cuộc chiến tranh có giới hạn đang trở thành cuộc chiến tranh toàn diện hay là một cuộc chiến tranh đầy đủ đã diễn ra rồi. Và sự lúng túng trong 12 giờ đầu tiên về cách làm thế nào để tiếp xúc với nhau có thể tiêu hủy một số cơ hội ổn định hóa các hành động trong phạm vi giới hạn.

Cần có sự quan tâm dành cho ích lợi của việc sử dụng trung gian và trọng tài. Để có được thỏa thuận trong việc chọn ra các trung gian có ảnh hưởng, người ta thường phải có sự hiểu biết từ trước, hoặc ít ra cũng có một tiền lệ, hay một truyền thống, hay một dấu hiệu của sự hoan nghênh. Thậm chí, nếu chúng ta loại trừ sự dàn xếp công khai cho tình huống, sự chứng tỏ của mỗi bên về việc chấp nhận vai trò của các trọng tài hay trung gian, mặc dù việc sử dụng họ chỉ đem lại lợi ích nhỏ nhoi, có thể giúp cho việc tạo ra một công cụ cực kỳ có giá trị trong một tình huống tồi tệ.

Nhưng tất cả những cố gắng như vậy có thể gặp khó khăn nếu một bên đối kháng không muốn tham gia các bước chuẩn bị. Không những một đối thủ có thể lẩn tránh việc đưa ra tín hiệu hăm hở muốn đi đến thỏa thuận, mà còn có thể một bên

trong cuộc chiến tranh tiêm tàng lại có lợi ích chiến thuật trong việc không kiềm chế chiến tranh và làm nghiêm trọng thêm khả năng tàn phá lẫn nhau khi chiến tranh nổ ra. Tại sao như vậy? Đó là vì chiến lược đe dọa, lừa gạt, và răn đe. Ý muốn phát động chiến tranh hay có những bước đi có thể dẫn đến chiến tranh, bất kể là sự xâm lược hay chống lại xâm lược, có thể phụ thuộc vào sự tự tin mà các nhà lãnh đạo quốc gia nghĩ là cuộc chiến có thể được kiểm soát trong giới hạn. Nói cụ thể hơn, sự sẵn sàng của Mỹ trong việc đáp trả bất kỳ sự xâm lược cục bộ nào bằng vũ khí nguyên tử sẽ phụ thuộc vào - và Nga cũng biết điều này - việc Mỹ chắc chắn đến đâu khi nghĩ rằng sự trả đũa đó có thể tự nó nằm trong giới hạn. Nghĩa là, nó phụ thuộc vào sự chắc chắn trong phán xét mà Mỹ và người Nga, khi cả hai đang rất cần biết được giới hạn trong đó mỗi bên sẵn sàng chấp nhận thua trong cuộc chiến mà không mở rộng giới hạn đó, sẽ cùng tìm ra những giới hạn đó và đi đến bằng lòng với chúng. Nếu khi đó, người Nga từ chối tham gia vào bất cứ hoạt động nào có thể dẫn đến khả năng của một cuộc chiến tranh hạn chế thì điều đó ngăn cản quyết tâm hành động của chính người Mỹ, họ có thể mạo hiểm mất đi những giới hạn như vậy, những giới hạn làm giảm sự đe dọa về hành động của Mỹ. Một trong hai người nhảy dù trong ví dụ của chúng ta có thể biết rằng người kia sẽ bắt cẩn với máy bay nếu anh ta tin chắc rằng họ sẽ gặp nhau và tự cứu được mình. Do vậy nếu người thứ nhất lảng tránh thảo luận về tình huống này, thì người thứ hai sẽ phải ngồi yên trên máy bay vì sợ gây ra sự lạc nhau chết người nếu rơi xuống địa phận bên dưới.

Bất kể vì lý do này hay chỉ vì những ngẩn cản thường thấy trong đàm phán nghiêm túc làm cho việc trao đổi trước không thể diễn ra, thì vẫn có một ý tưởng bối rối rút ra từ một trong các trò chơi đã nêu của chúng ta. Đó là cuộc đàm phán hay trao đổi với mục đích phối hợp các kỳ vọng không cần phải có sự qua lại: sự đàm phán đơn phương có thể đem lại sự phối hợp mà sẽ cứu giúp cả hai bên. Hơn nữa, ngay cả một bên không muốn sẽ không nhất thiết tự đặt anh ta vào vị trí mà không thể tiếp nhận được thông điệp. Hãy nhớ lại người đưa ra đề nghị chữ cái R trong một trò chơi mặc cả: chừng nào mà đối tác của anh ta vẫn nghe thấy - và rõ ràng là anh ta đã nghe thấy - chữ cái R là đề xuất duy nhất tồn tại, và không bị thách thức, thì đề xuất đó vẫn có thể phối hợp một cách đương nhiên bất kỳ phản đề nghị nào gần được cũng như đã được chấp nhận công khai. (Thậm chí sự từ chối của bên kia cũng không làm mất đi tính nổi trội mà còn đơn giả là chứng tỏ anh ta biết được điều đó, miễn là đối phương không đưa ra một phản đề nghị có thể gây ra sự bất định.) Nếu một trong hai người nhảy dù, ngay trước khi máy bay gấp sự cố và trong khi chưa có ai nghĩ đến việc phải nhảy dù, nói vu vơ rằng "Nếu tôi phải gấp ai đó ở dưới kia, tôi sẽ hướng đến ngọn đồi cao nhất có thể thấy được," thì người kia có thể nhớ lại câu nói này và hiểu rằng người thứ nhất cũng sẽ tin chắc là anh ta sẽ nhớ ra và sẽ tìm đến địa điểm đó, mặc dù lúc máy bay gấp sự cố anh ta thốt lên "Sao đại dột thế nhỉ," hay là "Tôi không leo đồi đâu, đau chân lắm." Khi có một số tín hiệu được cả hai bên rất cần đến và cả hai bên đều biết điều đó, thậm chí một tín hiệu yếu ớt và thiên vị cũng có thể bắt buộc các bên công

nhận, khi không còn tín hiệu nào khác. Khi tình huống bất ngờ ập đến trước mặt họ, thì những lợi ích của họ, mà ban đầu rất khác nhau trong trò chơi đe dọa và răn đe, sẽ trở nên trùng hợp về căn bản vì nhu cầu bức thiết có được một tiêu điểm cho thỏa thuận.

**PHẦN II**

**ĐỊNH HƯỚNG LẠI  
LÝ THUYẾT  
TRÒ CHƠI**



## 4

# TIẾN TỚI MỘT LÝ THUYẾT VỀ QUYẾT ĐỊNH PHỤ THUỘC LÃN NHAU

Về chiến lược xung đột thuần túy - các trò chơi có tổng lợi ích bằng không (zero-sum game) - lý thuyết trò chơi (game theory) đã đem lại nhiều hiểu biết và gợi ý quan trọng. Nhưng về chiến lược hành động nơi mà xung đột trộn lẫn với sự phụ thuộc lẫn nhau - các trò chơi có tổng lợi ích khác không (nonzero-sum game) bao gồm chiến tranh và đe dọa gây chiến tranh, đình công, thương lượng, răn đe tội phạm, chiến tranh giai cấp, chiến tranh chủng tộc, chiến tranh giá cả, và tổng tiền, sự xoay xở trong bộ máy quan liêu hay trong tắc nghẽn giao thông, và sự cưỡng ép con cái của mình - thì lý thuyết trò chơi truyền thống chưa đem lại được những hiểu biết hay gợi ý tương ứng. Đây là những "trò chơi" mà trong đó, mặc dù nhân tố của xung đột đem lại mối quan tâm sâu sắc, sự phụ thuộc lẫn nhau là một bộ phận của cơ cấu logic và đòi hỏi một số hình thức hợp tác hay nhượng bộ qua lại - đó là hình thức ngầm, nếu không phải là công khai - cho dù chỉ để tránh sự tàn phá lẫn nhau. Đây cũng là các trò chơi trong đó,

mặc dù tính chất bí mật có thể đóng một vai trò chiến lược, có nhu cầu thiết yếu đưa ra tín hiệu về các ý định và sự gắp nhau của các ý nghĩ. Cuối cùng, chúng là những trò chơi mà trong đó cái mà một người chơi có thể làm để tránh tổn thất chung sẽ có ảnh hưởng đến cái mà người chơi khác sẽ làm để tránh tổn thất đó, do đó không phải lúc nào cũng là lợi thế khi có được thế chủ động, kiến thức, hay sự tự do lựa chọn.

Lý thuyết trò chơi truyền thống, xét về phần lớn, đã vận dụng vào các trò chơi phụ thuộc lẫn nhau này (các trò chơi có tổng lợi ích khác không) các phương pháp và khái niệm đã được chứng minh là thành công trong việc nghiên cứu chiến lược xung đột thuần túy. Chương này và chương tiếp theo cố gắng mở rộng phạm vi của lý thuyết trò chơi, lấy trò chơi có tổng lợi ích bằng không làm một trường hợp giới hạn hơn là một điểm xuất phát. Sự mở rộng của lý thuyết này sẽ đi theo hai tuyến chính. Một là nhận diện yếu tố linh hôi và gợi ý trong việc hình thành những kỳ vọng nhất quán tương hỗ. Hai là (tình bày trong chương tiếp theo) nhận diện “các nước đi” cơ bản có thể xuất hiện trong các trò chơi chiến lược thực sự, và các yếu tố cấu trúc mà những nước đi này lấy làm chỗ dựa. Nó sẽ bao gồm các khái niệm như là “sự đe dọa,” “đảm bảo thực thi,” và khả năng trao đổi hay phá vỡ liên lạc.

Việc lý thuyết trò chơi đó kém phát triển theo hai tuyến này có thể phản ánh sự tập trung của nó vào trò chơi có tổng lợi ích bằng không. Những gợi ý và suy diễn, đe dọa và hứa hẹn không có tầm quan trọng nào trong lý thuyết đã được thừa nhận về trò chơi có tổng lợi ích bằng không. Các yếu tố đó không quan trọng bởi vì chúng nói lên mối liên hệ giữa hai người chơi mà, trừ phi hoàn toàn vô hại, phải là bất lợi cho

một bên, và anh ta có thể hủy bỏ nó bằng cách áp dụng một chiến lược tối thiểu hóa tối đa (minimax), được dựa trên một cơ chế gây ngẫu nhiên, nếu cần thiết. Do vậy “các chiến lược hợp lý” được hai bên theo đuổi trong một tình huống của xung đột thuần túy - như đã được diễn hình hóa bởi sự đuổi bắt và lẩn tránh - không thể được trông mong sẽ làm rõ được loại hành vi nào giúp ích cho sự nhượng bộ qua lại, hay là làm cách nào để khai thác sự phụ thuộc lẫn nhau phục vụ cho lợi ích đơn phương.

Nếu trò chơi có tổng lợi ích bằng không là trường hợp giới hạn của xung đột thuần túy, thì thái cực kia là gì? Nó phải là trò chơi “hợp tác thuần túy” trong đó những người chơi cùng nhau được hoặc thua, và có sở thích (preference) giống hệt nhau đối với kết cục của trò chơi. Bất kể họ giành được các phần cố định trong tổng số hay các phần thay đổi theo tổng số chung, họ phải xếp hạng tất cả các kết cục có thể xảy ra theo một cách hoàn toàn giống nhau, theo thang sở thích của từng người. (Và, để tránh mọi xung đột từ đầu, thì những người chơi phải thấy rõ ràng sở thích của họ là hoàn toàn giống nhau, để không có xung đột lợi ích khi hai bên trao đổi thông tin, dù là thông tin đúng hay thông tin sai.)

Cái gì về hợp tác thuần túy làm cho nó có liên quan đến lý thuyết trò chơi hay mặc cả? Câu trả lời không đầy đủ, chỉ để nói rằng đây không phải là trò chơi tầm thường, là nó có thể chứa đựng những vấn đề về nhận thức và trao đổi thông tin thuộc những dạng mà rất thường xuất hiện trong các trò chơi có tổng lợi ích khác không. Mỗi khi cơ cấu trao đổi thông tin không cho phép các bên tham gia phân chia công việc trước theo một kế hoạch công khai, thì việc phối hợp hành vi trong

quá trình của trò chơi sẽ không dễ dàng. Những người chơi phải hiểu biết lẫn nhau, tìm thấy những khuôn mẫu hành vi cá nhân của nhau để có thể dự đoán được hành động của mỗi bên. Họ phải thử nghiệm lẫn nhau về ý thức chung về khuôn mẫu hay tính đêùn đắn và biết khai thác những sáo ngữ, quy ước, và những ký hiệu ngẫu hứng để phát tín hiệu về những ý định của mình và đáp lại tín hiệu của nhau. Họ phải trao đổi với nhau bằng các gợi ý và bằng hành vi gợi ý. Hai xe cố gắng tránh đâm nhau, hai người cùng nhau khiêu vũ theo một bản nhạc lạ lẫm, hay những thành viên của một đội du kích bị lạc nhau trong chiến đấu phải phối hợp ý định của nhau theo cách này, cũng giống như những khán giả vỗ tay trong buổi hòa nhạc, những người mà tại một thời điểm nào đó “tán thành” về việc liệu có nên cùng vỗ tay mạnh để yêu cầu diễn lại hay cùng vỗ tay yếu dần.

Nếu *cờ vua* là một ví dụ chuẩn cho một trò chơi có tổng lợi ích bằng không, thì *trò chơi đố chữ* (charade) có thể làm điển hình cho trò chơi phối hợp thuần túy. Nếu *theo đuổi* là hình ảnh thu nhỏ của trò chơi có tổng lợi ích bằng không, thì *hẹn hò* cũng là như thế đối với trò chơi phối hợp.

Thí nghiệm của O. K. Moore và M. I. Berkowitz cho ta một sự kết hợp lý thú trong đó cả hai trường hợp giới hạn đều dễ nhận thấy.<sup>1</sup> Nó bao gồm một trò chơi có tổng lợi ích bằng không giữa hai đội, mỗi đội gồm ba người. Ba người của một đội có những lợi ích giống nhau nhưng, nhưng do tính chất

<sup>1</sup>O. K. Moore và M. I. Berkowitz, *Lý Thuyết Trò Chơi và Tương Tác Xã Hội* (Game Theory and Social Interaction), Văn phòng Nghiên cứu Hải Quân, Báo cáo Kỹ thuật, Hợp đồng Số SAR/NONR-609 (16) (New Haven, tháng 11-1956).

đặc biệt của trò chơi, họ không thể hành động như một thể thống nhất. Tính chất đặc biệt này là ba người của mỗi đội bị cách ly nhau và chỉ có thể liên hệ với nhau bằng điện thoại, và tất cả sáu máy điện thoại đều được kết nối trên cùng một đường dây, do đó mọi người có thể nghe được các thành viên của đội mình và đội khác nói chuyện với nhau. Các quy ước trước về mã trao đổi là không được phép. Ở đây, giữa hai đội chúng ta sẽ có trò chơi xung đột thuần túy, và giữa các đội viên với nhau chúng ta có trò chơi phối hợp thuần túy.

Nếu trong trò chơi này chúng ta bỏ đi "đội kia" và nếu ba người chơi chỉ đơn giản tìm cách phối hợp một chiến lược giành chiến thắng trong trò chơi về kỹ năng hay may rủi khi có khó khăn trong trao đổi, chúng ta sẽ có một trò chơi phối hợp thuần túy gồm ba người. Một vài "trò chơi" thuộc dạng này đã được nghiên cứu, cả bằng thực nghiệm và lý thuyết. Trong thực tế, có sự trùng lặp đáng kể tại điểm này giữa trò chơi có tổng lợi ích khác không và lý thuyết tổ chức hay lý thuyết trao đổi thông tin.<sup>2</sup>

---

<sup>2</sup> Một phân tích lý thuyết có chiều rộng về vấn đề phối hợp đã được Jacob Marschak thực hiện trong "Các Yếu Tố của Lý Thuyết Nhóm" (*Elements for a Theory of Team*) và "Tiến Đến một Lý Thuyết Kinh Tế về Tổ Chức và Thông Tin" (*Toward an Economic Theory of Organization and Information*), *Cowles Foundation Discussion Papers* (Báo cáo tham luận của Quỹ tài trợ Cowles), số 94 và 95 (loạt mới), và cùng với Roy Radner trong "Các Vấn Đề Cấu Trúc và Trao Đổi Vận Hành trong Các Nhóm" (*Structural and Operational Communication Problems in Teams*), *Cowles Foundation Discussion Papers, Economics*, số 2076. Các ví dụ về nghiên cứu thực nghiệm có liên quan có thể được tìm trong "Các Mô Hình Trao Đổi Thông Tin trong Các Nhóm Công Tác" (*Communication Patterns in Task-oriented Groups*) của Alex Bavelas; *Tính Động của Nhóm* (*Group Dynamics*) (Evanston, 1953) của D. Cartwright và A. F. Zander; "Giải Quyết Vấn Đề của Các Nhóm Nhỏ bằng Cách Sử Dụng Các Mạng

Các thí nghiệm được nêu trong Chương 3 cho thấy sự lựa chọn có phối hợp là có thể ngay cả trong trường hợp hoàn toàn không có sự trao đổi. Hơn nữa, chúng còn cho thấy rằng ó các tình huống mặc cả ngầm trong đó xung đột lợi ích trong việc lựa chọn hành động có thể bị áp đảo bởi sự cần thiết đơn thuần cho sự phối hợp một số hành động, và trong các tình huống đó, trường hợp giới hạn về phối hợp thuần túy sẽ có lập tính chất cốt lõi của trò chơi có tổng lợi ích khác không tương ứng.

Như vậy trong *việc giải quyết bài toán có sự phối hợp*, chúng ta sẽ có, với sự phụ thuộc của nó vào việc truyền tải và tiếp nhận các ý định hay kế hoạch, một hiện tượng làm nổi bật khía cạnh cốt lõi của trò chơi có tổng lợi ích khác không, và nó sẽ có mối quan hệ với trò chơi này chẳng khác đối với trò chơi có tổng lợi ích bằng không, tức là “trường hợp giới hạn.” Một cái là trò chơi hỗn hợp xung đột-hợp tác với tất cả phạm vi cho sự hợp tác bị loại bỏ. Cái kia là trò chơi hỗn hợp xung đột-hợp tác với sự xung đột bị loại bỏ. Trong trò chơi này phần thưởng dựa vào việc giữ bí mật, còn trong trò chơi kia, phần thưởng dựa vào việc để lộ thông tin.

---

Tuong Linh Khac Nhau” (Problem Solving by Small Groups Using Various Communication Nets) của G. A. Heise và G. A. Miller; *Các Nhóm Nhỏ* (Small Groups) (New York, 1955) của P. A. Hare, E. F. Borgatta và R. F. Bales; “Một Số Ảnh Hưởng của Phản Hồi lên Trao Đổi Thông Tin” (Some Effects of Feedback on Communication) của H. J. Leavitt và R. A. Mueller trong *Các Nhóm Nhỏ* (Small Groups); và “Một Nghiên cứu Thủ Nghiêm về Ảnh Hưởng của Ngôn Ngữ lên Tái Tạo Hình Thái Nhận Thức về Thị Giác” (An Experimental Study of the Effects of Language on the Reproduction of Visually Perceived Form) của L. Carmichael, H. P. Hogan và A. A. Walter, *Journal of Experimental Psychology*, 15:73-86 (tháng 2-1932).

Cần nhấn mạnh rằng trò chơi phối hợp thuần túy là *một trò chơi về chiến lược* theo ý nghĩa thuần túy kỹ thuật. Đó là một tình huống hành vi trong đó sự lựa chọn tối ưu cho hành động của mỗi bên phụ thuộc vào hành động mà anh ta kỳ vọng bên kia sẽ thực hiện, mà anh ta biết rằng nó lại phụ thuộc vào kỳ vọng của bên kia đối với hành động của anh ta. Sự phụ thuộc lẫn nhau của các kỳ vọng này chính là cái phân biệt trò chơi chiến lược với trò chơi may rủi hay trò chơi kỹ năng. Trong trò chơi phối hợp thuần túy các lợi ích là hội tụ. Trong trò chơi xung đột thuần túy các lợi ích là phân tán. Nhưng không trong một trò chơi nào, sự lựa chọn hành động diễn ra một cách khôn ngoan mà không xét đến sự phụ thuộc của kết cục vào những kỳ vọng tương hỗ của người chơi.<sup>3</sup>

Nhớ lại trường hợp nổi tiếng của Holmes và Moriarty trên các chuyến tàu khác nhau, không ai tiếp xúc trực tiếp với ai, mỗi người phải lựa chọn có xuống tàu tại ga tới hay không.

---

<sup>3</sup> Về điểm này, Carl Kaysen trong bài viết của ông về cuốn *Lý Thuyết về Trò Chơi và Hành Vi Kinh Tế* (Theory of Games and Economic Behavior) của Von Neumann và Morgenstern có nói “Lý thuyết các trò chơi chiến lược đó nói về các hành động của nhiều tác nhân, trong một tình huống mà tất cả các hành động phụ thuộc lẫn nhau và nói chung tại nơi mà không có khả năng nào cho cái mà chúng ta gọi là thông số hóa (parametrization) cho phép mỗi tác nhân (người chơi) hành xử như thể hành động của đối phương đã được kể đến. Trong thực tế, chính việc không có sự thông số hóa này mới là bản chất của trò chơi.” Cách nói tương tự đã được R. Duncan Luce và Howard Raiffa sử dụng trong cuốn *Trò Chơi và Các Quyết Định* (Games and Decisions) (New York, 1957): “Nói một cách trực giác, vẫn đề xung đột lợi ích, đối với người tham gia, là vấn đề đưa ra quyết định cá nhân trước hồn hợp giữa rủi ro và bất định, những bất định xuất phát từ sự không hiểu biết của anh ta về cái mà bên kia sẽ hành động” (trang 14). Tuy nhiên, sự chú ý của họ tập trung vào cuộc xung đột; trường hợp có sự trùng hợp trong sở thích thì họ loại bỏ xem như là tầm thường (trang 59, 88), và họ coi những người chơi đó như là một cá nhân (trang 13).

Chúng ta có thể xem xét ba loại lợi ích (payoff). Thứ nhất, Holmes sẽ được giải thưởng nếu hai người xuống tàu tại các ga khác nhau, Moriarty sẽ thắng nếu hai người xuống tàu cùng một ga. Đây là trò chơi có tổng lợi ích bằng không, trong đó sở thích của hai người chơi có mối tương quan hoàn toàn ngược nhau. Trường hợp thứ hai, Holmes và Moriarty cả hai sẽ được thưởng nếu họ xuống tàu tại cùng một ga bất kỳ nào. Đây là trò chơi phối hợp thuần túy, trong đó có sự tương quan thuận giữa các sở thích của hai người chơi. Trường hợp thứ ba là cả Holmes và Moriarty sẽ được thưởng nếu họ xuống tàu tại cùng một ga, nhưng Holmes được thưởng nhiều hơn nếu cả anh ta và Moriarty cùng xuống tàu tại một ga nhất định, Moriarty được thưởng nhiều hơn nếu cả hai cùng xuống tàu tại một ga nhất định khác, cả hai sẽ thua cuộc nếu họ không xuống tàu cùng một ga. Đây là trò chơi có tổng lợi ích khác không thường thấy, hay còn gọi là trò chơi “tương quan không hoàn toàn giữa các sở thích.” Đây là sự kết hợp giữa xung đột và phụ thuộc lẫn nhau đại diện cho các tình huống mặc cả. Bằng cách xác lập các hệ thống liên lạc và thông tin cụ thể cho các bên tham gia, chúng ta có thể làm phong phú thêm trò chơi hoặc làm cho nó trở nên tầm thường hoặc đưa thêm một lợi thế cho một bên trong các biến thể thứ nhất và thứ ba.

Yếu tố cốt lõi của trò chơi chiến lược có mặt trong cả ba trường hợp nói trên: sự lựa chọn tối ưu cho mỗi bên phụ thuộc vào điều mà anh ta kỳ vọng bên kia sẽ lựa chọn, với sự hiểu biết rằng bên kia cũng được chỉ dẫn tương tự, sao cho mỗi bên đều hiểu rằng mình phải cố gắng phán đoán cái mà người thứ hai phán đoán người thứ nhất sẽ phán đoán người thứ hai sẽ phán đoán và cứ tiếp tục như vậy, theo đường xoắn ốc quen thuộc của các kỳ vọng tương hỗ.

## *MỘT CÁCH PHÂN LOẠI LẠI CÁC TRÒ CHƠI*

Trước khi đi sâu hơn, chúng ta có thể phân loại lại các tình huống trò chơi cho có ích hơn. Việc chia đôi các trò chơi thành hai loại có tổng lợi ích bằng không và tổng lợi ích khác không làm mất đi tính chất đối xứng mà chúng ta cần có và không nhận diện được trường hợp giới hạn tương phản với trò chơi có tổng lợi ích bằng không. Các điểm cốt lõi của một cách phân loại trò chơi có hai người có thể được biểu diễn trên một đồ thị hai chiều. Giá trị của bất cứ kết quả cụ thể nào của trò chơi, dành cho hai người, sẽ được thể hiện bằng hai tọa độ của một điểm. Mọi kết quả có thể có của một trò chơi xung đột thuần túy sẽ được biểu thị bởi một số hay tất cả các điểm trên một đường có độ dốc âm, mọi kết quả của trò chơi với lợi ích chung thuần túy được biểu thị bởi một số hoặc tất cả các điểm nằm trên một đường có độ dốc dương. Trong trò chơi hỗn hợp, hay tình huống mặc cả, có ít nhất một cặp điểm chỉ ra độ dốc âm và ít nhất một cặp chỉ ra độ dốc dương.<sup>4</sup>

---

<sup>4</sup> Nếu bản chất của trò chơi làm cho người chơi thích sử dụng công cụ ngẫu nhiên trong việc chọn chiến lược của mình, hoặc làm cho khả thi rằng họ đảm phán một thỏa thuận có thể cưỡng chế được mà thực tế phụ thuộc vào một cơ chế ngẫu nhiên, như là rút thăm chặng hạn, thì sẽ có cơ hội cho sự hợp tác trong việc lựa chọn các chiến lược ngay cả khi có sự hoàn toàn bất đồng về thứ hạng của các kết cục. Trong trường hợp đó các điểm biểu thị trò chơi xung đột thuần túy phải thỏa mãn sự hạn chế nghiêm ngặt hơn tức là phải nằm trên một đường thẳng, với hai trục đo lường "độ thỏa dung" của người chơi theo nghĩa mà nay rất quen thuộc trong lý thuyết trò chơi. Sự hạn chế này cũng áp dụng cho

Chúng ta có thể giữ nguyên thuật ngữ truyền thống, đối với các trò chơi hoàn toàn thuần túy, bằng cách gọi chúng là các trò chơi có *tổng lợi ích cố định* và các trò chơi có *tỷ lệ lợi ích cố định*, lấy những thuật ngữ dài dòng như *tổng lợi ích biến đổi-tỷ lệ lợi ích biến đổi* để đặt tên cho tất cả các trò chơi ngoại trừ các trường hợp giới hạn. Chúng ta cũng có thể gọi chúng là các trò chơi tương quan hoàn toàn âm và các trò chơi tương quan hoàn toàn dương, để chỉ sự tương quan của các sở thích đối với các kết cục, và để lại cho trò chơi hỗn hợp phong phú

---

trò chơi lợi ích chung thuần túy, vì những người chơi hoàn toàn đồng ý với thứ tự của các *kết cục* có thể không đồng ý về lợi ích của, chẳng hạn, một điểm cụ thể nào đó với cơ hội 50-50 nằm giữa hai điểm nằm ngay ở trên và dưới nó. Như vậy các trò chơi xung đột "thuần túy hoàn toàn" và lợi ích chung, do không đem lại khuôn khổ cho sự hợp tác ở trường hợp đầu và không có khuôn khổ cho sự bất đồng ở trường hợp sau, sẽ phải cho thấy *các giá trị được kỳ vọng* của tất cả các chiến lược (ngẫu nhiên) hỗn hợp thích hợp, nằm dọc theo các đường dốc đi xuống và đường dốc đi lên tương ứng, với các trục được đo bằng "đơn vị thỏa dụng" đã được nói đến. Điều này đến lượt nó có nghĩa là các điểm biểu thị *các kết cục* phải nằm trên một đường thẳng.

Cũng như vậy, các trò chơi thuần túy không thể chấp nhận "thanh toán phụ." Nếu một trong các đối tác trong trò chơi lợi ích chung thuần túy đe dọa phá hoại kết quả nếu anh ta không được trả tiền - với giả sử rằng cấu trúc liên lạc và cưỡng chế của trò chơi cho phép điều này - thì xung đột lợi ích sẽ xuất hiện. Trên thực tế, điểm biểu thị đưa tiền hối lộ sẽ xuất hiện phía bên trái hay phải dưới bên phải của một điểm khác hay các điểm trên đường dốc đi lên, tạo ra cấu trúc của một trò chơi hỗn hợp. Và nếu một trong các bên tham gia trong trò chơi xung đột thuần túy có thể đe dọa phá hỏng hay đe nghị đền bù để khiến đối phương chịu thua trong trò chơi này, thì sẽ có cơ hội cho sự mập cá; sẽ không còn mối quan hệ xung đột thuần túy, và các điểm biểu thị tổn thất do đe dọa hay đền bù do hứa hẹn sẽ nằm ngoài đường dốc đi xuống. Nói cách khác, *mọi kết quả tiềm năng thích hợp* đều có thể xảy ra. (Hai trò chơi xung đột thuần túy xảy ra đồng thời, ngay cả khi chúng bị giới hạn nằm trên các đường thẳng, sẽ tạo cơ hội cho sự thương lượng trừ phi hai độ dốc của hai đường thẳng ngẫu nhiên bằng nhau.)

hơn cái tên có phần nhảm chán là “trò chơi tương quan không hoàn toàn.”

Khó khăn là ở chỗ tìm ra được một cái tên đủ phong phú cho trò chơi hỗn hợp trong đó có cả xung đột và phụ thuộc lẫn nhau. Đáng chú ý là chúng ta chưa có một từ ngữ rất hay nào dành cho khái niệm về *mối quan hệ* giữa những người chơi: trong trò chơi lợi ích chung chúng ta gọi họ là “các đối tác” và trong trò chơi xung đột thuần túy thì gọi họ là “các đối thủ” hoặc “các địch thủ”. Tuy nhiên, các mối quan hệ hỗn hợp có trong chiến tranh, đình công, thương lượng, v.v... đòi hỏi phải có một thuật ngữ có tính nước đôi hơn.<sup>5</sup> Trong phần còn lại của phần này, tác giả sẽ đề cập đến trò chơi hỗn hợp như là một *trò chơi mặc cả* hay *trò chơi động cơ hỗn hợp*, vì những thuật ngữ này có vẻ bám sát được tinh thần. “Động cơ hỗn hợp,” tất nhiên, không phải nói về việc thiếu hiểu biết rõ ràng về sở thích bản thân mà là tính nước đôi của mối quan hệ của anh ta với người chơi khác - sự hỗn hợp giữa sự phụ thuộc lẫn nhau và xung đột, giữa hợp tác và cạnh tranh. “Tổng lợi ích khác không” chỉ đến trò chơi hỗn hợp cũng như

---

<sup>5</sup> Cần nhấn mạnh rằng các trò chơi có tổng lợi ích khác không cũng có thể được xếp loại một cách thích hợp trong lý thuyết về quan hệ hợp tác cũng như trong lý thuyết xung đột, và để cung cấp hiểu biết cho các vấn đề như chiến tranh hạn chế, sẽ có lý khi sử dụng các từ ngữ nói lên lợi ích chung của các đối thủ và “quá trình mặc cả” có liên quan đến bản thân các hoạt động quân sự. Như sẽ được thấy trong Chương 9, ngay cả vấn đề tấn công bất ngờ cũng tương đồng về mặt logic với một bài toán về quy tắc hợp tác. Nếu lý thuyết về các trò chơi mang quá nhiều hàm ý thiên về xung đột, thì một cái tên chẳng hạn như lý thuyết về quyết định phụ thuộc lẫn nhau (theory of interdependent decision) có lẽ sẽ là một thuật ngữ trung lập bao hàm đầy đủ cả hai trường hợp giới hạn cũng như trường hợp hỗn hợp.

trò chơi lợi ích chung thuần túy. Và, bởi vì nó đặc trưng cho vấn đề và hoạt động có liên quan, *trò chơi phối hợp* có vẻ là một cái tên thích hợp cho việc chia sẻ hoàn toàn các lợi ích.

## CÁC TRÒ CHƠI PHỐI HỢP

Mặc dù phần lớn cuốn sách này là nói về trò chơi hỗn hợp, sự bàn luận ngắn về trò chơi phối hợp thuần túy, ngoài những điều đã nói trong Chương 3, sẽ giúp làm rõ đây là một trò chơi quan trọng bởi chính bản thân nó và sẽ giúp nhận diện một số tính chất nhất định của trò chơi hỗn hợp mà xuất hiện rõ ràng nhất trong trường hợp giới hạn của phối hợp thuần túy.

Hãy nhớ lại các bài toán phối hợp thuần túy trong Chương 3. Mỗi một bài toán rõ ràng đã đem lại một tiêu điểm nào đó cho lựa chọn có phối hợp, một manh mối nhất định cho phối hợp, một căn cứ cho sự hội tụ các kỳ vọng tương hỗ của các bên tham gia. Ở đó có sự lập luận rằng các manh mối phối hợp cùng loại sẽ là một lực lượng có sức mạnh không chỉ trong hợp tác thuần túy mà cả trong tình huống hỗn hợp bao gồm xung đột. Và trên thực tế, các thử nghiệm đã chứng minh rằng trong trường hợp hoàn toàn không có sự trao đổi, điều này hoàn toàn đúng. Nhưng có một số trường hợp trong đó bản thân sự phối hợp thuần túy - thủ tục *ngầm* trong việc nhận diện các đối tác và phối hợp kế hoạch với các đối tác đó - là một hiện tượng có ý nghĩa. Một ví dụ điển hình là sự hình thành các đám đông bạo loạn.

Thông thường điều cốt yếu của sự hình thành một đám

đóng là các thành viên tiềm năng cần phải biết không chỉ ở đâu và bao giờ sẽ tụ tập nhau mà còn khi nào cần hành động để có một sự phối hợp. Lãnh đạo công khai sẽ giải quyết vấn đề này. Nhưng người lãnh đạo thường dễ bị nhận diện và loại trừ bởi nhà cầm quyền muốn ngăn chặn hành động của đám đông. Trong trường hợp này vấn đề của đám đông là thống nhất hành động mà không có sự lãnh đạo công khai, là tìm ra một tín hiệu chung làm cho mọi người tin tưởng rằng nếu anh ta hành động, anh ta sẽ không hành động đơn độc. Do đó, vai trò của "các sự cố" có thể được coi là vai trò phối hợp. Đó là sự thay thế cho lãnh đạo và trao đổi công khai. Nếu không có cái gì đó giống như một sự cố xảy ra, thì khó có thể tạo ra được bất kỳ hành động nào của đám đông, bởi vì việc được miễn trách nhiệm đòi hỏi mọi người phải biết khi nào sẽ cùng hành động. Tương tự như vậy, một thành phố mà không có một địa điểm trung tâm "rõ rệt" hay một địa điểm nổi bật nào khác có thể gây khó khăn cho đám đông tập hợp một cách tự phát; không có địa điểm nào "rõ rệt" đến mức trở thành điều hiển nhiên đối với mọi người rằng đó là điều rõ rệt đối với tất cả mọi người. Hành vi bầy đàn (bandwagon), trong việc lựa chọn lãnh đạo, hay trong việc bỏ phiếu, cũng có thể phụ thuộc vào các tín hiệu "được linh hôi tương hỗ," khi một phần trong sở thích của mỗi người là ý muốn tham gia vào một đa số, hoặc ít nhất là muốn trông thấy một đa số kết thành một khối.<sup>6</sup>

---

<sup>6</sup> Một hiện tượng có liên quan gần gũi được hiểu rõ bởi những người tìm cách hòa nhập vào đám đông để tránh bị nêu tên, hay bị chọn ra để trừng phạt, hay được chọn ra để "tiến cử" vào một số vị trí mà ai cũng muốn tránh.

Hành vi phân cực quá mức có thể là kết quả không hay của sự phụ thuộc vào phái hợp ngầm và thủ đoạn. Khi những người da trắng và những người da đen nhận thấy rằng một khu vực nào đó sẽ “không tránh khỏi” trở thành một nơi bị chiếm giữ hoàn toàn bởi những người da đen, thì “điều không thể tránh khỏi” này trở thành một đặc trưng của kỳ vọng hội tụ.<sup>7</sup> Cái được người ta nhận thức trực tiếp nhất như là điều không thể tránh khỏi không phải là kết quả cuối cùng mà là *sự kỳ vọng*, mà đến lượt nó lại làm cho kết cục đó trở nên không tránh khỏi. Ai cũng kỳ vọng người khác kỳ vọng những người còn lại kỳ vọng ở kết quả đó, và tất cả mọi người đều bất lực không thể từ chối điều đó. Không có tiêu điểm nào ổn định ngoại trừ tại các thái cực. Không ai có thể kỳ vọng quá trình thỏa thuận ngầm sẽ kết thúc bằng con số 10, 30 hay 60%; không có tỷ lệ phần trăm cụ thể nào có thể chi phối được thỏa thuận hay tạo ra một điểm quy tụ. Nếu truyền thống chỉ ra con số 100%, thì truyền thống đó chỉ có thể bị phủ nhận bởi một thỏa thuận công khai. Nếu sự phối hợp phải dưới hình thức ngầm ngầm thì sự thỏa hiệp là không thể thực hiện được. Người ta bị phó mặc cho một hệ thống liên lạc sai sót mà dễ dàng dẫn đến “thỏa thuận” (một cách ngầm ngầm) dời đi nhưng không thể dẫn đến thỏa thuận ở lại. Hệ thống phân bổ hạn ngạch trong phát triển nhà ở, trường học, v.v... có thể

<sup>7</sup> Hiện tượng này, được gọi là “nước tràn ly” (tipping) đã được M. Grodzins phân tích trong “Tình Trạng Chia Tách ở Đô Thị Lớn” (Metropolitan Segregation), tạp chí *Scientific America*, 197:33-41 (tháng 10-1957). Một ví dụ vô hại về các kỳ vọng hội tụ một cách bùng nổ, dựa vào sự liên lạc ngầm có khả năng gây ra phấn khích, là trường hợp một tiếng cười nhỏ gây kích thích làm nổ ra tiếng cười lớn không kiểm soát được trong một đám đông đang lo âu. Một ví dụ quan trọng là sự sụp đổ của chế độ Batista, hay của Đệ tử Cộng hòa.

được coi là những cố gắng lấy một trò chơi công khai có đủ thông tin liên lạc và biện pháp cưỡng chế để thay thế cho trò chơi ngầm có “đáp án” cực đoan không mong muốn.

Trò chơi phối hợp có lẽ đứng đầu sau sự ổn định của các thể chế và truyền thống, và có thể cả hiện tượng về bản thân sự lãnh đạo. Trong số các bộ quy tắc có thể chỉ phối một cuộc xung đột, thì truyền thống chỉ đến bộ quy tắc mà mọi người có thể kỳ vọng người khác nhận thức được đó là ứng cử viên rõ ràng để áp dụng. Nó thắng thế một cách nghiêm nhiên so với những bộ quy tắc mà không dễ nhận diện bởi thỏa thuận ngầm. Hiệu lực của nhiều quy tắc trong nghi lễ và chừng mực xã hội, kể cả một số quy tắc (như quy tắc không được kết thúc câu bằng một giới từ) mà tầm quan trọng hay thẩm quyền của nó đã bị tước bỏ, có vẻ như phụ thuộc vào việc nó đã trở thành “các đáp số” cho trò chơi phối hợp: ai cũng kỳ vọng tất cả người khác kỳ vọng mọi người kỳ vọng vào sự tuân thủ, sao cho sự không tuân thủ sẽ gánh chịu hậu quả do quá lô liếu. Các kiểu quần áo và xe hơi thời thượng có thể cũng phản ánh một trò chơi trong đó người ta không ai muốn mình bị loại ra khỏi một đa số đã hình thành và không được tổ chức để ngăn việc hình thành những đa số. Khái niệm *vai trò* trong xã hội học, một khái niệm có liên quan rõ rệt đến các kỳ vọng của người khác đối với hành vi của một người, cũng như kỳ vọng của một người về cách cư xử của người khác đối với mình, có thể được giải thích một phần thông qua sự ổn định của “các kỳ vọng hội tụ,” của cùng một loại có trong trò chơi phối hợp. Người ta bị mắc vào một vai trò cụ thể, hay bởi vai trò của người khác, bởi vì nó là vai trò duy nhất có thể được nhận diện bởi một quá trình thỏa thuận ngầm trong các tình huống cụ thể.

Một ví dụ điển hình là *tinh thần đồng đội* (esprit de corps) (hay không có) của một đơn vị lục quân hay hải quân hay hệ thống giá trị của một trường học hay hội ái hữu. Đây là những tổ chức xã hội có tỷ lệ thay thế khá lớn nhưng duy trì được bản sắc riêng biệt của chúng ở mức độ mà đường như không phải là do cách tuyển chọn cẩn thận hay thiên vị. Tính cách cá nhân của từng thành viên trong các đơn vị này dường như xuất phát từ kỳ vọng hội tụ - kỳ vọng của mỗi người về những điều mà tất cả mọi người kỳ vọng ở mỗi người - với những kỳ vọng của người mới đến được uốn nắn kịp thời để giúp uốn nắn cho những người đến sau. Ở đây có một nhận thức về "khế ước xã hội," trong đó những điều khoản cụ thể được nhận thức và chấp nhận bởi từng thế hệ mới đến. Tác giả được biết rằng sự bền bỉ của một truyền thống trong một thực thể xã hội là một trong những lý do vì sao bản sắc pháp lý của một sư đoàn hay trung đoàn bộ binh - tên và số hiệu và lịch sử của nó - thường được cố tình lưu giữ lại khi sức mạnh của nó đã suy sụp tới mức gần như phải giải thể: truyền thống đi kèm với bản sắc pháp lý của nhóm người này là một tài sản đáng bảo tồn cho việc xây dựng lực lượng trong tương lai. Đây có thể chính là một hiện tượng dẫn đến việc thu thuế thu nhập thành công ở một số nước này nhưng không thành công ở nước khác: nếu những kỳ vọng tương hỗ thích hợp tồn tại, người ta sẽ kỳ vọng việc trốn thuế ở mức độ nhỏ không làm bận tâm chính quyền và có thể sau đó sẽ bù lại hoặc bởi một ý thức trung thực hơn hoặc bởi nỗi lo sợ bị bắt, cả hai điều đó cộng lại, lý giải cho các kỳ vọng về bản thân chúng.

**Bản chất của quá trình trí tuệ trong phối hợp.** Cần phải nhấn mạnh rằng sự phối hợp không phải là vấn đề di

phỏng đoán xem “một người bình thường” sẽ làm gì. Trong sự phôi hợp ngầm, người ta không tìm cách để phỏng đoán người khác sẽ làm gì trong một tình huống khách quan, mà người ta sẽ cố gắng phỏng đoán người khác sẽ phỏng đoán bản thân anh ta sẽ phỏng đoán người khác phỏng đoán, và tiếp tục như vậy đến vô tận. (“Gặp gỡ” ai đó trong mục đời tư của một tờ báo là một ví dụ điển hình.<sup>8</sup>) Lý lẽ này trở nên xa lạ với tình

<sup>8</sup> Cũng như việc gặp gỡ với ai đó trên cùng một tần số sóng vô tuyến mà người đó có thể gửi tín hiệu cho chúng ta từ không gian vũ trụ. “Chúng ta sẽ tìm trên tần số nào? Việc tìm kiếm theo một dải dài để được một tín hiệu yếu ớt của tần số xa la là một điều rất khó khăn. Nhưng, đúng trong vùng sóng vô tuyến được *đưa* chuộng nhất có một tiêu chuẩn tần số đặc nhất, khách quan, mà mọi nhà quan sát vũ trụ đều biết: đường bức xạ vô tuyến nổi tiếng tại 1420 megaxich của khí hydro trung hòa” (Giuseppe Cocconi và Philip Morrison, *Nature*, ngày 19-9-1959, trang 844-846). Lập luận này đã được John Lear phụ họa thêm: “Tất cả nhà thiên văn trên trái đất sẽ kêu lên ‘À, đương nhiên là 1420 megaxich! Đó là đường bức xạ vô tuyến của hydro trung hòa. Hydro là nguyên tố có nhiều nhất bên ngoài trái đất, những người láng giềng chúng ta có thể kỳ vọng nó sẽ được nhìn thấy thậm chí bởi những người chẳng biết gì về thiên văn học’” (“Cuộc Tìm Kiếm Cuộc Sóng Văn Minh trên Các Hành Tinh Khác” (The Search for Intelligent Life on Other Planets), *Saturday Review*, ngày 2-1-1960, trang 39-43). Chúng ta tìm kiếm tín hiệu gì? Cocconi và Morrison gợi ý một chuỗi gồm một số nguyên tố nhỏ các xung hay các tổng số học đơn giản.

Về điều này gợi ra một định hướng khác của các thí nghiệm này trong đó các đối tượng được yêu cầu đưa ra phán đoán, trong một chuỗi dài ngẫu nhiên các ánh sáng đỏ hoặc xanh, xem cái nào sẽ xuất hiện tiếp theo, đỏ hay là xanh. Các đối tượng hình như bám giữ phán đoán của mình dựa trên cơ sở một số khuôn mẫu mà họ linh hội được, một hành vi “bất hợp lý” nếu xét đến sự hiểu biết của họ rằng chuỗi được tạo ra bởi công cụ ngẫu nhiên. Nhưng, như Herbert Simon đã chỉ ra, “Con người không chỉ là con vật biết học; mà họ còn là con vật biết tìm ra các khuôn mẫu và hình thành khái niệm.” (“Các lý Thuyết Ra Quyết Định trong Kinh Tế và Khoa Học Hành Vi”, (Theories of Decisions – Making in Economics and Behavioral Science), *American Economic Review*, 44:272). Thế thì tại sao không đưa thêm vào cuộc thí nghiệm người làm ra khuôn mẫu phôi hợp, người sẽ làm ra các tín hiệu tùy thuộc vào các giới hạn khác nhau và sự can thiệp ngẫu nhiên, và để cho đối tượng chỉ

huống khách quan, trừ phi chứng nào mà tình huống khách quan có thể đưa ra mạnh mẽ nào đó cho sự lựa chọn có phối hợp. Sự giống nhau ở đây không phải chỉ là tìm cách bỏ phiếu theo đa số mà còn tìm cách bỏ phiếu cùng với một đa số khi ai cũng muốn nằm trong một đa số và mọi người đều biết điều đó - không phải để dự đoán ai sẽ là Hoa hậu Rheingold của năm 1960, mà là đi mua cổ phiếu hay bất động sản mà ai cũng kỳ vọng mọi người kỳ vọng mọi người sẽ mua. Đầu tư vào kim cương có thể là một ví dụ hoàn hảo. Lớn nhất trong tất cả có thể là vai trò tiền tệ của vàng, mà có lẽ chỉ có thể được giải thích như là "đáp số" của một trò chơi phối hợp. (Một phiên bản gia đình của trò chơi phối hợp là khi hai người đang nói chuyện điện thoại thì bị cắt; nếu cả hai cùng gọi lại, thì họ chỉ nhận được tín hiệu máy bận.)

---

biết đi tìm khuôn mẫu sử dụng kỹ năng của mình tìm ra khuôn mẫu do một đối tác hợp tác cài đặt hơn là dùng nó một cách phung phí cho một chuỗi ngẫu nhiên? Nếu, để tận dụng trí thông minh của người liên lạc này, chúng ta đưa thêm vào người thứ ba mà phần thưởng của anh ta liên hệ ngược với phần thưởng của đối tác hợp tác, người được phép ngăn chặn thông điệp và được phép thay đổi nó trong một giới hạn nhất định, chúng ta sẽ có một trò chơi rất giống với trò chơi của Moore và Berkowitz đã mô tả ở trên. Bằng cách bổ sung thêm các dữ liệu ngoài sự lựa chọn lưỡng tính đỏ và xanh có thể đem lại phạm vi rộng cho việc hình thành khuôn mẫu sáng tạo thực sự, một khuôn mẫu đang được chú ý trong tâm lý học Gestalt, thẩm mỹ học và thậm chí việc giải các bài toán ở bậc cao hơn. Simon nhận xét trong cùng một bài báo (trang 426) rằng thậm chí một máy vi tính cũng có thể được lập chương trình "để sử dụng cái gì đó giống như hình tượng hay sự mô phỏng trong việc lập kế hoạch chứng minh" các định lý hình học. Đây là sự tìm kiếm khuôn mẫu thực sự có ích. (Nó nhắc nhớ chúng ta rằng giả thuyết "bản chất ác ý" bởi các nhà lý thuyết trò chơi tổng lợi ích bằng không là không áp dụng được cho, chẳng hạn như, phát minh toán học. Thiên nhiên cho chúng ta những gợi ý. Thiên nhiên để lộ các bí mật của mình theo một khuôn mẫu làm cho chúng trở nên dễ phán đoán hơn rất nhiều so với việc cần có một cái máy quét hoàn hảo để tìm ra chúng.)

Xem xét trò chơi “chọn ra một con số dương.” Các thử nghiệm giống như trường hợp ở Chương 3 chứng minh rằng hầu hết người chơi, được yêu cầu chỉ chọn ra một con số, sẽ chọn các số như 3, 7, 13, 100 và 1. Nhưng khi được yêu cầu chọn ra một số giống con số mà người khác sẽ chọn khi những người khác cũng quan tâm đến việc chọn con số giống nhau, và ai cũng biết người khác cũng đang cố làm như vậy, thì động cơ sẽ khác đi. Sự lựa chọn chiếm ưu thế là con số 1. Và có vẻ như điều này khá logic: ở đây không có “con số được ưa thích” duy nhất. Sự đa dạng về các ứng viên như 3, 7, và v.v... là quá lớn, và không có cách hợp lý nào để chọn ra con số “được ưa thích nhất” hay nổi bật nhất. Nếu có ai đó hỏi con số nào, trong số các con số dương, là độc đáo rõ ràng nhất, hay *quy tắc lựa chọn nào sẽ dẫn đến các kết quả rõ ràng*, thì người đó sẽ rất chú ý đến một thực tế là không gian của tất cả các con số dương có một con số “đầu tiên” hoặc “nhỏ nhất.”<sup>9</sup>

<sup>9</sup> Có một đoạn Keynes viết (trang 156) đã được trích dẫn rất rộng rãi đáng để nêu lại nhằm chỉ ra rằng, trong khi nó đề cập đúng vấn đề đang bàn ở đây, cách tìm “đáp số” của nó không giống một tí nào cả: “Sự đầu tư chuyên nghiệp cũng giống với các cuộc cạnh tranh báo chí trong đó các đối thủ cạnh tranh phải chọn ra sáu bộ mặt đẹp nhất trong số một trăm bức ảnh, giải thưởng sẽ thuộc về đối thủ nào có sự lựa chọn gần nhất với sở thích trung bình của tất cả các đối thủ; do đó mỗi đối thủ cạnh tranh không phải chọn bộ mặt mà mình cho là đẹp nhất mà là bộ mặt mà anh ta nghĩ rằng có nhiều khả năng thu hút nhất sự chú ý của các đối thủ khác, tất cả họ đều nhìn vào vấn đề với cùng quan điểm như nhau. Đây không phải là trường hợp lựa chọn bộ mặt, là đẹp nhất theo đánh giá của một người, thực sự là đẹp nhất, cũng không phải bộ mặt theo ý kiến bình quân nghĩ rằng đẹp nhất. Chúng ta đã đến cấp độ thứ ba trong đó chúng ta dành trí tuệ của mình để phán đoán xem những ý kiến bình quân sẽ dự kiến kết quả bình quân là như thế nào. Và tôi tin rằng, sẽ có những người tính toán ở cấp độ thứ tư, thứ năm và cao hơn nữa” (J. M. Keynes, *Lý Thuyết Chung về Việc Làm*,

## Xác lập bài toán phối hợp theo lý thuyết trò chơi.

Ma trận lợi ích (payoff matrix) cho một bài toán phối hợp thuần túy sẽ có hình dạng như Hình 8. Một người chơi chọn hàng ngang, còn người kia chọn cột dọc, và họ nhận được phần thưởng được đánh dấu bằng con số nằm trong các ô là giao điểm của các sự lựa chọn của họ. Nếu với mỗi lựa chọn của một người chơi tương ứng với một lựa chọn của người kia mà cả hai đều “chiến thắng,” chúng ta có thể sắp xếp các cột sao cho tất cả các ô thắng nằm dọc theo đường chéo. Trong các ô đó là các phần thưởng số dương cho cả hai người chơi, trong các ô còn lại chúng ta ghi số không. (Nhằm phục vụ cho mục đích đang nói chúng ta chỉ ghi vào mỗi ô một con số cho phần thưởng của cả hai người chơi, mà không ảnh hưởng gì đến trò chơi.)

1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

H. 8

*Lãi Suất và Tiền Tệ* (The General Theory of Employment, Interest and Money) [New York, 1936], trang 156). Loại trò chơi này tính cờ chứng minh rằng sự tương quan thường thấy giữa biến thiên của tham số và các số lớn không đúng với trò chơi ngầm có nhiều cân bằng. Để thích ứng một cách “tham số hóa” với hành vi của người khác đòi hỏi trong trường hợp này là hành vi của họ có thể quan sát được, chứ không phải phỏng đoán. Trong trường hợp này, tính phi tham số của phối hợp ngầm vẫn tồn tại, bất kể số người chơi lớn đến đâu.

Nhưng chúng ta phải loại trừ một tiên đề có thể được đề xuất bởi sự tương đồng với các lý thuyết trò chơi khác, đó là (xin mượn từ ngữ của Luce và Raiffa) “việc đánh dấu” (labeling) các hàng, các cột và người chơi sẽ không làm thay đổi gì đến kết quả.<sup>10</sup> Đó chính là vì các chiến lược được “đánh dấu” theo một nghĩa nào đó – tức là chúng có các đặc điểm biểu tượng hay hàm ý vượt quá cấu trúc toán học của trò chơi – những người chơi có thể vượt lên trên sự may mắn và “thắng” trong các trò chơi này. Và cũng vì lý do đó mà các trò chơi này hấp dẫn và quan trọng.

Ngay cả trò chơi được mô tả trên Hình 8 mà trông có vẻ như có ý nghĩa tương trưng tối thiểu gắn với các hàng và các cột không phải là khó để “thắng,” nghĩa là, người chơi có thể làm tốt hơn nhiều so với chỉ dựa vào may rủi, nếu như trò chơi được mô tả trong một ma trận đã biểu thị. (Nếu chúng ta đưa vào trò chơi một dãy vô hạn các hàng và cột, thì hình như sẽ làm cho nó dễ hơn chứ không phải khó hơn. Trong trường hợp

---

<sup>10</sup> Việc đánh dấu *người chơi* đã được công khai loại trừ bởi Luce và Raiffa (trang 123-127) khi thảo luận các trò chơi hợp tác và trong thực tế đã bị Nash loại trừ trong giả thiết đối xứng của ông (J. F. Nash, “Bài Toán Mặc Cả” (The Bargain of Problem), *Econometrica*, 18:155-162 (1950), và “Các Trò Chơi Hợp Tác Gồm Hai Người” (Two Person Cooperative Games), *Econometrica*, 21:128-140 (1953)). Việc đánh dấu các chiến lược cho các trò chơi có tổng lợi ích khác không ngầm hay công khai bị ngầm loại bỏ bằng cách chỉ đề cập đến các trò chơi dưới dạng chính tắc, tức là phiên bản trừu tượng của nó được biểu thị bằng ma trận lợi ích (mà bản thân nó là một công cụ phân tích chứ không phải bộ phận của trò chơi, và do đó không đem lại sự xếp thứ tự trái-phải, cao-thấp, hay số thứ tự của các chiến lược thực tế). Một ví dụ hay, trong đó việc đánh dấu *người chơi* là yếu tố kiểm soát là về cuộc gọi điện thoại bị gián đoạn đã đề cập ở trên, với vấn đề ai là người phải gọi lại và ai là người phải chờ gọi đến.

đó trò chơi mới hoàn toàn đồng nhất với trò chơi đã nêu trước đó, "Chọn một số dương," nhưng, vì việc "đánh dấu" ở đây có khác, nên sẽ có ít cơ hội hơn cho các phe thiểu số tập trung vào các số 3, 7, 13 v.v...) Việc thiết lập ma trận đã làm giới hạn lựa chọn, vì nó chỉ tập trung sự chú ý vào cái "đầu tiên," "ở giữa," "cuối cùng" v.v.<sup>11</sup> Nếu các chiến lược không được đánh dấu theo một dãy liên tiếp, tức là, theo thứ tự các con số hay các chữ cái, mà được đặt bằng các tên riêng, và không được sắp xếp theo một trật tự cụ thể nào, thì chính những cái tên sẽ được dùng để phối hợp lựa chọn.

Và ở đây có một điều rất rõ ràng là các quá trình trí tuệ khi lựa chọn một chiến lược trong tình huống xung đột thuần túy và khi lựa chọn một chiến lược phối hợp là hoàn toàn khác nhau. Ít nhất là đúng như vậy nếu một bên sử dụng giải pháp "tối thiểu hóa tối đa," có thể ngẫu nhiên hóa nếu cần thiết, trong trò chơi có tổng lợi ích bằng không. Trong trò chơi phối hợp thuần túy, mục tiêu của người chơi là tiếp xúc với người chơi khác thông qua một số quá trình nội quan tưởng tượng, tìm kiếm những manh mối được người khác chia sẻ. Trong chiến lược tối thiểu hóa tối đa của trò chơi có tổng lợi ích bằng không - đáng chú ý nhất là với lựa chọn ngẫu nhiên hóa - mục tiêu bao trùm của một bên là tránh mọi sự gặp gỡ trong

---

<sup>11</sup> Đây là nét đặc trưng của nhiều chứng minh trong các thực nghiệm của tác giả đã nói ở phần trên, đại ý là định đề về "sự phụ thuộc vào các phương án không thích hợp" có thể không đáng tin cậy trong trò chơi ngầm, và vì lý do tương tự, không thể mong đợi nó đúng trong trò chơi mặc cả công khai. Các kết cục tiềm tàng có thể phù hợp với sự lựa chọn phối hợp, mặc dù bản thân chúng không được lựa chọn. Để biết thêm về nội dung và sự thảo luận về định đề này, xem Luce và Raiffa, trang 127.

suy nghĩ với bên kia, thậm chí là do vô tình.<sup>12</sup>

Để minh họa, giả sử tác giả sẽ lấy ra một lá bài trong bộ bài gồm 52 lá bài và bạn sẽ đoán lá bài mà tác giả lấy ra. Lý thuyết trò chơi truyền thống sẽ cho chỉ dẫn làm thế nào để tác giả lựa chọn với giả thiết rằng tác giả không muốn bạn đoán được ý của mình. Tác giả có thể lấy hú họa một lá bài và

---

<sup>12</sup> Các chiến lược ngẫu nhiên hóa, mặc dù vây vẫn có ích để có được sự phối hợp phân bổ các lá phiếu, chẳng hạn phân bổ lá phiếu cho một nhóm các ứng cử viên. Nếu có một đa số 55% trong số một trăm cử tri, và họ biết rõ điều đó; nếu có hai trong số sáu ứng cử viên được đa số đó ủng hộ; và nếu ba ứng cử viên chiếm được số phiếu lớn nhất sẽ được bầu vào hội đồng quản trị, thì sẽ có nguy cơ việc bỏ phiếu không phối hợp có thể sẽ tập trung quá nhiều phiếu của phe đa số vào lựa chọn thứ nhất (hoặc thứ hai), để lại cho hai ứng cử viên phe thiểu số thắng cử với mỗi người 22 phiếu. Nhưng nếu mỗi cử tri của phe đa số tung đồng xu để bỏ phiếu cho một trong hai ứng cử viên của đảng mình, thì khả năng một ứng cử viên chỉ nhận được 22 phiếu chỉ là một trong sáu. Nếu phe thiểu số cũng không có phương tiện phối hợp công khai và phải dựa vào công cụ ngẫu nhiên thì các cơ hội của phe đa số là rất thuận lợi. Chiến lược ngẫu nhiên hóa một phần cũng có thể được sử dụng để làm giảm phạm vi xung đột. Giả sử hai người, ngồi ở ghế Bắc và ghế Đông của chiếc bàn đánh bạc, phải chuyển sang ngồi bàn khác bên cạnh được kê theo cùng một hướng, họ phải chọn ghế ngồi ở bàn bên cạnh mà không được liên lạc với nhau, và mỗi người sẽ được thưởng 1 đô-la nếu họ lấy chiếc ghế cạnh nhau. Đây là một bài toán dễ phối hợp. Nhưng chúng ta hãy thay đổi phần thưởng, bằng cách cộng thêm 2 đô-la tiền thưởng cho người chơi nào ngồi vào bên tay phải của người kia trong trường hợp họ cùng ngồi cạnh nhau. Trò chơi này không có điểm cân bằng, các lợi ích không tương đồng, và không có cách xếp chỗ ngồi nào mà không tạo ra sự khuyến khích để dịch chuyển. (Ai cũng muốn bên kia hứa hẹn sẽ ngồi bên tay trái của mình, nhưng không thể.) Chiến lược ngẫu nhiên sẽ đem lại cho mỗi người chơi giá trị tối thiểu hóa tối đa 1 đô-la. Nhưng, nếu mỗi bên quyết định anh ta sẽ ngồi ở đâu trong một trò chơi lợi ích chung thuận túy, sau đó tung đồng xu để xem anh ta sẽ ngồi ở chỗ đó hay ngồi đối diện, những người chơi đảm bảo được rằng họ sẽ không chọn cùng một ghế hay ngồi đối diện nhau và sẽ chia sẻ cơ hội giành được phần thưởng. Đây là một cặp cân bằng trong chiến lược (hỗn hợp), có giá trị kỳ vọng là 2 đô-la cho mỗi người.

thách đố bạn có cách gì tốt hơn ngoài việc đoán mò lá bài của tác giả. Nhưng nếu đổi lại trò chơi tác giả *lại* muốn bạn đoán đúng lá bài đã chọn và bạn cũng biết rằng tác giả sẽ tìm cách rút lá bài sao cho tạo thuận lợi cho sự phán đoán của bạn, thì công cụ ngẫu nhiên không giúp gì cho sự hợp tác ngầm. Holmes có thể *xóa bỏ* việc đánh dấu các nhà ga bằng cách tung đồng xu để quyết định xem sẽ xuống tàu ở đâu, và Moriarty chỉ có cơ hội 50-50 để dự đoán kết quả tung đồng xu. Nhưng trong trường hợp trò chơi có lợi ích chung hai bên phải *sử dụng* ở mức độ nào đó việc đánh dấu các nhà ga để làm được tốt hơn là chỉ dựa thuần túy vào sự may rủi. Và cách sử dụng nó như thế nào sẽ phải dựa vào trí tưởng tượng hơn là logic, vào thơ ca hay sự hài hước hơn là toán học. Điều đáng lưu ý là lý thuyết trò chơi truyền thống không gắn “giá trị” vào trò chơi: việc người ta có thể phối hợp với nhau tốt đến đâu theo cách này là một điều mà, mặc dù hy vọng có thể tuân theo sự phân tích hệ thống, không thể tìm ra được bằng lập luận tiền nghiệm (từ nguyên nhân suy ra hậu quả - a priori - người dịch chú thích). Góc này của lý thuyết trò chơi *hiển nhiên* là phải dựa vào bằng chứng thực nghiệm.<sup>13</sup>

---

<sup>13</sup> Trong các trường hợp giống như thế này chúng ta chỉ cần xem xét câu hỏi *cái giá* nào mà những người chơi sẽ đồng ý trả cho một thông tin phối hợp, và các khuôn mẫu thông tin nào sẽ đem lại cơ hội gì cho sự phối hợp, là chúng ta đã rơi vào trung tâm của lý thuyết *nhóm* của Marschak.

Một cách vô tình, ở đây có một phiên bản “Sự bế tắc của các tù nhân” cho trò chơi này: hai kẻ đồng phạm, bị bắt trước khi dân xếp chứng minh ngoại phạm và bị thẩm vấn riêng rẽ, phải phối hợp nhau chứng minh chứng cứ ngoại phạm họ dung nén, nếu không sẽ bị kết tội. Một biến thể thú vị có thể được đưa ra bằng cách giả sử rằng việc thú nhận sẽ bị án nhẹ hơn so với có tội nhưng không chịu thú nhận. Mỗi người chơi có một chiến lược “tối thiểu hóa tối đa” về thú tội và không chỉ xem

Cần đặc biệt lưu ý là để khẳng định ảnh hưởng của các “nhân hiệu” (tức là, ảnh hưởng của các chi tiết tượng trưng và có hàm ý của trò chơi) và sự phụ thuộc của lý thuyết vào bằng chứng thực nghiệm không có liên quan đến vấn đề liệu lý thuyết trò chơi mang tính tiên đoán hay chuẩn mực – liên quan đến các khái quát hóa về sự lựa chọn trong thực tế hay chiến lược của sự lựa chọn đúng đắn. Điều khẳng định ở đây *không* phải là người ta chỉ đơn giản bị ảnh hưởng bởi các chi tiết tượng trưng mà là họ *nên* chịu ảnh hưởng đó vì mục đích có sự lựa chọn đúng. Lý thuyết chuẩn mực phải đưa ra các chiến lược mà ít nhất cũng không kém hơn so với khi người ta làm mà không có nó. Hơn nữa, lý thuyết đó không được từ chối hay xóa bỏ các chi tiết của trò chơi có lợi rõ ràng cho người chơi, hai hoặc nhiều hơn, và do đó những người chơi

xét bằng chứng ngoại phạm cụ thể nào sẽ tạo ra chiến lược ngoại phạm tốt nhất mà còn xem bằng chứng đó tốt đến đâu (xét về sự trùng khớp với lời khai của kẻ tòng phạm kia) và liệu họ có cùng nhau quyết tâm thực hiện hay không. Ma trận của trò chơi này là:

$\text{S}$	$\text{S}$	$\text{S}$	$\text{S}$
$\text{S}$	$\text{I}$	$\text{O}$	$\text{O}$
$\text{O}$	$\text{I}$	$\text{I}$	$\text{O}$
$\text{O}$	$\text{O}$	$\text{I}$	$\text{I}$

(Con số nằm dưới bên trái của mỗi ô là phần thưởng cho người chơi chọn hàng ngang, con số nằm trên bên phải là cho người chơi chọn cột dọc.)

không được loại bỏ hay xem thường các chi tiết đó trong các lợi ích chung của mình. Hai cặp nhảy xoay xở tìm chỗ trống trên sàn nhảy hoặc hai đội quân xoay xở để tìm đường ranh giới ngừng bắn đều có thể chịu thiệt hại vì các quá trình ra quyết định bị giới hạn trong các đặc điểm trừu tượng của hoàn cảnh.

Hệ quả đặc biệt của sự khái quát này là trò chơi theo dạng “chính tắc” (trừu tượng toán học) không tương đương về logic với trò chơi theo dạng “mở rộng” (cụ thể), mỗi khi chúng ta chấp nhận logic theo đó những người chơi phối hợp các kỳ vọng của họ với nhau. Như đã nói trong Chương 3, cách suy xét tương tự dường như cũng có mặt rõ ràng trong mặc cả công khai. Hệ quả thuật ngữ của những suy xét này là “không hợp tác” không phải là cái tên hay cho trò chơi phối hợp ngầm; nó vẫn có sự hợp tác theo cách riêng của nó và vẫn như vậy khi chúng ta đưa thêm vào sự xung đột và hình thành trò chơi có động cơ hỗn hợp ngầm. (Trong Phụ lục C có nêu lập luận rằng một số khái niệm đáp án (solution concept) quen thuộc trong lý thuyết trò chơi có thể được diễn giải dưới dạng khái niệm về phối hợp.)

## *GỢI Ý VÀ NHẬN THỨC TƯƠNG HỐ TRONG TRÒ CHƠI CÓ ĐỘNG CƠ HỖN HỢP*

Lý thuyết trò chơi phối hợp, mặc dù có sự hấp dẫn của nó, có sự hấp dẫn chủ yếu là vì nó soi sáng cho bản chất của trò chơi có động cơ hỗn hợp. Yếu tố phối hợp xuất hiện nổi bật

nhất trong trò chơi ngầm thuần túy, trong đó không có sự trao đổi thông tin cũng như không có bất kỳ một chuỗi các động thái nào mà hai người tham gia trò chơi chuyển đến cho nhau. Một ví dụ, tương tự bài toán 6 trong trang 62, sẽ như sau.

Một người chơi “định cư” ở thành phố Cincinnati, còn người kia ở San Francisco. Hai người có hai chiếc bản đồ nước Mỹ giống nhau và muốn chia bản đồ này thành hai phần. Mỗi người sẽ vẽ ra một đường chia nước Mỹ thành hai phần; đường phân chia có thể là đường thẳng hoặc đường cong, có liên hệ hoặc không có liên hệ về các yếu tố địa lý hay chính trị. Nếu hai người chia bản đồ theo cách khác nhau, thì không ai được thưởng gì cả. Nhưng nếu họ vẽ các đường phân chia trên bản đồ giống nhau, thì cả hai sẽ được thưởng. Phần thưởng cho mỗi người phụ thuộc vào cái nằm ở phần của mỗi người sau khi chia, tức là, phần có chứa thành phố nơi anh ta định cư. Chúng ta không cần làm rõ các phần thưởng này. Chúng có thể phụ thuộc một phần vào diện tích, một phần vào dân số, một phần vào sự giàu có về công nghiệp và tài nguyên nông nghiệp, v.v... và có thể phần nào khác nhau giữa hai người chơi. Nói cách khác, trong khi tất cả lnh thổ đều quý giá, nhưng không phải tất cả các bộ phận của lãnh thổ đều quý giá ngang nhau, và không có công thức nào định giá trị của chúng một cách rõ ràng. (Không có phương tiện để lựa chọn sự phân chia hoàn toàn đối xứng các giá trị của hai người chơi.)

Trong trò chơi này có một vấn đề phòi hợp bắt buộc. Mỗi người chơi chỉ có thể thắng nếu anh ta làm đúng y như kỳ vọng của người kia, biết rằng người kia cũng đang cố gắng tương tự để làm đúng như kỳ vọng của anh ta. Họ phải cùng nhau tìm ra đường phân chia mà theo một cách nào đó bản thân

nó gợi ý cho cả hai người hay có sức lôi kéo cho cả hai người. Không bên nào có thể “khôn ngoan” hơn bên nào mà không gây hại cho chính mình.

Các thí nghiệm trong Chương 3 cho thấy hai người chơi không hề tuyệt vọng khi đối mặt với trò chơi kiểu này. Trò chơi không đến nỗi khó khăn “vô tận” như sự vô tận của các đường phân chia lãnh thổ gợi ra. Một số biến thể của trò chơi chẳng khó khăn chút nào. Nhưng kết quả thành công của trò chơi lại phụ thuộc vào các loại nhân tố kiểm soát trò chơi phối hợp thuần túy này. Thực tế là, trong một số trò chơi loại này cả hai người đều “thắng” nhờ chọn chính xác kết quả như nhau như họ sẽ chọn nếu hệ thống giải thưởng cho họ các lợi ích giống hệt nhau, thay vì lợi ích xung đột. Vấn đề là tìm ra một số tín hiệu hay đầu mối hay căn cứ hợp lý để hai bên nhận ra sự lựa chọn “đúng,” với việc mỗi bên sẵn sàng tuân theo kỷ luật bởi tín hiệu hay đầu mối đó trong trường hợp chúng phân biệt đối xử với anh ta. Họ phải tìm ra các đầu mối của mình ở nơi có thể tìm ra. (Chẳng hạn, nếu bản đồ họ đang sử dụng chẳng may có rất nhiều đầu mối, gây khó khăn để tìm ra cái nào đặc biệt, thì trọng tài sẽ gợi ý bằng cách vẽ ra một con đường tùy ý giống nhau trên cả hai bản đồ, có thể được coi như là “người làm trung gian,” mặc dù nó có thiên vị đáng kể về một bên.)

Nhưng yếu tố phối hợp này, đặc biệt trong trường hợp không có xung đột, lại tỏ ra có liên hệ chặt chẽ đến *vấn đề trao đổi thông tin*. Trò chơi phối hợp thuần túy không những mất sức hấp dẫn mà gần như không còn là “trò chơi” nữa nếu những người chơi có thể phối hợp với nhau một cách chắc chắn không có khó khăn, và không tốn kém gì cả. Khi đó, câu

hỏi đặt ra là, nói chung, yếu tố phối hợp quan trọng đến đâu trong các trò chơi có động cơ hỗn hợp, vì nhiều trò chơi loại này có hình thức mặc cả công khai không ngăn cản giao tiếp.

Sự có mặt khắp mọi nơi của nguyên tắc phối hợp xuất phát từ hai xem xét riêng biệt. Thứ nhất, như đã được thảo luận ở Chương 3 đó là sự mặc cả ngầm cho ta một mô hình phân tích - có lẽ chỉ là sự tương tự nhưng cũng có thể là một sự nhận diện hiện tượng tâm lý hay trí tuệ thực tế - của quá trình "hợp lý" trong việc tìm kiếm thỏa thuận trong các tình huống mặc cả thuần túy, trong đó cả hai bên công nhận có một phạm vi rộng các kết cục có lợi cho cả hai bên so với không có thỏa thuận nào. Hiện tượng tâm lý của "nhận thức tương hỗ" mà có thể được chứng minh là đúng trong trường hợp mặc cả ngầm có một vai trò quan trọng trong việc phân tích mặc cả công khai. Vai trò đó chính là *sự phối hợp các kỳ vọng*.

Thứ hai, có nhiều quá trình mặc cả hay các tình huống trò chơi mà chúng ta cần phân tích ít nhất cũng một phần mang tính ngầm. Trong một số trường hợp, như việc tìm cách lái xe trong tình huống tắc nghẽn giao thông, lời nũi bị ngăn cản bởi điều kiện vật lý. Trong các trường hợp khác, như việc tìm kiếm sự hòa giải với người láng giềng thì lời nói bị ngăn cản bởi lợi ích của giữ bí mật riêng tư. Sự mặc cả trái phép, hay sự mặc cả ngoại giao mà có thể gây lúng túng cho cả hai bên tham gia nếu nó bị tiết lộ cho các nước khác, có thể là không được công bố đầy đủ. Nếu số người chơi trong một trò chơi là lớn, như trong quá trình mặc cả để quyết định đường ranh giới chủng tộc giữa các khu vực dân cư và nghề nghiệp, có thể sẽ không có cơ sở mang tính thể chế cho sự đàm phán công khai. Trong các trường hợp này, trong khi lời lẽ có thể là bộ phận của quá

trình mặc cả, nhưng hành động cũng là bộ phận của quá trình đó, và trò chơi là một “thủ đoạn” hơn chỉ là nói chuyện.

Hơn nữa, nếu những người chơi có sẵn các *bước đi*, sao cho việc tiếp tục dùng thủ đoạn là có lợi ngay cả trong quá trình đàm phán, và đặc biệt nếu một số thủ đoạn trở nên lộ rõ đối với người chơi khác chỉ sau một khoảng thời gian, thì sẽ không có lý do để cho rằng việc hoãn dùng thủ đoạn ngay lập tức sẽ được tôn trọng ngay từ đầu. Trong trường hợp đó, trò chơi vẫn diễn ra trong khi đàm phán tiếp tục. Nếu các bước đi chỉ có ý nghĩa tương trưng, chúng ta có thể đưa chúng vào quá trình trao đổi thông tin cùng với lời nói. Tuy nhiên, trường hợp điển hình là những nước đi có ý nghĩa chiến thuật, làm cho trò chơi trở nên khác biệt so với trước mà không thể đảo ngược được nữa, cũng như những nước đi có ý nghĩa chiến thuật đưa nó lên mức độ cao hơn cả lời lẽ thuần túy thậm chí về nội dung thông tin. Một người có thể nói đi nói lại là súng đã được lắp đạn mà không thể chứng minh được điều đó cho đến khi anh ta thực sự bóp cò. Một người có thể nói đi nói lại rằng anh ta coi một khu vực nào đó có ý nghĩa quan trọng về chiến lược nhưng không ai tin cho đến khi anh ta bỏ ra chi phí hoặc mạo hiểm để bảo vệ nó. Như vậy, những bước đi có thể làm lộ thông tin về hệ thống giá trị của người chơi hoặc về sự lựa chọn hành động có trong tay anh ta. Các bước đi có thể ràng buộc anh ta với một số hành động nhất định trong khi lời lẽ thường không được như vậy. Và các bước đi thường diễn ra với tốc độ được quyết định đơn phương, không phụ thuộc vào nghi thức thỏa thuận tại hội nghị.

Nói cách khác, các trò chơi mặc cả thường hay bao gồm

một quá trình động của sự nhẫn nhượng qua lại hơn là trao đổi thuận túy đạt đến đỉnh cao trong một thỏa thuận rõ ràng. Việc tìm kiếm các giới hạn trong chiến tranh hạn chế là một ví dụ hoàn hảo, và chúng ta có thể minh họa nó bằng việc bổ sung vào trò chơi đoán chữ như đã mô tả ở trên.

*Một minh họa về trò chơi ngầm.* Giả dụ hai người tham gia trò chơi của chúng ta với các bản đồ nước Mỹ trước mặt, mỗi người được giao 100 thẻ đánh bạc và được yêu cầu chơi trò chơi như sau.<sup>14</sup> Ở mỗi "nước đi," người chơi sẽ phân bổ năm thẻ cho các bang trên bản đồ của mình. Các nước đi của hai người sẽ được so sánh, và nếu cả hai người chơi đều đặt một thẻ vào cùng một bang như nhau, thì hai thẻ đó sẽ bị thu lại. Nếu một người đặt một thẻ và người kia đặt ba thẻ vào cùng một bang, thì mỗi người sẽ bị lấy lại một thẻ và chỉ để lại hai thẻ đại diện cho người chơi đã đặt ba thẻ, và cứ tiếp tục như vậy. Họ sẽ làm tương tự như vậy ở nước đi tiếp theo,

<sup>14</sup> Vì trong Chương 6 sẽ đưa ra đè xuất rằng các trò chơi như vậy trên thực tế có giá trị nghiên cứu cũng như giá trị minh họa, nên cần nhận thấy ngay từ đầu có một vấn đề đặc biệt trong việc khuyến khích người chơi tham gia trò chơi thí nghiệm có tổng lợi ích khác không. Trong trò chơi có tổng lợi ích bằng không, việc giành chiến thắng được do lường tượng đối với đối thủ trực tiếp, và sự thách thức trí tuệ và sự canh tranh tay đôi khuyên khích người chơi hướng về cách giành chiến thắng đúng đắn (và duy nhất). Nhưng đối với trò chơi có động cơ hỗn hợp, việc giành "chiến thắng" phải được thực hiện sao cho đạt được số điểm tuyệt đối của một người, chứ không phải số điểm tương đối so với người chơi với mình. Sự khuyến khích sẽ bị bóp méo nếu cuộc chơi chỉ có cạnh tranh tay đôi. Do vậy, nếu không có giải thưởng thực sự, trò chơi phải được tổ chức theo kiểu luân lưu hoặc theo sơ đồ có sự tham gia của hai người chơi theo một chuỗi trò chơi có hai người, với kết quả cuối cùng được quyết định bởi vị trí tương đối của số điểm tuyệt đối của một người. (Đó là lý do tại sao không có các trò chơi đoán chữ có tổng lợi ích khác không gồm hai người chơi.)

một lần nữa bằng năm cái thẻ. Lần này họ có sự lựa chọn hoặc đặt cái thẻ của mình vào các bang mới hoặc đặt thẻ vào các bang cũ đã có các thẻ được đặt trước. Nếu A đặt hai thẻ vào một bang tại đó B đã đặt một thẻ từ trước, thẻ của B sẽ bị thu hồi cùng với một thẻ của A, để lại một thẻ của A có mặt để "lấy" bang đó. Và cứ thế trò chơi tiếp tục cho đến khi cả hai người dùng hết toàn bộ số thẻ của mình. Sau đó trò chơi lại tiếp tục, và cứ mỗi nước đi một người chơi có thể dịch chuyển thẻ của mình từ một đến năm cái từ bang này sang bang khác, và cũng với số thẻ bằng nhau được thu hồi khỏi bang mà cả hai người cùng đặt thẻ của mình. Quá trình này tiếp tục cho đến khi cả hai người chơi báo cho trọng tài biết rằng họ muốn chấm dứt cuộc chơi.

Bây giờ phần thưởng sẽ được phân phối. Mỗi người chơi sẽ nhận được một đô-la cho một chiếc thẻ của mình còn lại trên bàn, tức là, cho số thẻ không bị thu hồi khi anh ta "lấy được" một bang hay "mất đi" một bang vào tay đối phương. Anh ta cũng sẽ nhận được tiền cho số bang mà anh ta "sở hữu," đó là những bang mà anh ta có các thẻ đặt trên đó cộng với số bang không có thẻ nhưng nằm trong khu vực chứa căn cứ địa của anh ta, tức được hoàn toàn bao quanh bởi các bang mà anh ta có thẻ đặt trên đó.

Những "phần thưởng" cho các bang được sở hữu này có giá trị cụ thể bằng đô-la gắn với từng bang trong số 48 bang. Chúng tuân theo một quy tắc nào đó có gợi ý về, chẳng hạn như, "giá trị kinh tế" của mỗi bang hay đại loại như vậy. Ở đây cả hai người đều không có sự giả định trước rằng các giá trị này là như nhau, hoặc thậm chí tương quan chặt chẽ với nhau, cho hai người chơi. Dân số có thể là một yếu tố quan

trọng về “giá trị” của các bang đối với người chơi này nhưng lại là tương đối không quan trọng về “giá trị” đối với người chơi khác. Không người chơi nào biết được hệ thống giá trị của người kia – hoặc có thể chỉ biết chút ít về nó, chẳng hạn yếu tố nào là quan trọng nhưng không biết là quan trọng đến đâu. Mỗi người chơi phải nghiên cứu để biết được hệ thống giá trị của người kia thông qua sự quan sát các nước đi của người kia.

Ở đây chúng ta có một trò chơi động cơ hỗn hợp, một trò chơi diễn ra bởi một quá trình hỗ trợ lẫn nhau – một loạt các nước đi trong quá trình chơi trong đó người chơi chịu chịu tổn thất chung nếu họ phối hợp không tốt. Họ sẽ mất tiền nếu không đoán được bên kia đặt thẻ vào chỗ nào cho mỗi nước đi, trong các trường hợp đó họ không muốn mất tiền do tranh nhau một bang. Mỗi bên mất ít nhất một đô-la khi một người lấy được một bang của người kia, và mỗi người sẽ mất hơn một đô-la nếu người bị mất đi một bang tìm cách lấy lại bằng cách đặt thêm nhiều thẻ vào bang đó. Và họ không chỉ mất đi một đô-la cho mỗi đô-la bị tịch thu, mỗi người sẽ còn lại ít “thẻ” hơn xét theo góc độ của các bang mà mình có quyền lợi, và họ có thể phải để trống một số bang không ai chiếm giữ nếu số thẻ còn lại của họ không đủ khi trò chơi kết thúc.

Bây giờ những người chơi trò chơi này “mặc cả” với nhau như thế nào? Bằng cách này hay cách khác, thực ra họ có đưa ra các đề nghị và phản đòn đề nghị. Họ chấp nhận, từ chối, trả đũa, và thậm chí tìm cách để chuyển cho nhau sự đe dọa và hứa hẹn.<sup>15</sup> Nhưng nếu chúng ta không cho phép họ giao tiếp

---

<sup>15</sup> Điều này đã được chứng minh trong các thử nghiệm sơ bộ với một trò chơi như vậy.

bằng lời nói dưới bất kỳ hình thức nào, họ sẽ phải trao đổi những ý định và đề nghị của họ bằng các khuôn mẫu hành vi. Mỗi bên sẽ phải cảnh giác trước những biểu hiện thủ đoạn của bên kia, và mỗi bên phải đủ sáng tạo để chuyển được ý định của mình khi anh ta muốn chuyển. Nếu một người chơi thực sự muốn một bang cụ thể nào đó, bởi vì nó đặc biệt có giá trị cao đối với anh ta đến nỗi anh ta phải bám giữ lấy nó và bảo vệ nó bằng mọi cách, và tồn nhiều đô-la cho cái mình yêu thích, trước khi bên kia chịu từ bỏ, thì sẽ tốt hơn cho cả hai người nếu họ nhận ra trước ai là người mong muốn bang đó nhiều nhất. Và nếu một bên thực sự sẵn sàng để nhân nhượng một phần lớn lãnh thổ làm "vật trao đổi" để đổi lấy bộ phận lãnh thổ mà anh ta muốn có nhất, anh ta không những phải làm cho bên kia thấy rõ điều đó mà còn phải phân định các giới hạn của nó bằng khuôn mẫu hành động của mình.

Nhưng những khuôn mẫu này từ đâu tới? Cấu trúc toán học của trò chơi không giúp được nhiều về vấn đề này, đặc biệt do chúng ta đã cố tình làm cho hệ thống giá trị của mỗi người chơi trở nên quá không chắc chắn đối với bên kia để làm cho những suy xét đối xứng, công bằng, và những suy xét khác có thể có ích. Giả sử rằng, họ tìm ra các khuôn khổ của họ nằm trong những cái như là các biên giới tự nhiên, các nhóm chính trị thân quen, các đặc điểm kinh tế của các bang có thể nằm trong các hệ thống giá trị của họ, tâm lý học Gestalt, và mọi sáo ngữ hay truyền thống mà họ có thể tìm ra cho mình trong quá trình tham gia trò chơi.<sup>16</sup>

---

<sup>16</sup> Nếu cây có trái của hàng xóm xòe cành sang bên sân của nhà tôi và tôi hái tất cả trái cây nằm về phía bên nhà mình, láng giềng của tôi có thể nhận rõ "đề nghị" của tôi là gì và hiểu rõ anh ta sẽ phải nhận

**Trao đổi công khai.** Vậy giờ chúng ta hãy thay đổi luật chơi sao cho những người chơi có thể nói chuyện thoải mái với nhau. Điều này sẽ làm cho trò chơi khác đi như thế nào? Ở một số khía cạnh, nó sẽ làm tăng hiệu quả của người chơi. Những vật trao đổi cụ thể này có thể được nhận diện so với trước kia là quá phức tạp để có thể đưa ra đề nghị trong một hoàn cảnh không thuận tiện. Cũng có thể những người chơi sẽ tránh một số cuộc va chạm không chủ ý do đặt các thẻ vào cùng một bang mà đã làm cho họ mất đi một số đô-la. Chúng ta không thể biết chắc họ sẽ cùng nhau tránh được cuộc chạy đua tốn kém để giành lấy các bang; bởi vì lợi thế là người đầu tiên có mặt tại một bang là đủ lớn để kích thích những người chơi tiếp tục chơi thậm chí ngay cả khi họ đang nói chuyện. Và họ không có cách gì để thuyết phục lẫn nhau rằng điều họ nói với nhau là thật trừ phi họ chứng minh điều đó qua cách mà họ chơi. (Chúng ta sẽ để cho họ nói cho nhau họ đánh giá các bang như thế nào, nhưng chúng ta công khai không phạt những lời nói không thật, và chúng ta cũng không cho những người chơi bằng chứng bằng văn bản về hệ thống giá trị của họ để họ có thể cho nhau xem.)

Như vậy việc đưa vào sự trao đổi bằng lời không bị ngăn cản có thể không làm thay đổi lớn tính chất của trò chơi, mặc dù kết cục cụ thể sẽ khác. Sự phụ thuộc của hai người chơi

---

nhưng gì trong tương lai nếu anh ta không trả đũa. Nhưng nếu thay vào đó, tôi hái cùng một số lượng trái cây từ cả hai bên một cách ngẫu nhiên hay hái với một số lượng, chẳng hạn, tỷ lệ với số người trong gia đình tôi, thì anh ta sẽ có ít khả năng nhận ra tôi có ý đồ gì trong đầu. (Anh ta có thể sẽ buộc phải chống lại hay trả đũa nhiều hơn nếu tôi hái một phần thay vì hái tất cả trái cây nằm bên phần nhà tôi, vì tôi đã không phân định giới hạn ý định của mình.)

vào việc truyền đạt các ý định của họ cho nhau và nhận thức được các ý định của nhau, về hành vi theo khuôn mẫu có thể dự đoán được và sự chấp nhận các luật chơi hay các giới hạn, phần lớn vẫn là như cũ.

Sự tương phản với trò chơi có tổng lợi ích bằng không và việc tự làm hại mình trong một đáp số tối thiểu hóa tối đa là điều rất nổi bật ở đây. Với một kết quả tối thiểu hóa tối đa, trò chơi có tổng lợi ích bằng không bị rút xuống còn là một công việc hoàn toàn đơn phương. Người chơi không những không cần liên lạc với đối phương của mình, anh ta thậm chí không cần biết đối phương của mình là ai, hay có đối phương hay không. Chiến lược ngẫu nhiên là chiến lược hoàn toàn phản trao đổi. Nó là một phương tiện có tình phá hủy mọi khả năng xác lập trao đổi, đặc biệt là trao đổi các ý định, dù có tình hay không. Đó là phương tiện xóa bỏ mọi chi tiết khỏi trò chơi trừ cấu trúc toán học về lợi ích và xóa bỏ mọi quan hệ liên lạc của những người tham gia trò chơi.

Trong cờ vua việc các quân cờ có trông giống con ngựa, mục sư, con voi, lâu đài hay vỏ bánh hamburger hay không chẳng có gì là quan trọng. Cũng như vậy, bất kể trò chơi này được gọi là "cờ vua," "cuộc nội chiến," hay "bất động sản," hay bất kể các ô vuông bị làm méo trông giống với sự phân chia chính trị hay địa lý cũng không có gì quan trọng. Nó cũng không có gì quan trọng nếu những người chơi có biết gì về nhau hay liệu họ có nói cùng một ngôn ngữ và có cùng văn hóa hay không. Nó cũng không có gì quan trọng ai là người từng chơi trò chơi này trước đây và kết quả ra sao. (Nếu điều đó là quan trọng, một trong những người chơi sẽ có động cơ

để loại bỏ ảnh hưởng của các tình tiết này, và một chiến lược tối thiểu hóa tối đa, ngẫu nhiên hóa nếu cần, sẽ loại bỏ nó.)

Nhưng ta có thể thay đổi ma trận lợi ích trong cờ vua, làm cho nó trở thành trò chơi có tổng lợi ích khác không mà người chơi được thưởng không chỉ đối với các phần mà họ chiếm được mà còn các phần còn lại sau khi kết thúc, cũng như các ô vuông mà họ chiếm giữ, sao cho cả hai người chơi có lợi ích trong việc làm giảm đến mức thấp nhất "tổng số" phần nắm giữ phá hoại giá trị chung. Ta có thể làm cho mỗi người chơi không biết chắc được các ô vuông nào và con cờ cụ thể nào bên kia đánh giá cao nhất. Và ta có thể đặt thời gian cho các nước đi để không bên nào trì hoãn nước đi của bên kia nhằm nói chuyện với anh ta.

Bây giờ sẽ là sự khác biệt đối với người chơi cờ nếu chúng ta gọi trò chơi là "chiến tranh" hay "đi tìm vàng," hoặc các con cờ trông giống con ngựa, người lính, người thám hiểm, hay các trẻ em săn lùng trứng Phục sinh, hay chiếc bản đồ nào hay bức tranh nào được đặt lên bàn và các ô vuông bị làm méo mó thành các hình thù khác nhau ra sao, hay câu chuyện dẫn nhập cho người chơi được kể trước khi bắt đầu cuộc chơi.

Bây giờ chúng ta đã dàn xếp trò chơi sao cho những người chơi phải *mặc cả* cách chơi để có kết quả, hoặc bằng lời nói hoặc bằng các nước đi nối tiếp, hoặc cả hai. Họ phải tìm cách điều khiển hành vi của mình, trao đổi các ý định của nhau, tự cho phép mình tìm kiếm những điểm gặp gỡ trong suy nghĩ, một cách ngầm hay công khai, để tránh làm tổn hại đến lợi ích tiềm tàng của nhau. "Những chi tiết vô tình" có thể tạo thuận lợi cho các bên tìm ra khuôn mẫu hành vi được biểu thị;

và mức độ mà các nội dung *biểu trưng* của trò chơi - những sự gợi ý và hàm ý - gợi ra các thỏa hiệp, các giới hạn, và các luật lệ cần phải có để tạo sự khác biệt. Nó cần phải có, bởi vì nó sẽ giúp cho cả hai bên không giới hạn mình trong phạm vi cấu trúc trừu tượng của trò chơi trong việc tìm kiếm các khuôn mẫu về nước đi có tính ổn định, không phá hoại lẫn nhau, và có thể nhận ra được. Quá trình tâm lý và trí tuệ cơ bản là sự tham gia vào việc tạo ra *các truyền thống*, và các thành phần từ đó các truyền thống được tạo ra, hoặc các nguyên liệu trong đó các truyền thống tiềm tàng sẽ được nhận thức và công nhận chung, không hề trùng hợp với các nội dung toán học của trò chơi.<sup>17</sup>

Kết quả trò chơi được xác định bởi các kỳ vọng mà mỗi người chơi nghĩ người kia sẽ chơi như thế nào, khi mỗi người biết rằng các kỳ vọng của họ có tính tương hỗ đáng kể. Những

---

<sup>17</sup> Một ví dụ hay là câu hỏi liệu một đường phân chia rõ ràng có thể được vạch ra giữa vũ khí nguyên tử và các vũ khí khác hay không, câu trả lời theo báo cáo hiện nay là không có nếu lấy sức nổ làm tiêu chuẩn, các cấp độ của sức nổ là chồng lấn lên nhau. Tuy nhiên vẫn có sự khác biệt nếu có đủ số lượng người suy nghĩ như vậy, và đúng là họ nghĩ như vậy. Đó chính là sự khác biệt được tạo ra xuất phát từ chính cấu trúc của kỳ vọng: đó là *truyền thống* 10 năm coi vũ khí nguyên tử là khác biệt; người ta tin như vậy và tin người khác cũng tin như vậy, và ngay cả những người bác bỏ sự khác biệt vẫn sẽ nín thở, mỗi khi một quả bom nguyên tử nổ tung trong chiến tranh, vì họ không thể giải thích sự khác biệt thông qua sức nổ. Đó chỉ đơn thuần là sự khác biệt thông thường, giống như sự khác biệt làm cho việc bắt giam không phải là sự trừng phạt "độc ác và bất bình thường" hay làm cho, chẳng hạn, sự đại diện của trường đại học tại Quốc hội hoàn toàn thích hợp với nền dân chủ của Anh nếu nó luôn luôn tồn tại nhưng không thích hợp nếu nó đã phải được lập lại sau một khoảng thời gian 10 năm. Sự khác biệt của vũ khí nguyên tử cũng có thể là một cái được cố tình cung cấp hay cố tình làm mờ nhạt đi theo thời gian, như mọi truyền thống khác. (Điểm này sẽ được trình bày sâu thêm ở Phụ lục A.)

người chơi cần phải cùng nhau tìm ra và cùng nhau chấp nhận kết quả hay cách chơi sẽ làm cho kết quả được xác định. Họ phải cùng nhau tìm ra “luật chơi” hoặc cùng nhau gánh chịu các hậu quả.

Một ví dụ rất hay của vấn đề trao đổi ý định là việc làm cho bên kia hiểu rõ, một cách thuyết phục, khuôn mẫu trả đũa được dự định cho các hành động cụ thể mà một bên đề nghị xem là “vượt giới hạn.” Nếu không có sự trao đổi thông tin đầy đủ, khả năng của một bên để chuyển đi một khuôn mẫu ý định như vậy sẽ phụ thuộc không những vào các tư liệu tình huống để hình thành các biên giới và giới hạn mà còn phụ thuộc vào khả năng của bên kia nhận ra công thức (Gestalt) trả đũa khi anh ta nhìn thấy một ví dụ về nó. Tiền lệ lịch sử và văn học, sự ngụy biện pháp lý và đạo lý, toán học và thẩm mỹ học, cũng như những sự tương đồng thuộc mọi ngóc ngách của cuộc sống, có thể tạo ra một thực đơn từ đó một bên lựa chọn khuôn mẫu trả đũa có thể nhận ra được cũng như sự diễn giải khuôn mẫu dự định của bên kia. Ngay cả khi có sự trao đổi đầy đủ bằng lời, thì tình huống này có thể chẳng khác gì mấy. Các khuôn mẫu hành động có sức mạnh hơn lời nói.

Như vậy ảnh hưởng mà các chi tiết gợi ý của một trò chơi có thể có tác động đến kết quả và sự phụ thuộc của người chơi vào những manh mối và tín hiệu mà trò chơi đem lại có ích không chỉ cho việc nghiên cứu bằng cách nào những người chơi thực sự hành động trong trò chơi có tổng lợi ích khác không. Lập luận ở đây không phải là những người chơi chỉ phản ứng đối với các tính chất phi toán học của trò chơi mà là họ phải tính đến vai trò của chúng, từ đó ngay cả một lý thuyết chuẩn tắc - lý thuyết về *chiến lược* của các trò chơi -

phải nhận thấy rằng những người chơi hợp lý có thể sẽ cùng nhau lợi dụng điều này. Và ngay cả khi một người chơi hợp lý nhận ra rằng cấu hình của những chi tiết này bất lợi cho anh ta, anh ta cũng có thể nhận thức hợp lý rằng anh ta không trông cậy vào - bên kia sẽ kỳ vọng một cách hợp lý rằng anh ta tuân theo kỷ luật của các gợi ý xuất phát từ những chi tiết cụ thể của trò chơi và sẽ có những hành động mà, với tổn thất chung cho cả hai bên, làm cho anh ta phải hợp tác.<sup>18</sup>

<sup>18</sup> Cần phải nói thêm rằng khái niệm sức hút nội tại hay sự tập trung vào các kết cục cụ thể trong một tình huống mặc cả hay trong bài toán phối hợp thuần túy nhận được sự ủng hộ và làm rõ bởi các bằng chứng thực nghiệm của các nhà tâm lý học theo trường phái Gestalt. Các công trình của họ về nhận thức các hình dạng vật chất là rất thích hợp. Chẳng hạn, các hình thù không hoàn chỉnh được đưa ra cho những người hiếm thị một phần, và họ thường nhận ra hình thù đầy đủ của chúng hơn chỉ là một bộ phận. Nhưng các hình thù cụ thể mà họ "hoàn chỉnh" cho chính họ lại theo một số nguyên tắc nhất định về sự giản dị, và các hình dạng "đơn giản" không quen thuộc lại được hoàn chỉnh trong khi các hình dạng rất quen thuộc nhưng không đơn giản lại không được hoàn chỉnh. Koffka đề cập đến hiện tượng này là "tổ chức tự phát trong các hình dạng đơn giản." Chúng ta sống giữa những hình chữ nhật méo mó, nhưng cái mà chúng ta "nhìn thấy" là các hình chữ nhật, không khác gì các hình chữ nhật hoàn hảo, bởi vì "hình chữ nhật thực sự là một hình thù được tổ chức tốt hơn hình thù kém chính xác hơn nó." Nói về các tính chất tối thiểu-tối đa của các quá trình không chuyển động, Koffka cho rằng các quá trình tâm lý sẽ có những tính chất này: "Vì chí ít chúng ta có thể lựa chọn các tổ chức tâm lý xuất hiện dưới điều kiện đơn giản và sau đó dự đoán chúng phải có tính đều đặn, tính đối xứng, tính đơn giản. Kết luận này được dựa trên nguyên tắc của sự đồng hình (isomorphism) theo đó các đặc tính của các quá trình sinh lý học cũng có các khía cạnh của các quá trình nhân thức tương ứng." Và, "Như thế chúng ta có được một nguyên tắc chung, mặc dù có phần còn mơ hồ, để chỉ dẫn chúng ta nghiên cứu tổ chức tâm lý... Nguyên tắc... có thể được xác lập như thế này: tổ chức tâm lý sẽ luôn luôn là "tốt" chừng nào điều kiện chỉ phối cho phép. Trong định nghĩa này tính ngữ "tốt" không được định nghĩa. Nó bao gồm các tính chất

*Một thí nghiệm giả định.* Minh họa cho những gì tác giả suy nghĩ trong đầu, thí nghiệm giả định sau đây cần có thể được xem xét. (Hy vọng rằng, một số thí nghiệm như vậy sẽ được tiến hành.) Ở đây, nó là một sự tương tự về khái niệm, hay có thể hiểu là một phép thử thực nghiệm về hiện tượng tâm lý xảy ra trong quá trình mặc cả.

Giai đoạn thứ nhất của thử nghiệm này là chế ra một cái máy, có thể dựa theo nguyên tắc của máy kiểm tra nói dối, sẽ ghi chép hoặc đo lường "sự nhận ra" của một người hoặc sự tập trung chú ý của anh ta hay sự đề phòng hay sự kích động của anh ta. Cái mà chúng ta muốn là một cái máy có thể đo được, khi mà người chơi rà soát một dãy các kết quả có thể theo một thứ tự nào đó, mức độ mà các kết quả cụ thể gây sự chú ý của anh ta hay tạo ra sự kích động trong quá trình mặc cả thực tế diễn ra.

Với cái máy đó, chúng ta thiết lập một trò chơi mặc cả. Để cho đơn giản, chúng ta sẽ làm ra trò chơi trong đó có một số phần thưởng sẽ được chia khi đạt được thỏa thuận về cách phân chia. Chúng ta cho trò chơi đầy đủ "nội dung chủ đề" để có chỗ cho sự tranh luận, sự ngụy biện, các căn cứ khác

---

như là tính đều đặn, tính đối xứng, tính đơn giản và các tính chất khác mà chúng ta sẽ gặp phải trong quá trình thảo luận (K. Koffka, *Các Nguyên Lý của Tâm lý học Gestalt* (Principles of Gestalt Psychology), (London, 1955).

Nếu nhận thức cá nhân và "tổ chức" về các hình dạng tuân theo các hạn chế này, quá trình "nhận thức tương hỗ" và "các hình thức tổ chức tương hỗ" khi có sự hội tụ của các kỳ vọng sẽ phải phụ thuộc vào các hạn chế tương tự ít nhất với mức độ khắt khe như thế. Và, do trò chơi có tổng lợi ích khác không đòi hỏi cái gọi là "tổ chức hình dạng" chung tối hậu, một lý thuyết chiến lược chuẩn tắc (chứ không chỉ tâm lý học mô tả) sẽ phải kể đến các hạn chế này.

nhau v.v..., tức là, cho nhiều thứ hơn là chỉ một khoảng toán học với trung điểm dễ thấy.

Bây giờ chúng ta có hai người chơi được nối vào cái máy của mình sao cho mỗi người có thể nhìn thấy đồng hồ đo trên máy của mình, mỗi người có thể nhìn thấy đồng hồ đo trên máy của người kia, và mỗi người đều biết rõ rằng cả hai người đều biết rằng họ có thể nhìn thấy đồng hồ đo của nhau. Nói cách khác, họ cũng nhận thức được rằng cả hai đều có thể nhìn thấy các phản ứng của nhau đối với những kết quả cụ thể khi chúng xuất hiện trên máy. Chúng ta sẽ sử dụng một thiết bị thăm dò cơ học, có thể chuyển động qua lại trong phạm vi các kết quả có thể xảy ra, chỉ đèn, nháy đèn, hoặc tập trung vào kết quả có khả năng, hết cái này đến cái khác. Có thể nó sẽ theo một hành trình đều đặn, hoặc ngẫu nhiên. Chúng ta hãy để cho chiếc máy thăm dò, và hãy để những người chơi quan sát khi máy dò, quan sát đồng hồ đo của mình và đồng hồ của người khác, và quan sát cả bộ mặt của nhau nếu họ muốn.

Cuối cùng, chúng ta sẽ đi sâu vào trò chơi, và ở đây có thể có một vài biến thể của trò chơi. Một khả năng lý thú sẽ là loại bỏ mặc cả công khai và chỉ cần cho máy hoạt động, chạy qua lại hay vòng quanh các kết quả có thể. Chúng ta sẽ quan sát xem liệu các phản ứng được ghi lại của hai người chơi cuối cùng có dẫn đến sự tương đồng tại một kết quả nào không, theo nghĩa là các phản ứng bắt buộc, có thể nhận diện được của họ đạt mức cao nhất đối với kết quả cụ thể như nhau trong số tất cả những người mà thiết bị dò tìm cho thấy phản ứng của họ. (Vì mục đích kiểm soát, chúng ta có thể yêu cầu một lần từng người chơi tham gia vào cuộc dò tìm mà không

có mặt người kia, để có được ý niệm về phản ứng của từng người chơi một cách độc lập với sự tương tác giữa hai người.) Nếu sự hội tụ diễn ra, chúng ta chắc chắn nhận diện được một hiện tượng đầy ý nghĩa, dù chúng ta có thể cho rằng đó *chính* là quá trình mặc cả tâm lý hay không. Chúng ta sẽ chứng minh được: (a) rằng những người chơi sẽ phản ứng lại trước nội dung của tình huống mặc cả và (b) các phản ứng của họ phụ thuộc vào sự tương tác qua lại xuất phát từ thực tế là mỗi người chơi có thể nhìn thấy phản ứng của người khác và mỗi người cũng biết rằng phản ứng dễ nhận thấy của mình sẽ nói lên những kỳ vọng của chính mình. (Tác giả phỏng đoán rằng, giống như vợ của Lot, những người chơi thường không thể rời sự chú ý của mình khỏi các kết quả cụ thể, thậm chí các kết quả bất lợi, và rằng một nỗ lực rõ ràng để bỏ qua một "tiêu điểm" thường làm tăng sức mạnh tiêu điểm).<sup>19</sup>

Một biến thể khác có thể là để cho những người chơi mặc cả công khai trong khi dò tìm kết quả và đo lường, bằng thiết bị dò tìm không ngừng đưa ra những phản ứng bên ngoài của họ trong quá trình trao đổi mà cả hai bên đều thấy rõ. (Trong trường hợp này, chúng ta thậm chí còn cho một người chơi

<sup>19</sup> Nhận xét sau, do Koffka trích dẫn, có thể khó tin nhưng nói rất đúng vấn đề: "Khi một chuyên gia... theo dõi một trận đấu bóng đá một cách chăm chú, anh ta sẽ dễ ý thấy rằng thủ môn, đứng trước một khung thành khá lớn, thường bị đá vào người hơn là do cú đá tinh cờ của đối phương, ngay cả khi người ta tính đến thực tế rằng thủ môn bắt cú lúc nào anh ta có thể cũng sẽ tìm cách ngăn chặn quả bóng cho bằng được. Thủ môn là một điểm nổi trội trong không gian thu hút đôi mắt của cầu thủ đối phương. Nếu cú đá xảy ra khi đôi mắt của người đá nhắm vào thủ môn thì quả bóng thường rơi vào gần anh ta. Nhưng khi người đá biết cách chỉnh lại tầm nhìn của mình, để làm thay đổi "trọng tâm" nổi bật từ thủ môn sang một điểm khác trong không gian, thì trọng tâm mới này cũng sẽ có sức hút như thủ môn đã có trước đó."

được viện dẫn bằng chứng của phản ứng trông thấy trên đồng hồ đo nếu anh ta muốn như là một chiến thuật mặc cả, chẳng hạn để nói với đối thủ của mình rằng anh ta “rõ ràng” không thể chịu được cái giá 60 đô-la như lời người kia đang đòi, trong khi dấu hiệu từ huyết áp của anh ta cho thấy trong thâm tâm anh ta sẵn sàng chấp nhận cái giá 40 đô-la.)

Thí nghiệm này có thể dựa trên ba giả thuyết. Thứ nhất, một người chơi riêng lẻ sẽ có “các phản ứng” có thể nhận diện vật lý được sau khi quan sát các kết cục khác nhau trong số một dãy các kết cục có thể có của trò chơi và những phản ứng này sẽ phải khác biệt rõ rệt trong số những phản ứng khác nhau. Thứ hai, những phản ứng này, khi người chơi biết rằng chúng lộ rõ trước con mắt đối thủ của mình, sẽ biến thiên theo cách có thể gọi ý sự mặc cả, tức là, những phản ứng của hai người chơi, khi lộ rõ trước mặt cả hai người, sẽ tương tác với nhau theo kiểu của một “quá trình mặc cả.” Thứ ba, hiện tượng được đo lường này, mà chúng ta so sánh nó với quá trình mặc cả, là một bộ phận hay nằm trong, hay có liên quan đến, chính quá trình mặc cả được định nghĩa theo cách thông thường. (Một thí nghiệm theo loại này, như đã được mô tả, có thể sẽ đặc biệt lý thú cho trường hợp trò chơi có nhiều hơn hai người.)

Cuộc thí nghiệm vẫn chưa được tiến hành và chưa được viện dẫn như bằng chứng. Nó được mô tả ở đây nhằm cung cấp một cách thể hiện trên thực tế của hệ thống lý thuyết mà tác giả suy nghĩ trong đầu khi nói đến “sự hội tụ” của các kỳ vọng và để gợi ý rằng sự hội tụ mà cuối cùng sẽ xảy ra trong quá trình mặc cả có thể phụ thuộc vào tính động của bản thân quá trình và không chỉ đơn thuần vào các dữ kiện tiền nghiệm của trò chơi.

### *Một số đặc tính động của các giải pháp tiêu điểm.*

Sự phụ thuộc của một kết quả “tiêu điểm” vào một số đặc điểm phân biệt nó về chất so với các kết quả xung quanh có những xem xét động quan trọng. Chẳng hạn, nó thường làm cho những sự nhẫn nhượng nhỏ ít có khả năng hơn so với những nhân nượng lớn, nó thường có nghĩa rằng tiêu điểm có sức thuyết phục như là một kết quả được kỳ vọng chính xác hơn là xấp xỉ. Nếu một người mặc cả đã kiên trì đòi hỏi 50% mà không thành công, sự thỏa hiệp với 47% là ít có khả năng, vì sự nhẫn nhượng nhỏ có thể là một dấu hiệu của sự sụp đổ. Các nguyên tắc định tính khó có thể thỏa hiệp, và các tiêu điểm nói chung phụ thuộc vào các nguyên tắc định tính. Người ta không thể mong đợi làm thỏa mãn một kẻ xâm lược bằng cách cho hắn một vài dặm vuông đất trên lãnh thổ của họ. Hắn ta biết rằng cả hai bên đều biết rằng hai bên kỳ vọng bên kia sẽ rút lui cho đến khi tìm được một biên giới mới có sức thuyết phục mà có thể biện giải được.

Trong thực tế, một tiêu điểm cho thỏa thuận thường có tính chất của mình dựa vào một thực tế rằng những nhẫn nhượng nhỏ thường không thể được, rằng những sự xâm lấn nhỏ sẽ dẫn đến sự xâm lấn nhiều hơn và lớn hơn. Người ta vạch đường phân chia ranh giới tại những nơi có địa hình dễ nhận thấy hoặc căn cứ vào một số nguyên tắc rõ ràng dựa vào câu hỏi tu từ, “Nếu không phải ở đây, thì ở đâu?” Nhẫn nhượng càng rõ ràng là sụp đổ bao nhiêu, thì tiêu điểm càng có sức thuyết phục bấy nhiêu. Luận điểm này cũng được minh họa trong trò chơi mà chúng ta tự chơi với chính mình khi chúng ta muốn bỏ hút thuốc hoặc uống rượu. “Chỉ uống một ít thôi,” là một lời thỏa hiệp không vững chắc mà ai cũng biết, và số người bỏ hút thuốc lá hoàn toàn nhiều hơn so với số người cố

gắng đạt được một thỏa hiệp ổn định chỉ hút với số lượng nhỏ mỗi ngày. Một khi nguyên tắc khởi thủy không còn nữa, thì cũng sẽ mất niềm tin đối với tất cả những điểm dừng khác, và những sự kỳ vọng sẽ quy về sự sụp đổ hoàn toàn. Chính việc nhận ra điều này sẽ giữ cho sự chú ý tập trung vào một điểm là cai hoàn toàn.

Đôi khi bản thân tiêu điểm cũng bất ổn định một cách cố hữu. Trong trường hợp đó, nó sẽ không trở thành kết quả mà chỉ là một dấu hiệu chỉ ra nơi nào sẽ có kết quả. Điều này thường đúng cho trường hợp "bỏ phiếu thử" tại cơ quan lập pháp hoặc một "vấn đề thử nghiệm" xuất hiện trong các quan hệ giữa những người chơi khi trò chơi đang diễn ra. Thông thường đó là một sự thách thức, hay thách đố, hay một hành động ngang bướng mà, về bản chất của nó, sẽ tạo ra hoặc là phản ứng khuất phục của bên kia hoặc là bị buộc phải rút lui. Chính cái thành phần phụ của trò chơi lại làm biểu tượng cho bản thân trò chơi, tạo ra một khuôn mẫu cho các kỳ vọng vượt qua thực chất của tiêu điểm liên quan. Đôi khi điều này là cố ý định và tạo thành một chiến thuật có tình. Trong các trường hợp khác hành động hay vấn đề đó phát triển một ý nghĩa biểu tượng không cố ý, làm cho sự thỏa hiệp là không thể.

Sự công nhận ngoại giao chế độ Cộng sản tại Trung Quốc, sự tuyên thệ trung thành tại các trường đại học, việc giải quyết cuộc đình công tại một ngành công nghiệp then chốt, sự đầu hàng của đám đông trước một kẻ phá bĩnh tại một buổi tiệc rượu, hay cuộc bỏ phiếu cho một dự luật tại một hội nghị chính trị, tất cả đều có thể có ý nghĩa như điều nói trên. Đôi khi đúng là kết quả đối với vấn đề cụ thể này chính là căn cứ cho việc quyết định các vấn đề khác như thế nào, ví dụ như

một cuộc bỏ phiếu thử sẽ chỉ cho thấy một cách chính xác sự phản đối trước một giải pháp sẽ lớn đến đâu. Nhưng thường thì một vấn đề cụ thể không đại diện cho phần còn lại của trò chơi, nó chỉ đòi hỏi sự công nhận ngầm như là một dấu hiệu cho tất cả cái khác sẽ diễn ra, sao cho mỗi bên chỉ còn là con tin hoặc người được hưởng lợi của các kỳ vọng tương hỗ đã được thiết lập.

Thông thường hiện tượng này được nhận diện như là một tín hiệu thực sự trong một trò chơi phối hợp. Các thành viên của một liên minh không có tổ chức thường nhận thấy những tiềm năng của hành động phối hợp mà không biết chắc chắn rằng một "thỏa thuận" là có tồn tại cho phối hợp hành động. Mọi người đều muốn biết những người khác sẽ hành động như thế nào và liệu những người khác sẽ làm cái điều mà anh ta biết mình cũng phải làm hay không. Cuộc bỏ phiếu thử tại cơ quan lập pháp hay một số hành động đồng thời trong một nhóm người, như cuộc biểu tình lớn, thường là một biện pháp để "phê chuẩn" sự tồn tại của một liên minh và sự chứng minh rằng mọi người kỳ vọng tất cả những người khác hành động phối hợp. Nhưng ngay cả trong một trò chơi có hai người, mà điển hình là một cuộc thách đấu, hiện tượng lẩn át hay quy phục về tâm lý có thể chứng tỏ là tương đồng về tâm lý với kết quả của một trò chơi mặc cả.

Quá trình này, theo đó những nước đi cụ thể trong một trò chơi hay những lời đề nghị và sự nhân nhượng đạt được tầm quan trọng tương trưng với tư cách là chỉ dấu cho biết các kỳ vọng sẽ hội tụ về đâu trong phần còn lại của trò chơi, có vẻ là một lĩnh vực trong đó tâm lý học thực nghiệm có thể đóng góp cho lý thuyết trò chơi.

**Tính thích hợp thực nghiệm của các tiêu điểm toán học.** Chúng ta cần tránh cho rằng bất cứ cái gì nhà phân tích có thể nhận thức được cũng sẽ được những người tham gia trò chơi nhận thức được, hay bất cứ cái gì đưa ra gợi ý chi phối đối với nhà phân tích thì cũng sẽ như vậy đối với người chơi. Đặc biệt, các đặc tính của trò chơi thích hợp cho các đáp số toán học phức tạp (trừ phi giải pháp tương tự cũng có thể đạt được bằng cách khác, ít phức tạp hơn) có thể sẽ không có sức mạnh làm hội tụ các kỳ vọng và gây ảnh hưởng đến kết quả của trò chơi. Chúng có thể có sức mạnh đó chỉ khi nào những người chơi coi nhau như các nhà toán học. Đây có thể là sự giải thích trên cơ sở thực nghiệm của các “đáp số” như của Braithwaite, Nash, Harsanyi, và những người khác. Đó là các thuộc tính toán học của một trò chơi, giống như các thuộc tính thẩm mỹ, các thuộc tính lịch sử, các thuộc tính pháp lý và đạo lý, các thuộc tính văn hóa, và tất cả các chi tiết gợi ý và hàm ý khác, có thể phục vụ cho sự tập trung các kỳ vọng của một số người chơi vào một số kết quả nhất định. Nếu bản thân hai người chơi là các nhà toán học về lý thuyết trò chơi, có thể họ sẽ cùng nhau nhận thức và bị ảnh hưởng mạnh mẽ bởi các kết quả tiềm tàng có những thuộc tính toán học hấp dẫn. Mỗi người có thể vượt lên trên, và biết rằng người bên kia cũng sẽ vượt lên trên nhiều chi tiết ngẫu nhiên mà, đối với những người chơi không phải là nhà toán học, có thể thích hợp cho việc tập trung các kỳ vọng nhiều hơn so với một số thuộc tính định lượng của trò chơi.

(Trong nhiều trường hợp các thuộc tính toán học này có thể là tính đơn nhất hay đối xứng có thể có các định nghĩa phi toán học và cả sự hấp dẫn phi toán học, hoặc có thể trùng

hợp với các điểm phân biệt được một cách định tính có thể được biện giải bằng một phương pháp phi toán học có sức thuyết phục không kém.)

Như vậy các đáp số toán học là một loại trong một nhóm các tác động có sức tập trung các kỳ vọng, nhưng chúng tác động thông qua cùng một cơ chế tâm lý - sức mạnh gợi ý có thể đưa các kỳ vọng về điểm hội tụ - như các loại khác. Khi hai vợ chồng bị lạc nhau trong cửa hàng bách hóa họ sẽ vội tìm đến phòng Đồ Thất Lạc dựa vào sự hiểu biết tương hỗ ngầm và hài hước rằng đó là địa điểm "rõ ràng" để tìm gặp nhau, còn hai nhà toán học trong cùng hoàn cảnh đó - ai cũng biết rằng cả hai đều là nhà toán học - có thể sẽ đi tìm một vị trí duy nhất về mặt hình học hơn là một vị trí chỉ dựa vào từ ngữ.

Điểm mấu chốt ở đây không phụ thuộc vào liệu, theo "các luật chơi" của lý thuyết trò chơi, một người chơi hợp lý phải được giả định là có hiểu biết toán học nhiều như anh ta cần có hay không. Ở đây chúng ta nói về những nhận định được chia sẻ giữa những người tham gia trò chơi, sự bận tâm, sự tập trung cao độ, và sự nhạy cảm đối với những gợi ý, chứ không phải nói về nguồn lực họ có thể lấy ra khi cần thiết. Nếu hiện tượng "thỏa thuận hợp lý" về cơ bản là hiện tượng tâm lý - sự hội tụ của các kỳ vọng - thì sẽ không có giả định cho rằng lý thuyết trò chơi toán học là cần thiết cho quá trình đạt được thỏa thuận, do đó không có cơ sở để cho rằng toán học là nguồn cảm hứng chủ yếu trong quá trình hội tụ. (Vấn đề này sẽ được bàn thêm trong Phụ lục B.)

Người ta có thể đồng ý hoặc không đồng ý với bất kỳ giả thuyết cụ thể nào về việc những kỳ vọng của người mặc cả

được hình thành như thế nào, hoặc trong quá trình mặc cả hoặc trước đó, và hoặc bởi bản thân sự mặc cả hay bởi các sức mạnh khác. Nhưng rõ ràng rằng kết quả của quá trình mặc cả sẽ được mô tả một cách sớm nhất, trực diện nhất, và mang tính thực nghiệm nhất thông qua hiện tượng các kỳ vọng hội tụ ổn định. Bất kể một người đồng ý công khai về kết quả mặc cả hay đồng ý ngầm hay chấp nhận vì không còn cách nào khác, anh ta sẽ phải, nếu khôn ngoan, thấy rằng mình không thể làm tốt hơn và nhận ra rằng bên kia cũng phải đáp ứng lại suy nghĩ đó. Như vậy cơ sở của một kết quả, vốn chỉ là sự lựa chọn phối hợp, cần phải được mô tả phân tích bằng khái niệm về sự hội tụ của các kỳ vọng.

*Trao đổi thông tin chủ quan.* Vai trò của “nước đi biếu lộ” trong một trò chơi có sự hỗ trợ lẫn nhau được tăng cường bởi lý do là trong các trò chơi có động cơ hỗn hợp, ngược lại so với các trò chơi có tổng lợi ích bằng không mà những người tham gia biết rằng sẽ có tổng lợi ích bằng không, sẽ có nhiều khả năng mỗi bên không chắc chắn về hệ thống giá trị của bên kia. Trong trò chơi có động cơ hỗn hợp, các bước đi sẽ có nội dung *thông tin*.

Chúng ta cũng không thể xác lập một trường hợp chung của trò chơi mặc cả trong đó mỗi bên biết trước sở thích của bên kia. Việc cho rằng mỗi bên đều biết được ma trận lợi ích “thực sự” của bên kia thường là đã đặt ra giả thiết phi thường về những dàn xếp thể chế cho trò chơi. Lý do là một số yếu tố nhất định trong một trò chơi mặc cả là điều *vốn dĩ không thể biết được* đối với một số người tham gia trò chơi, trừ trường hợp có những điều kiện đặc biệt. Làm sao chúng ta có

thể biết được người Nga không thích một cuộc chiến tranh tổng lực đến mức nào trong đó cả hai bên sẽ bị tiêu diệt? Chúng ta không thể biết được. Và lý do chúng ta không thể biết *không* phải đơn thuần chỉ vì người Nga nhất thiết không muốn chúng ta biết được điều đó. Ngược lại, các tình huống có thể xuất hiện trong đó họ lại muốn chúng ta biết sự thật đó. Nhưng bằng cách nào họ có thể làm chúng ta biết được điều đó? Bằng cách nào họ có thể làm cho chúng ta tin rằng điều họ nói với chúng ta là thật? Làm thế nào để tù nhân đang bị tra tấn vì một bí mật mà anh ta thực sự không biết có thể thuyết phục kẻ bắt giữ rằng anh ta không hề biết bí mật đó? Làm thế nào để Trung Quốc, nếu họ thực sự quyết tâm thu hồi Đài Loan bằng cái giá của một cuộc chiến tranh tổng lực, có thể thuyết phục chúng ta rằng họ không thể bị răn đe bằng mọi hình thức và rằng mọi sự đe dọa của chúng ta chỉ làm cho cả hai bên cam kết một cuộc chiến tranh tổng lực?<sup>20</sup>

---

<sup>20</sup> Việc thiếu phương tiện để kiểm chứng sự thật chính là cơ sở của trò chơi mồi mọc trong đó mỗi người chơi gắn một giá trị dương vào hạnh phúc của người khác, như khi hai vợ chồng thảo luận với nhau có nên xem phim hay không, mỗi bên muốn làm bất cứ điều gì bên kia muốn làm và muốn tỏ ra là mình muốn, biết rằng bên kia cũng sẽ tương tự thể hiện sở thích theo phỏng đoán cái mà một bên muốn làm v.v... Cũng có một lĩnh vực riêng của lý thuyết trò chơi liên quan đến mối quan hệ giữa người với người trong đó sự biểu lộ công khai hay sự công nhận hệ thống giá trị của một người có ảnh hưởng đến các giá trị; sự nhận biết của tôi về việc người láng giềng không thích tôi có thể làm tôi buồn bực một ít, cũng như sự nhận biết của anh ta về sự nhận biết của tôi, nhưng nếu chúng tôi buộc phải thừa nhận sự thật đó một cách công khai, thì sự buồn bực sẽ rất lớn. "Nghi lễ xã hội", như nhận xét của Erving Goffman, "nhắc nhở những người đàn ông chờ nên yêu cầu các cuộc hẹn hò vào đêm giao thừa quá sớm trong mùa lễ hội này nếu không thì các cô gái sẽ thấy khó tim ra cờ lịch sự để từ chối." "Về Giữ Thể Dien," *Tâm lý học: Tạp chí Nghiên cứu Các Quá Trình Giao Tiếp* (Psychiatry: Journal for the Study of Interpersonal Processes), 18:224 (1955).

Trong các trường hợp đặc biệt thông tin có thể được chuyển đi. Trong một trò chơi hú cát, trong đó “hệ thống giá trị” của mỗi người chơi được chứa đựng trong các con bài hoặc các thẻ, anh ta chỉ cần lật mặt con bài lên (nếu luật chơi cho phép hoặc nếu anh ta và đối thủ của anh ta có thể cùng nhau đánh lừa trọng tài). Trong một xã hội mà người ta tin tưởng tuyệt đối vào sức mạnh siêu đẳng sẽ trừng phạt sự sai trái khi được yêu cầu làm điều đó và mọi người đều biết tất cả những người khác cũng tin như thế, “tôi xin thề nếu không thì tôi chết” là một công thức đầy đủ để truyền đi sự thật một cách tự nguyện. Nhưng đó là những trường hợp đặc biệt. Nếu chúng ta muốn có một “trường hợp khái quát” đó phải là trường hợp trong đó có ít nhất một số điều không biết về hệ thống giá trị của nhau, hay sự lựa chọn chiến lược của mỗi bên, chỉ vì những sự thật đó vốn dĩ không thể biết được hoặc không thể trao đổi được.

Von Neumann và Morgenstern đã minh họa khái niệm *đáp số* của họ cho trò chơi có tổng lợi ích khác không bằng một ví dụ của một người bán hàng A, sẵn sàng bán ngôi nhà của mình với cái giá từ 10 trở lên, và hai người muốn mua B và C, sẵn lòng trả giá tối đa tương ứng là 15 và 25.<sup>21</sup> (Con số là của tác giả.) Điều sáng tạo của đáp số này là ở chỗ C có thể trả cho B một phần tiền tiết kiệm của mình nếu, nhờ B rút khỏi cuộc mua bán, C có thể mua được ngôi nhà với giá thấp hơn 15. Hai bên đưa ra đề nghị – và sự giới hạn này có sẵn trong khái niệm về *đáp số* – là B sẽ nhận được từ C tối đa là 15-10=5. Điều đáng chú ý về đòi hỏi thông tin trong đáp số này

---

<sup>21</sup> J. Von Neumann và O. Morgenstern, Lý Thuyết Trò Chơi và Hành Vi Kinh Tế (Theory of Games and Economic Behavior), (Princeton, 1953), trang 564ff.

không phải là giá bảo lưu 15 của B là điều mà anh ta có trình bày sai, mà là trong đời thường anh ta không thể nào nói lên sự thật một cách thuyết phục nếu anh ta muốn như vậy. Khái niệm “đáp số” này – bằng giả thiết có thông tin đầy đủ – không chỉ loại trừ sự tham gia của những tay đầu cơ (trừ phi họ thực sự muốn ngôi nhà đến mức cùng nhau chia sẻ trong đáp số), mà nó còn giả định rằng C có thể nhận thấy, hoặc B có thể tiết lộ, một sự thật chủ quan, mà D và E (những tay đầu cơ bị lôi cuốn do nhận thấy rằng B có lợi nhuận thuần túy do việc mặc cả một đồ vật mà anh ta không bao giờ làm chủ nó trước hay sau cuộc mua bán) không thể lừa gạt được.

Chắc chắn còn có các trường hợp đặc biệt, trong đó một bên có thể giả thiết rằng bên kia cũng giống như mình về các giá trị cơ bản và do đó có thể đánh giá các giá trị của bên kia bằng cách áp dụng tính chất đối xứng đơn giản. Nhưng có rất nhiều trường hợp thú vị khi người ta chơi với một đối thủ là một người hoàn toàn khác biệt. Người cha của một đứa trẻ bị bắt cóc sẽ không thể nào đoán được giá tiền chuộc thấp nhất là bao nhiêu nếu ông ta là kẻ bắt cóc. Không phải điều dễ dàng đối với một sĩ quan Anh hay Pháp có thể đoán được mức độ hình phạt phải khủng khiếp đến mức nào mới có thể răn đe được anh ta nếu như anh ta là một tên khủng bố Mau Mau hoặc Algeria. Thật khó cho một người con trai có thể đoán nếu anh ta là cô gái mà anh ta muốn hẹn hò, thì anh ta sẽ được yêu đến mức độ nào, hoặc thật khó cho khách hàng tại một hiệu ăn để biết rằng anh ta sẽ ghét cái cảnh ở đây như thế nào nếu anh ta ở địa vị người phục vụ bàn.

Đây là một trong những lý do vì sao lời nói không thể thay thế được nước đi. Nước đi ở khía cạnh nào đó có thể làm thay

đổi cả trò chơi, bằng cách gây ra các chi phí dễ thấy, rủi ro, hay làm giảm phạm vi lựa chọn về sau. Nước đi mang theo nội dung thông tin, hay *bằng chứng*, có tính chất khác với thông tin do lời nói mang lại. Lời nói có thể rẻ tiền còn nước đi thì không (trừ phi “lời nói” được thực hiện dưới dạng các lời đe dọa, lời hứa, cam kết v.v... mà có thể đảm bảo thực thi được, và điều đó dù sao cũng sẽ được phân tích theo hướng *nước đi* hơn là trao đổi thông tin). Sự thỏa hiệp lẫn nhau cuối cùng đòi hỏi phải có, nếu để kết cục mang tính hiệu quả, sự phân chia lợi ích phù hợp với “lợi thế tương đối,” tức là, cái mà một bên nhân nhượng phải là cái mà anh ta muốn nó ít hơn đổi phương, so với cái mà anh ta muốn qua trao đổi. Do đó mỗi bên cần trao đổi hệ thống giá trị của mình với sự trung thực nhất định, mặc dù mỗi bên cũng có thể có lợi bằng sự lừa gạt. Mặc dù các thủ đoạn của một bên là mơ hồ trong việc bộc lộ các hệ thống giá trị của mình và thậm chí còn là sự lừa dối cố tình, chúng vẫn có tính bằng chứng mà lời nói thuần túy không có được.

Tính bất định mà thường được coi là hiện hữu đối với các hệ thống giá trị của mỗi bên cũng làm giảm sự hữu ích của khái niệm *đối xứng* toán học như là một nguyên lý chuẩn tắc hay dự đoán. Sự đối xứng toán học không thể được nhận thức nếu người ta chỉ tiếp cận một nửa số đại lượng có liên quan. Ở phạm vi mà tính đối xứng có ích cho những người chơi trong việc nhường nhịn các động thái cho nhau, thì nó sẽ có xu hướng đối xứng về định tính hơn, loại phụ thuộc vào bối cảnh hữu hình hơn là các giá trị ngầm.

## 5

## SỰ THỰC THI, LIÊN LẠC VÀ CÁC NƯỚC ĐI CHIẾN LUỢC

Mỗi khi chúng ta nói về sự răn đe, hăm dọa nguyên tử, cân bằng sức hủy diệt (balance of terror), hay thỏa thuận những vùng trời mở để giảm nỗi lo sợ tấn công bất ngờ; khi chúng ta mô tả quân đội Mỹ tại châu Âu như đội quân tiền tuyến (trip wire) hay khung cửa kính chống đạn [cho nước Mỹ], hoặc khi ta dành cho kẻ thù ở thế yếu một sự rút lui trong danh dự; khi chúng ta nói đến sự bất khả thi của một hăm dọa khi nó tàn khốc tới mức kẻ đưa ra hăm dọa sẽ không dám thực hiện, hay quan sát cảnh các tay lái xe taxi luôn được mọi người tránh xa vì chúng có tiếng là không bận tâm đến chuyện va quẹt hay làm trầy xước xe, thì rõ ràng là chúng ta đang đi sâu vào lý thuyết trò chơi. Tuy vậy lý thuyết trò chơi chính thức đóng góp rất ít vào việc làm rõ các ý tưởng này. Tác giả cho rằng lý thuyết trò chơi tổng-lợi ích-khác-không có thể đã đánh mất một lĩnh vực hứa hẹn nhất của nó do bị dính chặt vào sự phân tích quá trừu tượng. Do trừu tượng hóa các hệ thống thông tin liên lạc và thực thi và do quan niệm rằng tính chất

đối xứng hoàn toàn giữa những người chơi là trường hợp chung hơn là trường hợp riêng, lý thuyết trò chơi có thể đã đi quá đích mà tại đó công việc bổ ích có thể đã được thực hiện và có thể đã bỏ sót một số thành phần quan trọng của các trò chơi có tổng lợi ích khác không điển hình. Do tập trung quá nhiều vào giải pháp cho trò chơi có tổng lợi ích khác không, lý thuyết trò chơi đã không công bằng với một số tình huống trò chơi hoặc các mô hình trò chơi nhất định, và [cũng không công bằng] đối với “các nước đi” mang tính đặc thù của trò chơi chiến lược có tổng lợi ích khác không.

Ví dụ, “mô hình” nào phản ánh đúng đắn nhất trong các tranh luận về sự trả đũa ô ạt? Những điều kiện nào là cần thiết cho một sự đe dọa hiệu nghiệm? Cái gì trong lý thuyết trò chơi tương ứng với tình huống mà ai cũng biết “bắt gấu đằng đuôi”; làm thế nào để chúng ta nhận diện được ma trận lợi ích, hệ thống liên lạc, và hệ thống bảo đảm thực thi của trò chơi? Những chiến thuật nào có thể giúp những người đi bộ làm cho cảnh lái xe ô-tô sợ hãi, hay các nước nhỏ làm cho nước lớn sợ hãi; bằng cách nào để chúng ta diễn đạt nó bằng các thuật ngữ của lý thuyết trò chơi? Cơ cấu thông tin hay liên lạc nào, hoặc phức hợp động cơ nào làm cho những chú chó, những kẻ đàn độn, những đứa trẻ, những người cuồng tín, và những người tử vì đạo trở nên miễn dịch đối với những sự đe dọa?

Chiến lược báp bênh của chiến tranh lạnh và sự bế tắc hạt nhân đã thường xuyên được diễn đạt bằng một dạng trò chơi tương tự: có hai kẻ thù, mỗi người đứng trong tầm bắn cung tên tầm thuốc độc của người đối diện phía bên kia khe núi, thuốc độc có tác dụng chậm đến nỗi một bên vẫn kịp bắn bên

kia trước khi chết;<sup>1</sup> một người chăn cừu rượt đuổi một con sói đến cùng đường nơi nó không còn cách nào khác ngoài chống cự lại, còn người chăn cừu thì không muốn quay lưng trước con thú; một kẻ truy đuổi được trang bị duy nhất một quả lựu đạn trong tay vô tình ở rất sát với nạn nhân và không dám sử dụng vũ khí của mình; hai người láng giềng, mỗi người kiểm soát chất nổ ở tầng hầm của bên kia, đang cố tìm kiếm sự an toàn chung thông qua việc bố trí các công tắc điện và các ngòi nổ.<sup>2</sup> Nếu chúng ta có thể phân tích các cấu trúc của các trò chơi này và phát triển sự hiểu biết bằng các mô hình tiêu chuẩn, thì chúng ta có thể cung cấp những hiểu biết về các vấn đề thực tế bằng cách sử dụng lý thuyết của chúng ta.

Để minh họa, [ta hãy xem xét] một mô hình minh họa trong đó 20 người bị cướp hoặc đòi tiền chuộc chỉ bởi một kẻ duy nhất có súng và sáu viên đạn. Họ có thể không chế hắn ta nếu họ dám liều chết sáu người trong số họ, nếu họ có phương tiện để quyết định sáu người nào sẽ hy sinh. Họ có thể đánh bại hắn mà không mất mát gì nếu họ có thể biểu lộ được sự cương quyết sẽ làm điều đó, nếu họ có thể đồng thời cam kết *hứa sẽ* tránh dùng hình phạt tử hình một khi họ bắt được hắn. Hắn ta có thể ngăn chặn được sự đe dọa của họ, nếu hắn chứng tỏ quyết tâm rõ ràng sẽ bắn bất chấp mọi sự đe dọa của họ, hoặc hắn có thể chứng tỏ rằng hắn không thể tin vào lời hứa

---

<sup>1</sup> Đối chiếu với C. W. Sherwin, "Giữ hòa bình bằng kỹ thuật quân sự" (*Securing Peace Through Military Technology*), *Bulletin of the Atomic Scientists*, 12:159-164 (tháng 5-1956)

<sup>2</sup> Đối chiếu với Herman Kahn và Erwin Mann, "Lý thuyết trò chơi" (*Game Theory*), RAND Corporation, Paper, P-1166 (Santa Monica, 1957), trang 55. Các tác giả nghiên cứu một số vấn đề bao gồm chất nổ, ngòi nổ, và sự ngăn chặn.

của họ. Nếu họ không thể đưa ra được sự đe dọa của mình - chẳng hạn, nếu hắn không hiểu ngôn ngữ của họ - thì họ không thể giải giáp hắn bằng lời nói được. Họ cũng không thể đưa ra đe dọa trừ phi họ tự mình thỏa thuận với nhau; do đó nếu hắn ta có thể đe dọa sẽ bắn bất cứ hai người nào nói chuyện với nhau, thì hắn có thể ngăn chặn được thỏa thuận này. Nếu 20 người này không thể tìm ra cách để chia sẻ rủi ro, sẽ không có ai là người ra tay trước để thực hiện sự đe dọa, từ đó không có cách nào làm cho sự đe dọa có sức thuyết phục; và nếu hắn ta có thể tuyên bố một công thức bắn, chẳng hạn ai cử động trước sẽ bị bắn trước, hắn có thể răn đe được họ trừ phi họ có cách cùng nhau hành động mà không có ai là người "đầu tiên." Nếu 14 trong số 20 người có thể áp đảo sáu người còn lại và ép họ phải tiến lên trước, họ có thể chứng tỏ được rằng họ sẽ không chế được hắn; nếu như vậy, sự đe dọa sẽ thành công và tên cầm súng sẽ đầu hàng, và ngay cả sáu người "có thể bị hy sinh" kia cũng giành được thắng lợi chỉ vì họ không có khả năng tự giúp mình thoát nạn. Nếu 20 người có thể không chế hắn nhưng không có cách nào để hắn trốn thoát, thì lời hứa miễn trừ phạt có thể là cần thiết; nhưng nếu họ không thể từ chối tư cách của họ để nhận diện hắn ta và làm chứng chống lại hắn về sau, thì có thể cần thiết để hắn bắt một con tin. Điều này, đến lượt nó, phụ thuộc vào khả năng của 19 người kia có thể thực thi thỏa thuận của mình để bảo vệ, bằng cách giữ im lặng, bắt kẻ ai hiện là con tin v.v... Khi chúng ta đã nhận diện những thành phần quan trọng trong nhiều trò chơi kiểu này chúng ta có thể ở vị trí thuận lợi hơn để hiểu được cơ sở quyền lực của một kẻ chuyên quyền

không được lòng dân, hay một thiểu số chi phái được tổ chức tốt, hay các điều kiện cho cuộc binh biến thành công.

Chương này sẽ cố gắng đưa ra các nước cờ điển hình và các yếu tố cấu trúc đáng được khảo sát trong lý thuyết trò chơi. Chúng bao gồm những nước cờ như “đe dọa,” “hứa hẹn,” “phá hủy liên lạc,” “trao quyền quyết định,” v.v... và các yếu tố cấu trúc như các điều khoản về thông tin liên lạc và đảm bảo thực thi.

### *MỘT NƯỚC ĐI MINH HỌA*

Một ví dụ về “nước đi” chuẩn là sự cam kết, đã được phân tích khá kỹ ở Chương 3. Nếu môi trường thể chế cho phép một người mua tiềm năng đưa ra lời chào mua “cuối cùng” với điều kiện chịu sự trừng phạt khắc nghiệt trong trường hợp anh ta sửa đổi lời chào mua của mình - để tự cam kết - người bán vẫn còn một quyết định khác trừ quyết định duy nhất được cân nhắc kỹ của mình: bán với cái giá được đề nghị hoặc không bán. Khả năng cam kết sẽ biến một tình huống mặc cả không xác định thành một trò chơi có hai nước đi; một người chơi đưa ra sự cam kết, và người kia đưa ra quyết định cuối cùng. Trò chơi trở thành trò chơi xác định.<sup>3</sup>

---

<sup>3</sup> Trong ví dụ về bất động sản Von Neumann và Morgenstern đã được nêu ở phần trước (trang 116) người mua B (người đưa ra mức giá cao nhất là 15) có thể nâng giới hạn mà anh ta có thể rút ra từ người mua C (giá cao nhất đưa ra là 25) nếu anh ta có thể tìm ra được cách để ràng buộc mình mua ngôi nhà với giá 20 và giữ lại ngôi nhà hoặc đập đi (nghĩa là, để không bán lại cho C với cái giá bị lỗ) trừ phi anh có được

Nước đi cụ thể này, đã được phân tích kỹ ở Chương 3, được nhắc lại ở đây chỉ như là một sự minh họa đơn giản của một nước đi điển hình. Như đã lưu ý ở Chương 3, sự tồn tại cũng như tính hiệu quả của nước đi này phụ thuộc vào cấu trúc thông tin liên lạc của trò chơi và khả năng của người chơi tìm ra cách để tự cam kết, để "bảo đảm thực thi" sự cam kết đối với chính bản thân anh ta. Hơn nữa, chúng ta đã cho phép cơ cấu nước đi của trò chơi là bất đối xứng; "người thắng" là người có được sự cam kết hoặc, nếu cả hai đều có thể, là người cam kết trước. (Chúng ta có thể xem xét trường hợp đặc biệt là hòa, nhưng chúng ta sẽ không lấy kết quả hòa là kết luận đương nhiên, do giả thiết về đối xứng.)

Nhưng, dù chúng ta đã làm cho trò chơi "xác định" theo cái nghĩa rằng chúng ta sẽ không có khó khăn trong việc nhận diện "giải pháp," mỗi khi chúng ta đã nhận diện được ai trong số hai người chơi có thể tự cam kết trước, thì nó vẫn là một trò chơi *chiến lược*. Mặc dù người thắng là người đạt được sự cam kết trước, trò chơi không giống như cuộc chạy đua bằng chân mà ai chạy nhanh nhất là người thắng. Sự khác biệt là sự cam kết không tự dung chiến thắng theo luật chơi, cả về phương diện vật lý lẫn pháp lý. Kết quả vẫn phụ thuộc vào

---

phần lớn, chẳng hạn 20-*P* - *P* là giá cuối cùng mà C trả. Tức là, B thay đổi giá "thực" cao nhất của mình, nhờ đó nâng giới hạn mà anh ta có thể rút được từ C. Tất nhiên, D và E cũng có thể tìm cách làm tương tự; và người đầu tiên nhận được cam kết thích hợp, hay người tìm ra cách làm nếu chỉ có một người trong số họ có thể làm được, sẽ là người thắng. Nếu D, người không gắn giá trị bản thân vào ngôi nhà, cam kết trả 22, anh ta sẽ là một thành viên đích thực của trò chơi với cái giá đăng ký thực sự 22; tính đích thực của anh ta sẽ còn lớn hơn cả B lúc ban đầu, nếu như sự cam kết có thể chứng minh được mặc dù sự định giá chủ quan thì không.

người chơi thứ hai, mà người chơi thứ nhất không thể kiểm soát trực tiếp. Sự cam kết [của người chơi thứ nhất] là một nước đi *chiến lược*, một động thái kích thích người chơi thứ hai đưa ra sự lựa chọn có lợi cho mình. Nó hạn chế sự lựa chọn của người chơi thứ nhất bằng cách gây ảnh hưởng đến kỳ vọng của anh ta [về hành vi của người thứ hai và kết quả cuối cùng].

Khi năng cam kết của một người trong trò chơi kiểu này là tương đương với “nước đi đầu tiên.” Và nếu khung khổ thể chế không có các phương tiện bảo đảm tính không thể thay đổi của cam kết, theo cái nghĩa pháp lý hay hợp đồng, một người có thể thực hiện được điều giống như vậy bằng một hành vi không thể đảo ngược [mà hành vi này] sẽ làm giảm quyền tự do lựa chọn của anh ta. Một người có thể tránh được hậu quả không mong muốn của việc không thể bảo đảm cam kết của mình bằng cách chủ động thực hiện trước một hành vi nào đó mà nếu [sau này] thất bại anh ta sẽ tự hại mình. Luce và Raiffa đã chỉ ra rằng chiến thuật giống như vậy có thể được dùng để chống lại chính mình khi anh ta muốn, chẳng hạn, ăn kiêng nhưng lại không tin vào [khả năng tự nhịn ăn nhịn uống của] mình. “Anh ta tuyên bố ý định của mình cho mọi người, hoặc đánh cược rằng anh ta sẽ không từ bỏ ăn kiêng, để sau đó anh ta sẽ *không* thể tự tiễn đổi ý và làm theo cái dạ dày của anh ta muốn vào lúc *đó*.<sup>4</sup> Tương tự như thế, việc một người cố tình thực hiện một chuyến đi nghỉ trong rừng sâu mà không mang theo thuốc lá cũng là một dạng hành vi chứ không phải một dạng cam kết.

---

<sup>4</sup> Games and Decisions (Trò chơi và Quyết định), trang 75.

## NHỮNG ĐE DỌA

Đặc tính nổi bật của một đe dọa là ở chỗ người đe dọa khẳng định rằng anh ta sẽ làm, trong một tình huống bắt buộc, cái điều mà anh ta không muốn làm nếu tình huống bắt buộc đó xảy ra, tình huống bắt buộc này bị quyết định bởi hành vi của bên thứ hai. Giống như một cam kết bình thường, đe dọa là việc [tự] hủy bỏ lựa chọn, từ bỏ các giải pháp khác. [Việc tự hủy bỏ các lựa chọn khác này] sẽ làm cho người đe dọa bất lợi hơn trong trường hợp chiến thuật này bắt thành. Cả sự đe dọa lẫn sự cam kết đều được xây dựng trên khả năng có thể ép buộc được người chơi thứ hai nhận thức rằng người chơi thứ nhất đã thay đổi động cơ hành động của mình. Giống như cam kết bình thường, đe dọa chỉ có thể ép buộc được người chơi kia chừng nào mà nó chuyển đến người kia chí ít một số dáng dấp của nghĩa vụ; nếu tôi đe dọa nổ tung cả hai chúng ta trừ phi bạn đóng cửa sổ, thì bạn sẽ biết rằng tôi sẽ không làm như vậy trừ phi tôi đã tự đẩy mình vào chỗ không còn lựa chọn nào khác<sup>5</sup>.

Tuy nhiên sự đe dọa khác với sự cam kết bình thường, ở chỗ nó làm cho hành động của một người lệ thuộc vào hành

<sup>5</sup> Trong ngôn ngữ hàng ngày, "đe dọa" cũng thường được dùng đối với trường hợp trong đó một người chỉ đơn thuần nói với một đối thủ, hoặc nhắc nhở anh ta, rằng người đó sẽ có hành động gây đau đớn cho đối phương nếu anh ta không tuân theo, điều rõ ràng là người đó đã có động cơ để làm điều đó. Việc "dọa" sẽ gọi cảnh sát bắt một người đột nhập là thuộc loại này, còn việc đe dọa sẽ bắn anh ta lại không phải. Nhưng có vẻ sẽ tốt hơn nếu dùng từ ngữ khác cho các trường hợp này - tác giả gợi ý dùng "cảnh báo" hơn là "đe dọa" - vì "đe dọa" hoặc là hời hợt, và không cấu thành một hành động, hoặc là truyền đi một thông

động của người kia. Trong khi sự cam kết gắn vào hành động của chỉ một bên, sự đe dọa gắn vào sự đối phó, sự phản ứng của bên kia. Sự cam kết là một biện pháp để giành được nước đi đầu tiên trong trò chơi mà nước đi đầu tiên tạo ra lợi thế; còn đe dọa là sự cam kết đối với chiến lược cho nước đi thứ hai.

Do đó sự đe dọa chỉ có thể có hiệu quả khi nào trò chơi là một loại trong đó nước đi đầu tiên thuộc về bên kia hoặc một bên có thể ép bên kia đi trước. Nhưng nếu một bên phải, theo ý nghĩa cơ học, đi trước hay đi đồng thời với bên kia, anh ta vẫn có thể ép bên kia có “nước đi đầu tiên” tương đương về pháp lý bằng cách gắn vào sự đe dọa của mình đòi bên kia cam kết trước anh ta sẽ hành động như thế nào – nếu trò chơi có các cấu trúc thông tin và cưỡng chế làm cho lời hứa có tính khả thi và bên bị đe dọa không thể hủy bỏ trước. Kẻ tấn công mà nạn nhân giàu có của hắn lại không mang tiền theo chặng lợi dụng được cơ hội đã có trừ phi hắn có thể bắt nạn nhân làm con tin trong khi chờ trả tiền chuộc; và thậm chí điều đó cũng sẽ không có kết quả trừ phi hắn có thể đưa ra cam kết

---

tin thực và có liên hệ đến những tình huống với một cơ cấu thông tin và cơ cấu liên lạc đáng được phân loại riêng biệt. Trong trường hợp sau nó là một nước đi có lợi chung, loại trừ kết quả cả hai bên không mong muốn nhờ cải thiện sự hiểu biết của bên thứ hai. Sự tương đồng chính giữa trường hợp “cảnh báo” và “đe dọa,” là những khó khăn có thể có trong việc truyền đi thông tin thực một cách đáng tin cậy, trong việc truyền đi bằng chứng để khẳng định một bên đã có, như đã nói, động cơ để làm như anh ta cảnh báo sẽ làm. Thực ra, nếu sự đe dọa có tính chất như (như nó thường thấy) hành động cam kết không song hành với hành động truyền đạt thông tin – nếu sự cam kết diễn ra trước khi truyền đạt thông tin, với những bằng chứng đáng tin cậy, đến bên bị đe dọa – thì hành động cam kết đầu tiên trong quá trình đe dọa sẽ làm thay đổi cơ cấu động cơ “thực sự,” và hành động thứ hai [là hành động truyền tin] trên thực tế [chỉ] là một “cảnh báo.”

thuyết phục rằng sẽ trả lại con tin theo cách khiến hắn không bị nhận diện hay bị bắt.

Sự thật là *một số* loại cam kết, hoặc chí ít là có vẻ như cam kết, phải đứng sau sự đe dọa và được thông tin đến bên bị đe dọa một cách thành công lại mâu thuẫn với một ý tưởng khác thường xuất hiện trong lý thuyết trò chơi. Đó là ý tưởng cho rằng sự đe dọa là điều đáng mong muốn, hay có thể chấp nhận được, hay có vẻ hợp lý vì phản ứng của bên bị đe dọa chỉ gây tổn thất tồi tệ hơn cho chính anh ta hơn là tổn thất cho bên đưa ra đe dọa. Đây là quan điểm của Luce và Raiffa, là những người mô tả đe dọa bằng cụm từ, "Nó sẽ làm anh đau nhiều hơn là làm tôi đau," cho thấy rõ việc đưa ra đe dọa phụ thuộc vào sự so sánh tác dụng của nó đối với hai bên. Trong trường hợp cả hai người chơi tìm cách đưa ra những đe dọa có vẻ hợp lý, các tác giả (Luce và Raiffa) cho rằng kết quả sẽ trở nên không xác định được, mà phụ thuộc vào "tính cách mặc cả" của những người chơi; "và việc phỏng đoán điều gì sẽ xảy ra mà không dựa trên các phân tích đầy đủ về tâm lý và kinh tế của những người chơi là việc làm ngu ngốc."<sup>6</sup>

<sup>6</sup> Trang 110-11, 119-20, 143-44. Morton A. Kaplan, khi vận dụng lý thuyết trò chơi vào quan hệ quốc tế cũng cho rằng "mọi tiêu chí làm tăng trọng lượng cho lập trường đe dọa của các bên đều có liên quan đến việc so sánh tác dụng của nó đối với các bên." (Xem *System and Process in International Politics* (Hệ thống và quá trình trong chính trị quốc tế) [New York, 1957].) Luce và Raiffa một phần có thể đã bị dẫn dắt đi đến quan điểm rằng chỉ có một trong số người chơi là có sự đe dọa "có vẻ hợp lý," bằng cách giới hạn bài thảo luận tóm tắt của họ vào các ma trận  $2 \times 2$ . Không thể biểu thị, bằng ma trận  $2 \times 2$ , một trò chơi trong đó cả hai người chơi đều có lợi ích trong việc đưa ra đe dọa. Về cơ bản sự đe dọa là một tuyên bố đáng tin cậy về một sự lựa chọn có điều kiện cho nước đi thứ hai. Nó chỉ có lợi khi nếu nó đem lại phần thưởng lớn hơn so với hoặc là nước đi đầu tiên hoặc là nước đi thứ hai đơn độc và khi một bên có thể làm cho bên kia thực sự đi trước hoặc phải đưa ra lời

hứa trước. (Nếu nước đi thứ hai đơn độc là tốt, thì sự đe dọa là không cần thiết; và nếu nước đi thứ nhất là tốt, thì một bên chỉ cần đưa ra cam kết chiến lược của mình vào ~~một~~ lựa chọn vô điều kiện, chứ không phải sự cam kết đối với một lựa chọn có điều kiện.) Nhưng nếu thứ tự ưu tiên này đúng cho một bên trong một ma trận  $2 \times 2$ , nó không thể đúng cho bên kia. Các ma trận thực sự do Luce và Raiffa sử dụng khi trao đổi quan điểm không cho thấy chiến lược đe dọa "có vẻ hợp lý" đối với người chơi số 2, không phải vì kích thước tuyệt đối của cái được hay cái mất của anh ta lớn hơn của người chơi số 1 mà là vì lý do đơn giản hơn nhiều rằng người chơi số 2 không sử dụng đe dọa. Anh ta sẽ thắng nếu anh ta đi nước đầu tiên; anh ta cũng thắng nếu đi nước thứ hai; và anh ta thắng với các nước đi đồng thời như cho thấy trong trò chơi. Lợi ích duy nhất của anh ta trong tuyên bố mang tính đe dọa là để ngăn chặn sự đe dọa của đối phương; và vì lý do anh ta chỉ cần một sự cam kết vô điều kiện đối với chiến lược được ưu tiên của anh ta - nghĩa là, sự tương đương pháp lý của "nước đi đầu tiên" trước sự đe dọa của đối phương. Chiến thuật "đe dọa" của J. F. Nash, vận dụng cho các trò chơi mặc cả đã đem lại nhiều kết quả hiệu nghiệm liên tiếp - hoặc có thể làm được điều đó, bằng thỏa thuận về kết quả rút thăm - sẽ khác với sự đe dọa được thảo luận ở đây, ở chỗ người đe dọa không đòi hỏi, trong trường hợp có tổn thất chung, một kết quả cụ thể mà chỉ *một số* kết quả cụ thể trong phạm vi hữu hiệu; nghĩa là, anh ta thay đổi điểm không tương ứng với "không có thỏa thuận." Động cơ đối với sự đe dọa đó là sự kỳ vọng về một kết quả cụ thể xác định bằng toán học mà tâm điểm được thay đổi bởi sự thay đổi về lợi ích tương ứng với trường hợp không có thỏa thuận. Đây là một loại đe dọa được giả định bởi Luce và Raiffa (trang 139) trong trò chơi "bất đối xứng." Cấu trúc pháp lý ngầm của trò chơi hình như thích hợp với các cam kết không thể thay đổi (nếu không, cam kết đầu tiên sẽ dễ dàng đem lại thắng lợi trong trò chơi cho bên này hoặc bên kia). Mỗi bên đều chịu sự chi phí của "sự bất lực" pháp lý mà anh ta luôn luôn có thể, bằng hành động công khai theo thỏa thuận với đối tác của mình về mọi kết quả, lẩn tránh sự cam kết của mình. Bằng cách đó, những cam kết có thể thay đổi chỉ có thể chuyển về điểm không - "nguyên trạng" trừ phi đạt được một thỏa thuận công khai về một số kết quả. Sự "bất đối xứng" xuất hiện trong trò chơi do Luce và Raiffa đưa ra là một đặc điểm của hệ thống pháp lý cụ thể tồn tại ẩn [trong cấu trúc của trò chơi mà họ đưa ra]. Trong thực tế nó có thể tương ứng với, chẳng hạn, việc cố tình gây ra sự phản đối của xã hội đối với việc không đạt được thỏa thuận. Sự phản đối như vậy sẽ tạo ra chi phí hay sự trừng phạt (có thể bất đối xứng giữa những người tham gia trò chơi) ngoài chi phí của việc không đạt được thỏa thuận, nhưng đối với công chúng không quan tâm đến nội dung của thỏa thuận mà chỉ quan tâm đến việc có đạt được thỏa thuận hay không.

	I	II
i	2 1	1
ii	0 0	0 0

	I	II
i	10 9	9 10
ii	8 0	0 0

Nhưng vấn đề lại vừa đơn giản hơn và chính xác hơn điều đó. Hãy xem xét ma trận bên tay trái trên Hình 9, nơi mà Cột được giả định là có “nước đi đầu tiên.” Không có sự đe dọa, Cột “thắng” một cách dễ dàng. Anh ta chọn chiến lược I, buộc Hàng phải chọn giữa các gói lợi ích 1 hoặc 0; Hàng chọn chiến lược i, đem đến cho Cột lợi ích bằng 2. Nhưng nếu chúng ta cho phép Hàng đưa ra đe dọa, anh ta tuyên bố rằng anh ta sẽ chọn chiến lược ii trừ phi Cột chọn II; nghĩa là, anh ta cho Cột sự lựa chọn [giữa hai tổ hợp] (ii,I) hoặc (i,II) bằng việc tự cam kết đối sẽ thực hiện chiến lược lựa chọn có điều kiện. Nếu Cột đi trước và chọn I, tất nhiên, Hàng sẽ ưu tiên chọn I; và cả hai đều biết điều đó. Chiến thuật [đe dọa] thành công chỉ khi nào Cột tin rằng Hàng *buộc phải* chọn ii trong trường hợp [Cột chọn] I.

Hoặc anh ta tin điều này, hoặc anh ta không tin. Nếu anh ta không tin, thì “sự đe dọa” chẳng có nghĩa gì đối với anh ta cả; anh ta tiếp tục đi tiếp và chọn nước đi đầu tiên “tốt nhất” của mình, bằng cách chọn I. Nếu anh ta tin rằng Hàng buộc phải đi theo chiến lược (i,II) hoặc (ii,I), Cột muốn 1 hơn 0 và chọn II. Nhưng điều này cũng đúng với mọi chữ số mà chúng ta có thể đưa vào ma trận phản ánh cùng thứ tự ưu tiên. Điều đó cũng đúng với ma trận thứ hai ở bên tay phải. Ma

trận thứ hai khuếch đại tính chất cơ bản của sự đe dọa nhiều hơn ma trận đầu, vì sự trừng phạt Hàng do sự lựa chọn bất hợp lý của Cột là lớn hơn trong trường hợp này; nhưng đối với việc chơi hợp lý và có thông tin đầy đủ, Hàng không cần phải lo lắng. Sự ưu tiên của Cột là rõ ràng; và một khi Hàng đã cho anh ta một cặp để lựa chọn - (ii,I) và (i,II) - thì không có gì nghi ngờ về điều mà Cột sẽ làm. Nếu tôi dọa sẽ làm vỡ sọ của mình làm vấy tú tung lên bộ đồ mới của bạn đang mặc nếu bạn không cho tôi lát bánh mì nướng cuối cùng, thì bạn sẽ đưa ngay cho tôi lát bánh mì đó hoặc không đưa tùy thuộc vào việc bạn có biết tôi có định làm việc đó hay không, cũng đúng như trường hợp khi tôi chỉ dọa sẽ hất trứng khuấy vào bạn.<sup>7</sup>

Vấn đề ở đây là chúng ta có thừa nhận hay không trò chơi có "các nước đi," nghĩa là, liệu có thể một người chơi hay cả

---

<sup>7</sup> Edward Banfield chỉ cho tôi trích dẫn về Bháts và Charáns của tây Ấn Độ, được tôn kính như là các thi sĩ. "Tại Guzerát họ vận chuyển số tiền lớn bằng vàng nén đi qua các vùng đất rộng nơi mà lực lượng tháp tùng manh cung không đủ để bảo vệ. Họ cũng đạt thỏa thuận giữa những người cầm đầu của họ và thậm chí với chính phủ.

"Sức mạnh của họ xuất phát từ sự thánh thiện của tính cách và quyết tâm mãnh liệt. Nếu một người đàn ông mang theo của cải bị tiếp cận, anh ta tuyên bố rằng anh ta sẽ tự hành hạ mình (trága), như tên nó được gọi như vậy: hoặc nếu không được [người khác] tuân thủ, anh ta sẽ đưa ra đe dọa tương tự trừ phi nó được thực hiện. Nếu anh ta không được quan tâm, anh ta sẽ dùng dao găm rạch chân tay mình, nếu mọi biện pháp thất bại anh ta sẽ đâm vào tim mình; hoặc anh ta sẽ chặt đầu đứa con của mình trước; hoặc bằng những sự bảo đảm khác đối với thỏa thuận ai sẽ là người đầu tiên bị chặt đầu bởi những người đồng hành. Sự xỉ nhục của những hành động này, và nỗi sợ hãi có vết máu của thi sĩ trên đầu họ nói chung, làm cho kẻ bướng bỉnh nhất cũng trở nên biết điều. Sự trung thành của họ là mẫu mực, và họ không bao giờ ngần ngại hy sinh tính mạng của mình để hiến dâng cho sự nghiệp mà tầm quan trọng của đẳng cấp của họ phụ thuộc vào đó" (Ngài Mountstuart Elphinstone, *History of India* (Lịch sử Ấn Độ) [in lần 7; London, 1889], trang 211).

hai người có hành động trong quá trình diễn ra trò chơi mà làm thay đổi bản thân trò chơi một cách không thể đảo ngược hay không – một cách nào đó làm thay đổi ma trận phản thưởng, trật tự của các lựa chọn, hoặc cấu trúc thông tin của trò chơi. Nếu trò chơi theo định nghĩa của nó không chấp nhận bất cứ kiểu nước đi nào, trừ nước đi có sự thỏa thuận chung và từ chối thỏa thuận, thì có thể đúng là “những tính cách” của người chơi sẽ quyết định kết quả, theo cái nghĩa rằng những kỳ vọng của họ trong một trò chơi “không có nước đi” sẽ hội tụ bởi một quá trình hoàn toàn tâm lý. Nhưng, nếu sự đe dọa còn là cái gì khác ngoài sự khẳng định quyết tâm được chuyển đến đối phương, chúng ta có thể hỏi cái đó là gì. Và nếu có, thì cái đó hẳn có liên quan đến một số ý tưởng về sự cam kết – thật hay là giả.

“Sự cam kết” sẽ được giải thích một cách khái quát ở đây. Nó bao gồm những hành vi giúp cho một bên ở vào tình thế mà không thể không thực hiện cam kết (như khi một lái xe hăm dọa xe khác bằng cách tự lái quá nhanh để không thể dừng đúng lúc được); những hành vi [của người thứ nhất] làm thay đổi quyết định cuối cùng [mà người thứ hai] không thể sửa khi kết cấu động cơ của người thứ hai thúc đẩy anh ta thực hiện đúng cam kết sau hành vi của người thứ nhất (chẳng hạn như khi [bên này] có tình trao thẩm quyền trừng phạt cho những kẻ tàn ác [làm cho bên kia sợ hãi mà phải thực hiện cam kết], hoặc khi một người chuyển các đòi hỏi đèn bù bảo hiểm và nghĩa vụ [trả tiền bảo hiểm] của anh ta cho một công ty bảo hiểm); những hành vi của một bên mà [tác dụng của nó] chỉ “làm suy giảm” lợi ích của chính họ trong tình huống họ không thực hiện cam kết, sao cho ngay cả khi việc thực

hiện cam kết dẫn tới những hậu quả có hại cho cả hai giờ đây cũng trở nên hấp dẫn hơn là không thực hiện cam kết (như khi một bên sắp đặt để bản thân bị bêu riếu trước công chúng là một thằng hèn nếu anh ta không thực hiện cam kết, hay khi anh ta đặt cửa sổ gắn tấm kính dày trước cửa hàng của mình, hoặc khi một bên tập trung phụ nữ và trẻ em vào một khu vực đặc biệt trong lãnh địa mà anh ta đã đe dọa – không được thuyết phục cho lắm – phía bên kia sẽ bảo vệ nó bất cứ giá nào). Một ví dụ thường ngày của Erving Goffman, người đã nhắc nhở chúng ta rằng “những người bán hàng, đặc biệt “những người bán hàng rong” trên phố, biết rằng nếu họ chọn cách bán khiến cho họ sẽ bị mất thẻ điện, trừ phi khách hàng miễn cưỡng mua, khách hàng có thể bị mắc bẫy bởi sự quan tâm chi tiêu cố và mua để giữ thẻ điện cho người bán và tránh dẫn tới cảnh thường thấy.”<sup>8</sup>

Tuy nhiên, có một số cách mà ý tưởng cam kết đe dọa có thể được nói lỏng có ích. Một là thừa nhận rằng cam kết “vững chắc” chung quy là phải có sự trùng phạt hoàn toàn hiệu nghiệm khiến cho trong mọi hoàn cảnh người cam kết sẽ muốn thực hiện điều mình đã cam kết. Chính sự trùng phạt với mức độ vô hạn (hoặc chí ít hơn mức cần thiết) khiến cho người ta chỉ còn biết gắn bó một cách tự nguyện, không lay chuyển và công khai với điều mà họ đã cam kết làm. Khái niệm này có thể được nói lỏng bởi giả định rằng sự trùng phạt có quy mô

---

<sup>8</sup>Bài viết của Goffman là một công trình nghiên cứu xuất sắc về mối liên quan của lý thuyết trò chơi với kỹ thuật chơi và là một minh họa đầu tiên về nội dung lý thuyết trò chơi phong phú liên quan đến các cấu trúc hành vi như nghỉ thức, phép lịch sự, thông lệ ngoại giao và – với ngữ ý – luật pháp.

hữu hạn và không nhất thiết phải lớn đến mức phải kiểm soát trong mọi trường hợp. Ở Hình 10, Cột sẽ thắng nếu anh ta có nước đi đầu tiên, trừ phi Hàng có thể tự cam kết chọn i. (Cam kết giành “nước đi đầu tiên” cho Hàng.) Nhưng, nếu sự cam kết [của Hàng] có nghĩa là gắn sự trùng phạt hữu hạn với sự lựa chọn hàng ii và chúng ta biểu thị điều này trong ma trận bằng cách trừ các phần thưởng của Hàng trong sự lựa chọn ii một số khẩn nhất định đại diện cho sự trùng phạt, thì chúng ta sẽ thấy sự cam kết chỉ có hiệu lực nếu mức trùng phạt là lớn hơn 2. Nếu không Cột sẽ thấy rõ là phản ứng của Hàng đối với II sẽ là ii, bất chấp sự cam kết. Trong trường hợp này sự cam kết chỉ là sự mất mát mà Hàng sẽ tự chuốc lấy, do đó anh ta sẽ tránh nó.

	I	II
i	2 4	1
ii	4 2	3

Hình 10

Tương tự với trường hợp đe dọa. Trên Hình 11 không có các đe dọa, giải pháp sẽ là tại (iii,II) bắt kể luật chơi đòi hỏi Hàng chọn trước, Cột chọn trước, hay cả hai chọn đồng thời. Mỗi người chơi có thể thắng nếu anh ta đi nước thứ hai và đổi mặt với đối phương bằng một đe dọa.<sup>9</sup>

<sup>9</sup> Nếu một người chơi, chẳng hạn Cột không thể ép buộc Hàng đi nước đầu tiên theo nghĩa cơ học, anh ta có thể làm như vậy theo nghĩa “pháp lý” bằng cách đe dọa lựa chọn I trừ phi Hàng hứa sẽ chọn ii. Sự phân

	I	II	III
i	-5      -5	-2      -1	-2      -1
ii	-4      -3	0      3	2      2
iii	-4      -3	1      1	3      0

Hình 11

Cột sẽ dọa chọn I nếu Hàng chọn iii, Hàng sẽ đe dọa chọn i nếu Cột chọn II. Nhưng nếu sự đe dọa được bảo đảm bởi một hình phạt, giới hạn dưới đối với bất kỳ hình phạt hợp lý nào mà Cột đòi sẽ là 4; bất kỳ hình phạt nhỏ hơn 4 sẽ làm anh ta ưa thích II hơn là I khi mà Hàng chọn iii. Giới hạn dưới đối với một hình phạt hợp lý cho sự không tuân thủ của Hàng sẽ là 3. Nếu khi đó, tình hình diễn ra là trường hợp trong đó các hình phạt có một "mức độ" duy nhất, một mức độ nhỏ hơn 3 sẽ vô tác dụng và kết quả sẽ là tại iii,II; một mức độ lớn hơn 4 là thích hợp cho cả hai bên, và "người chiến thắng" sẽ là

tích đầy đủ trường hợp này đòi hỏi phải chú ý đến các hình phạt đối với những hứa hẹn cũng như sự đe dọa. Do sự dàn xếp về mặt vật chất và thể chế đối với các hứa hẹn (nghĩa là, các cam kết đối với bên thứ hai) nói chung là hoàn toàn khác về bản chất so với những cam kết đơn phương (nghĩa là, các cam kết mà người chơi thứ hai không thể tự mình hủy bỏ được), những hình phạt có được có thể cũng sẽ khác hẳn nhau như giữa đe dọa và hứa hẹn - khái quát hơn, như sự khác nhau giữa người chơi thứ nhất và người chơi thứ hai. Những phần thưởng cụ thể được nêu ở Hình 4 có thể đòi hỏi những hình phạt ít nhất là 1 – 5 lời hứa của Cột hay Hàng. Lưu ý rằng trong trường hợp lời hứa được đưa ra bởi sự đe dọa, nó sẽ là một lợi thế cho người đe dọa có thể đòi hỏi hình phạt và là một bất lợi cho nạn nhân có thể bị hình phạt do sự vi phạm thỏa thuận của mình, nghĩa là phải tuân thủ.

người có thể đe dọa trước; quy mô giữa 3 và 4 chỉ có thể sử dụng cho Hàng, tức người chiến thắng. Trong trường hợp sau người chơi nào bị thiệt hại nhiều hơn bởi sự đe dọa không thành của mình chính là người không thể đe dọa - theo nghịch lý là anh ta không thể trút một hình phạt thích đáng lên đầu mình.

Lưu ý rằng sự so sánh “bị đau hơn” trong trường hợp này không phải nói về liệu Hàng hay Cột sẽ bị đau hơn bởi cái mà Hàng đe dọa mà là liệu Hàng sẽ bị đau hơn bởi việc thực hiện sự đe dọa của mình so với Cột hay không nếu thay vào đó Cột đưa ra đe dọa *của mình*. Quả thực, trong ma trận phần thưởng cụ thể như được thể hiện, sự đe dọa *thành công* của Hàng là sự đe dọa mà sẽ làm anh ta đau *hơn* khi thực hiện đe dọa đó so với Cột, trong khi sự đe dọa tiềm tàng *không thành công* của Cột sẽ làm anh ta đau *ít hơn* khi thực hiện nó so với Hàng.

Một cách nói lỏng khái niệm đe dọa nữa là thay đổi giả thiết của chúng ta về tính chất hợp lý. Giả sử với xác suất Pr dành cho người chơi R, và xác suất Pc cho người chơi C, họ sẽ mắc sai lầm, sẽ đi nước đi không hợp lý, hoặc sẽ có hành động không đoán trước được do người thứ hai không biết rõ lợi ích của người thứ nhất.<sup>10</sup> Điều này sẽ cho chúng ta một trò chơi trong đó cái được và cái mất có thể xảy ra khi một người tự cam kết với một đe dọa phải tính đến khả năng sự đe dọa được cam kết đầy đủ này rốt cuộc không được lắng nghe. Nếu khi đó, cái mất tiềm tàng- sẽ xảy ra do sự trùng phạt - đối với người này lớn hơn so với người kia, có thể có các tình huống đối xứng - các xác suất P bằng nhau và các mức trùng phạt là như nhau đối với hai người chơi - trong đó một người chơi

<sup>10</sup> Các tình huống loại này được khảo sát ở Chương 7 và 9.

có thể thấy có lợi khi đưa ra đe dọa và người chơi khác thì không như vậy, khi tính đến xác suất của “sai số.” (Một cách tính toán gần giống như vậy sẽ phải dùng đến trong tình huống nếu cả hai người chơi đều có cơ hội để thực hiện đe dọa, và tồn tại nguy cơ cả hai bên cùng cam kết vì lý do bên này không nhìn thấy rõ bên kia có cam kết hay không để dừng lại mà cứu nguy cho cả hai.)

Việc sửa đổi khái niệm đe dọa - mà điểm chính là sửa đổi giả định về tính duy lý ẩn trong đó - có phần đi theo hướng của chỉ tiêu “làm đau hơn.” Về tổng thể, mặc dù lý thuyết trò chơi làm tăng thêm tầm nhìn về chiến lược mặc cả bằng cách nhấn mạnh một chân lý nổi bật rằng sự đe dọa *không* phụ thuộc vào việc bên đe dọa chịu tổn thất ít hơn bên bị đe dọa nếu sự đe dọa phải được tiến hành thay vì bằng cách phóng đại sự thật có thể có trong án tượng trực giác đầu tiên. Nhưng sự đe dọa chiến tranh, đe dọa cái giá của chiến tranh, đe dọa làm hỏng bộ đồ; đe dọa gây “cãi lộn”; hầu hết các đe dọa của xã hội có tổ chức để truy tố tội phạm và hành động phi pháp; và các khái niệm về tống tiền và răn đe nhìn chung khó có thể hiểu được trừ phi phủ định tiêu chí so sánh lợi ích. Đúng là những bất đối xứng trong tình huống đe dọa, cũng như giữa hai người tham gia trò chơi, làm cho các đe dọa trở thành đe tài phong phú cho nghiên cứu; nhưng những bất đối xứng hợp lý bao gồm những cái trong hệ thống thông tin, trong khả năng thực thi các đe dọa và những lời hứa, trong tốc độ của cam kết, trong tính duy lý của các phản ứng được kỳ vọng, và cuối cùng (trong một số trường hợp) trong các tiêu chuẩn gây tổn thất tương đối.

## NHỮNG LỜI HỨA

Những lời hứa được bảo đảm thực thi không thể coi là điều dĩ nhiên. Những thỏa thuận cần phải được thể hiện dưới dạng có thể cưỡng chế được và bao gồm các loại hành vi có thể cưỡng chế được. Khả năng được bảo đảm thực thi phụ thuộc ít nhất vào hai điều - một quyền lực nào đó có thể thực hiện trừng phạt hay cưỡng bức, và một khả năng để nhận rõ liệu có cần phải trừng phạt hay cưỡng bức không. Những cuộc tranh luận sau chiến tranh về các đề nghị giải trừ quân bị và các kế hoạch giám sát cho thấy điều đó khó khăn như thế nào, ngay cả khi nếu cả hai rất mong muốn đạt được một thỏa thuận có thể cưỡng chế được hay tìm ra các biện pháp cưỡng chế có sức thuyết phục. Vấn đề trở nên phức tạp khi không bên nào tin vào bên kia và mỗi bên đều cho rằng không có ai tin vào ai và do đó không có ai có thể mong đợi sự tuân thủ của bên kia. Nhiều vấn đề kỹ thuật trong việc kiểm soát vũ khí sẽ biến mất nếu có các biện pháp chấp nhận được, làm cho lời hứa cưỡng chế được hoặc nếu tất cả các quốc gia trên thế giới đều ủng hộ vô điều kiện một đấng quyền lực siêu phàm nào đó. Nhưng vì sự không tuân thủ có thể sẽ không bị phát hiện, những lời hứa tuân thủ không thể được bảo đảm thực thi ngay cả khi sự trừng phạt được bảo đảm. Vấn đề được nhân gấp đôi bởi thực tế là sự trừng phạt không thể được bảo đảm, trừ phi sự trừng phạt đó có thể được đơn phương thi hành bởi bên kia đối với hành động vi phạm thỏa thuận ban đầu. Hơn nữa, một số thỏa thuận có vẻ như được mong muốn

lại bị bỏ xó vì không có điều khoản thực hiện; các thỏa thuận không gây bất lợi cho các bên chỉ được thi hành khi nó được quy định bằng những quy định khách quan có thể giám sát khách quan.

Những lời hứa thường được coi là những cam kết song phương (theo hợp đồng), được đưa ra như sự có đi có lại mà thường là sự đáp lại lời hứa. Nhưng cũng có động cơ cho một lời hứa đơn phương khi nó khiến bên kia muôn lựa chọn vì lợi ích chung. Trong ma trận ở phía tay trái trên Hình 12, nếu các lựa chọn đều tiến hành đồng thời, chỉ có một *cặp* lời hứa là có thể có hiệu quả; trong ma trận ở phía tay phải, lời hứa của Hàng đem lại phần thưởng cho chính mình: Cột có thể chọn II một cách an toàn, đạt được các kết quả cao hơn cho cả hai người. (Nếu, trong ma trận bên tay trái, các nước đi là lần lượt, người chơi nào chọn nước đi *thứ hai* sẽ phải có khả năng đưa ra lời hứa. Nếu những người chơi đồng ý với nhau về thứ tự các nước đi và chỉ một trong hai người có thể đưa ra lời hứa, họ có thể đồng ý rằng bên kia sẽ đi trước. Những lời hứa này, tương phản với những lời hứa của ma trận bên tay phải, phải đặt điều kiện trên lựa chọn của người chơi thứ hai. Một lời hứa đơn phương, vô điều kiện là đủ trong trường hợp ma trận phía tay phải, nhưng không có tác dụng trong trường hợp ma trận bên tay trái với các nước đi lần lượt.) Người chứng kiến một tội ác sẽ có động cơ tự hứa hẹn đơn phương khi kẻ phạm tội sẽ giết anh ta nếu tiết lộ.<sup>11</sup> Một quốc gia biết rằng mình đang

---

<sup>11</sup> Ý tưởng này được sử dụng nhiều trong tác phẩm *Thứ Bảy ẩm ướt* (Wet Saturday) của John Collier, gần đây được Alfred Hitchcock dựng lại trên truyền hình. Một người tình cờ nghe trộm một vụ giết người bị [kẻ giết người] dí súng đầu và bắt phải tự giữ im lặng bằng cách tự để lại dấu

	I	II	
i	0	2	-1
ii	-1	1	1

	I	II	
i	0	2	-1
ii	0	0	1

H. 12

ở trong tình trạng bị đe dọa bởi loại vũ khí tấn công bất ngờ có sức tàn phá lớn có thể có lý do để đơn phương từ bỏ nó - nếu có cách nào khả thi để làm như thế - để ngăn chặn từ đầu nỗi lo của kẻ thù về việc phải tấn công trước trong khi vẫn còn cơ hội.

Định nghĩa chính xác về lời hứa - chặng hạn, để phân biệt với sự đe dọa - là không rõ ràng. Có vẻ như lời hứa là một sự cam kết (có điều kiện hoặc vô điều kiện) mà bên thứ hai hoan nghênh, một sự cam kết có lợi chung cho hai phía, như một trò chơi thể hiện trên Hình 12. Nhưng Hình 13 biểu thị một tình huống trong đó Hàng phải kết hợp cả sự đe dọa và hứa hẹn; anh ta đe dọa sẽ chọn ii nếu Cột chọn I và hứa hẹn chọn i trong trường hợp Cột chọn II. Lời hứa bảo đảm cho Cột một phần lợi ích 4 thay vì 0, một khi anh ta đã chọn II, và theo nghĩa đó anh ta có lợi; với việc Hàng chịu thiệt 1 đơn vị [chọn i để lấy 4 thay vì chọn ii để lấy 5]. Nhưng, nếu Hàng không thể đưa ra lời hứa, Cột sẽ thắng 5; anh ta thắng là vì sự đe

---

vân tay của anh ta và các bằng chứng buộc tội khác, để cho nếu xác nhận bị phát hiện thì chính anh ta sẽ bị buộc tội giết người. Tuy nhiên, lẽ ra anh ta đã có thể khăng khăng đòi tạo ra các bằng chứng để chia sẻ tội ác với kẻ giết người thực sự, như đã xảy ra, anh ta đã bị lừa một cách tệ hại. (*Short Stories from the "New York"* [London, 1951], trang 171-178).

dọa không có hiệu lực nếu không có lời hứa, và sự đe dọa sẽ không xảy ra. Sự đe dọa chọn ii để chống lại I không có lợi; nó không thể buộc Cột chọn II, vì sự lựa chọn II sẽ để lại cho anh ta một kết quả tại ii,II, bằng 0 thay vì 1. Sự đe dọa của Hàng có thể có kết quả nếu có lời hứa kèm theo; hiệu quả của lời hứa là làm cho sự đe dọa có kết quả, Cột nhận được 4 thay vì 5, Hàng được 5 thay vì 2. Người ta không thể ép gián điệp, kẻ mưu đồ hay những người mắc các bệnh xã hội tự để lộ mình đơn thuần bằng lời đe dọa của sự truy quét gay gắt bất cứ lúc nào; người ta còn phải hứa miễn giảm tội cho những ai ra trình diện.<sup>12</sup>

Có lẽ một định nghĩa tốt hơn sẽ làm cho lời hứa thành sự cam kết được kiểm soát bởi bên thứ hai, nghĩa là, một sự cam kết mà bên thứ hai có thể cưỡng chế hoặc từ bỏ tùy mình lựa chọn. Nhưng ở đây thời điểm rất quan trọng. Lời hứa vừa được thảo luận sẽ có kết quả sau khi sự đe dọa được hoàn toàn cam kết; nhưng nếu nạn nhân của lời hứa (Cột) có thể bội ước lời hứa trước, sao cho Hàng hiểu rằng Cột kỳ vọng kết quả số 0 nếu anh ta chọn II, bản thân lời hứa sẽ bị ngăn chặn. Và, nếu sự đe dọa và lời hứa được tính toán trước sao cho không thể tách rời “về pháp lý” hoặc nếu chúng được thực hiện bởi một số mưu mẹo không thể đảo ngược, thì định nghĩa trở nên mơ hồ. (Trên thực tế, định nghĩa sẽ bị phá vỡ mỗi khi có được sự tương đương của một lời hứa bởi một số hành vi không thay đổi được thay vì bởi một sự cam kết “pháp lý”.)

---

<sup>12</sup> Liên quan đến vấn đề này là việc cho quyền miễn tố khiếu nại cho một nhân chứng kín tiếng không còn viện vào nỗi sợ tự buộc tội mình [nếu nói ra], và như thế đặt anh ta đứng trước hình phạt thông thường về tội xem thường tòa án [nếu không nói].

	I	II
i	2      5	4      4
ii	1      1	5      0

H. 13

Kỳ thực, mỗi khi số lựa chọn lớn hơn hai, sự đe dọa và lời hứa nhiều khả năng sẽ bị trộn lẫn trong bất kỳ "khuôn khổ phản ứng lại" nào mà một bên có thể đưa ra đối với bên kia. Do đó có lẽ điều tốt nhất là hãy coi sự đe dọa hay hứa hẹn là những cái tên đặt cho các khía cạnh khác nhau của cùng một chiến thuật tự cam kết có lựa chọn và có điều kiện, mà trong một số trường hợp đơn giản nhất định có thể được nhận diện dưới dạng lợi ích của bên thứ hai.

**Các kế hoạch cưỡng chế.** Các thỏa thuận sẽ không thể cưỡng chế được nếu không có một quyền lực bên ngoài để cưỡng chế chúng hoặc nếu việc không tuân thủ vốn dĩ không thể phát hiện được. Do đó xuất hiện vấn đề tìm ra các hình thức thỏa thuận, hay những điều kiện để thỏa thuận khiến hai bên không muốn lừa gạt, hay làm cho việc lừa gạt bị lật tẩy, hoặc có các hình phạt có thể cưỡng chế được. Mặc dù không cần loại trừ khả năng "tin cậy" giữa hai bên, nhưng sự tin cậy đó cũng không phải là điều đương nhiên có; và ngay cả bản thân sự tin cậy cũng có thể được nghiên cứu một cách hữu ích bằng các thuật ngữ lý thuyết trò chơi. Sự tin cậy thường có được bởi sự liên tục của mối quan hệ giữa các bên và sự công

nhận của mỗi bên rằng điều anh ta có thể đạt được bằng sự lừa gạt trong những tình huống nào đó không thể sánh được với giá trị truyền thống của sự tin cậy làm cơ sở cho hàng loạt thỏa thuận có thể có trong tương lai. Cũng tương tự như vậy, "sự tin cậy" cũng có thể tạo lập đối với một trường hợp đột xuất đơn lẻ nếu nó có thể được chia thành nhiều bước để tăng dần sự tin cậy.

Tuy nhiên, có những tình huống trò chơi cụ thể mà tự nó có thể có những thỏa thuận cưỡng chế được. Một là thỏa thuận phụ thuộc vào sự phối hợp hay bổ sung lẫn nhau theo một kiểu nào đó. Nếu hai người không đồng ý với nhau về địa điểm để ăn tối; nếu hai kẻ tòng phạm bắt đồng ý kiến về bằng chứng ngoại phạm chung; hoặc nếu các thành viên của một doanh nghiệp kinh doanh hoặc một đội bóng đá tranh chấp với nhau về cái giá nào họ sẽ mời chào hay chiến thuật nào họ sẽ sử dụng, thì mặc dù vậy họ vẫn có một lợi ích bao trùm đối với sự nhất quán đến cùng trong các hành động của họ. Một khi thỏa thuận chính thức đạt được, nó sẽ tạo ra một tiêu điểm khả dĩ duy nhất cho sự phối hợp ngầm cần thiết sau đó; bây giờ không một ai có sự ưu tiên đơn phương để làm bất cứ cái gì khác ngoài cái điều anh ta được kỳ vọng để làm. Trong trường hợp không có các phương tiện cưỡng chế nào khác, thì các bên cũng nên cố gắng tìm ra các thỏa thuận mang tính chất của các kỳ vọng phụ thuộc lẫn nhau, thậm chí phải nhập vào bản thỏa thuận một số yếu tố nhất định mà mục đích duy nhất của nó là tạo ra hình phạt nghiêm ngặt nếu không phối hợp. Việc xé chiếc bản đồ tìm chậu báu làm hai hay để cho một bên giữ súng và bên kia giữ đạn được là một ví dụ quen thuộc.

Việc sử dụng *con tin* là một kỹ thuật lâu đời đáng để nghiên cứu trong lý thuyết trò chơi, cũng như thông lệ uống rượu vang trong cùng một ly hay tổ chức cuộc gặp của các băng nhóm tại những nơi công khai để không bên nào có thể trốn thoát nếu một bên có ý định thảm sát. Việc chỉ dùng những người nghiện ma túy để làm đại lý hay nhân viên trong băng nhóm ma túy là một ví dụ khá trực diện về trường hợp con tin đơn phương.

Có lẽ sự trao đổi dân số qua lại giữa các nước vốn căm ghét lẫn nhau, hay một thỏa thuận di chuyển các cơ quan cai trị của cả hai nước đến một hòn đảo duy nhất nơi họ cư trú trong các khu nhà xen kẽ nhau trong thành phố có thể được dùng đến nếu cả hai bên đều cùng hết sức nỗ lực để tránh một cuộc tương tàn. Trở ngại chính đối với việc trao đổi các con tin, với giả thiết hành vi duy lý, là khả năng không thể biết vốn có của mỗi bên về hệ thống giá trị của bên kia như đã nói ở phần trên. Một ông vua gửi cô con gái của mình sang làm con tin tại cung điện của kẻ thù có thể không xua tan được nỗi lo ngại của kẻ thu rằng ông ta thực sự không thương yêu gì đứa con gái của mình. Chúng ta cũng có thể bảo đảm với người Nga rằng sẽ không có tấn công phủ đầu từ phía Mỹ bằng cách có chương trình tựa như “một năm ở nước ngoài” cho các cháu mẫu giáo: nếu tất cả các cháu học sinh Mỹ năm tuổi đều phải đến học mẫu giáo tại Nga - trong các cơ sở của Mỹ được xây dựng vì mục đích đó, sẽ được thiết kế đơn thuần cho mục đích “con tin” chứ không phải mục đích trao đổi văn hóa - và nếu số nhóm học sinh cho từng năm đến nơi trước khi lớp học sinh cũ tốt nghiệp và rời Nga, thì sẽ không mấy may có khả năng nào là nước Mỹ có thể khởi xướng tấn công nguyên tử

vào nước Nga. Chúng ta không thể biết chắc rằng người Nga sẽ tin chắc điều này. Chúng ta cũng không thể tin chắc rằng một chương trình có đi có lại như vậy sẽ là một sự ngăn chặn đối với chính phủ Nga; thật không may, ngay cả khi chính phủ Nga thực sự bị ràng buộc và không thể làm hại các trẻ em Nga, thì có vẻ họ giàn như không thể thuyết phục được chúng ta tin vào điều đó. Mặc dù vậy, trong nhiều tình huống tấn công bất ngờ một sự hứa hẹn đơn phương vẫn tốt hơn là không có; và các ý tưởng về các con tin có thể rất đáng để xem xét, ngay cả khi sự trao đổi đối xứng xem ra là không tồn tại.<sup>13</sup>

Quả thực, ý tưởng con tin là sự đồng nhất về logic với ý tưởng rằng một hiệp định giải trừ quân bị giữa các nước lớn có thể sẽ có hiệu quả nhiều hơn (và có thể chịu sự kiểm soát về kỹ thuật nhiều hơn) nếu nó được liên hệ với các vũ khí và các cơ cấu *phòng thủ*. Để tránh phòng thủ, trong thực tế, là làm cho toàn bộ dân chúng của mình thành con tin mà không cần bận tâm việc đặt bản thân họ thuộc quyền sở hữu của người khác. Như vậy chúng ta có thể phó mặc con em của chúng ta cho người Nga và có cái quyền tương tự như vậy đối với các trẻ em người Nga, không chỉ bằng việc trao đổi chính thân thể họ, với sự bất tiện lớn lao và cả sự vi phạm các quyền hiển định, mà còn bằng việc đồng ý không bảo vệ họ để người

---

<sup>13</sup> Định nghĩa chính xác về con tin có một chút khó khăn. Chúng có vẻ như thích hợp với cả sự đe dọa và hứa hẹn: Các sư đoàn Mỹ đóng quân tại châu Âu chủ yếu để chứng tỏ rằng nước Mỹ không thể tránh việc dính líu vào cuộc xung đột ở châu Âu có thể được coi là các con tin; nếu họ không phải thế, thì vợ con của họ trở thành sự cam kết có sức thuyết phục hơn hoặc “trip wire” hơn là bản thân quân đội. Như một quy luật chung, những kẻ xâm lược có thể buộc phải tránh mùa có nhiều khách du lịch tại các nước họ muốn chiếm, để tránh khiêu khích các nước bỗng dưng có những con tin bất ngờ.

khác có thể làm hại họ một cách dễ dàng như nằm trong tay anh ta. Như thế thì “sự cân bằng sức hủy diệt” mà thường được nói tới – nếu, trên thực tế, nó tồn tại và ổn định – tương đương với một sự trao đổi tổng thể của tất cả các con tin có thể được. (Sự tương đồng ở đây đòi hỏi sự cân bằng phải ổn định, nghĩa là không bên nào có thể, bằng sự tấn công bất ngờ, phá hủy sức mạnh đánh lại của bên kia mà chỉ có thể gây ra tổn thất lớn lao cho dân thường.)<sup>14</sup>

**Từ chối cưỡng chế.** Việc cưỡng chế các lời hứa cũng tựa như ảnh hưởng của bên thứ ba là bên muốn làm cho kết quả hiệu quả trở nên khó đạt đến hơn đối với hai người chơi kia. Một biện pháp mạnh để cấm các hoạt động bất hợp pháp thường là đặt chúng ngoài vòng pháp luật, sao cho các thỏa thuận [giữa những người phạm pháp] trở nên không thể cưỡng chế được. Việc không cưỡng chế các thỏa thuận đánh bạc hay các thỏa thuận hạn chế buôn bán hoặc các thỏa thuận về vận chuyển rượu trong thời kỳ bị cấm luôn luôn là một phần của quá trình ngăn cản các hoạt động này. Tất nhiên, đôi khi việc ngăn cấm theo kiểu này lại chuyển quyền lực to lớn vào tay của kẻ nào có khả năng cưỡng chế các thỏa thuận hoặc làm cho các lời hứa có thể cưỡng chế.<sup>15</sup> Việc từ chối cấp bản quyền các nhãn hiệu rượu trong thời kỳ cấm rượu có nghĩa là chỉ có

<sup>14</sup> Khái niệm này được trình bày kỹ hơn ở Chương 10.

<sup>15</sup> Người ta lập luận rằng một chức năng quan trọng của kẻ buôn gian bán lận là đôi khi giúp cưỡng chế các thỏa thuận nằm ngoài pháp luật. Bên nào chủ động giảm giá trong buôn bán quần áo tại Chicago có thể bị trừng phạt bằng các vụ nổ - thủ lao trả cho các vụ nổ này do tổ chức ấn định giá trả - theo R. L. Duffus, "Chức năng của kẻ buôn gian bán lận" (The Function of the Racketeer), *New Republic* (ngày 27-3-1929), trang 166-168.

những băng nhóm lớn hơn mới có thể bảo đảm chất lượng rượu của mình và từ đó giúp chúng phát triển sự kiểm soát độc quyền đối với ngành kinh doanh này. Tương tự như vậy, các luật pháp bảo vệ thương hiệu và nhãn hiệu có thể được xem là những công cụ để tạo thuận lợi cho kinh doanh dựa trên cơ sở các thỏa thuận không có trên giấy tờ.

## *TÙ BỎ QUYỀN CHỦ ĐỘNG*

Điều gì khiến các đe dọa hoặc cam kết bình thường trở thành một chiến thuật khó sử dụng? Một đối tượng lý thú để nghiên cứu là việc tìm ra một phương tiện đảm bảo các cam kết, một “hình phạt” sẵn có để trừng phạt việc không thực hiện của chính mình. Có một nhóm các chiến thuật có liên quan bao gồm việc dẫn mình vào một vị trí mà không còn sự lựa chọn có hiệu quả nào khác về cách hành xử hay phản ứng của mình. Mục đích của những chiến thuật này là loại bỏ quyền chủ động khó xử, làm cho kết quả chỉ phụ thuộc vào sự lựa chọn của bên kia.

Đây là một kiểu chiến thuật mà Bộ trưởng Ngoại giao John Foster Dulles tìm kiếm trong đoạn văn sau:

Trong tương lai, có lẽ chúng ta sẽ không cần phải dựa nhiều vào khả năng ngăn chặn của sức mạnh trả đũa trên quy mô lớn... Như thế, ngược với thập kỷ 1950, ở thập kỷ 1960, có thể các quốc gia ở xung quanh Trung-Xô sẽ có được sức mạnh quân sự để đánh bại một cuộc tấn công tổng lực bằng vũ khí thông thường, và như vậy có thể đối đầu với bất

kỳ kẻ xâm lược nào và đẩy [những kẻ xâm lược này] vào chỗ phải lựa chọn giữa hoặc chịu thất bại hoặc tự mình khởi xướng chiến tranh hạt nhân chống lại nước phòng thủ. Như vậy thì tình thế có thể đảo ngược theo cái nghĩa là thay vì [thông thường] những nước tự vệ phải dựa vào sức mạnh trả đũa tổng lực bằng nguyên tử để bảo vệ mình, những kẻ xâm lược tương lai sẽ không thể thực hiện một cuộc xâm lược thành công bằng vũ khí thông thường, và chính họ phải cân nhắc hậu quả của việc gây ra cuộc chiến tranh hạt nhân.<sup>16</sup>

Sự khác nhau giữa loại ngăn chặn mà ông ta gán cho những năm 1950 và loại mà ông gán cho những năm 1960 là ở chỗ ai là người phải đưa ra quyết định sau cùng; và sự khác biệt là quan trọng bởi vì Hoa Kỳ không thể tìm thấy sự đáng tin cậy từ phía đối phương, hoặc không thể làm cho mình trở nên đáng tin [trong mắt kẻ thù] - một phương tiện cam kết thuyết phục đối với sự đe dọa trả đũa ô ạt chống lại các kiểu xâm lược nhất định.

Có một thời kỳ, ngay sau khi quả bom nguyên tử đầu tiên được thả xuống, khi có một số phỏng đoán của giới báo chí về liệu bầu khí quyển trái đất có sức chịu đựng có giới hạn đối với phân hạch hạt nhân hay không; ý tưởng này đã bị lan truyền về phản ứng dây chuyền khủng khiếp có thể phá hủy

<sup>16</sup> J. F. Dulles, "Thách thức và đáp lại trong chính sách của Hoa Kỳ" (Challenges and Response in U.S. Policy), *Foreign Affairs* (tháng 10-1957). Lời văn tương tự cũng được Dean Acheson sử dụng (*Power and Diplomacy - Sức mạnh và Ngoại giao*. [Cambridge, Mass., 1958], trang 87-88) khi trao đổi về vai trò của lực lượng phòng thủ quy mô lớn tại châu Âu: bằng cách đòi hỏi kẻ thù một cuộc tấn công quy mô lớn, thay vì quy mô nhỏ, nó làm cho ông ta tin rằng sự trả đũa sẽ diễn ra, bởi vì "hắn sẽ phải quyết định thay chúng ta... Sự phòng thủ ở châu Âu với mức độ như thế này sẽ vượt qua quyết định mạo hiểm mọi cái từ phòng thủ đến phản công."

bầu khí quyển trái đất khi một số lượng lớn bom nguyên tử bị nổ. Một số người đề nghị rằng, nếu điều này là đúng và nếu chúng ta có thể tính toán với độ chính xác mức độ tối hạn của sức chịu đựng đó, chúng ta có thể trung hòa các vũ khí nguyên tử mãi mãi bằng một chương trình có sự chuẩn bị kỹ để kích nổ  $n-1$  quả bom một cách công khai.

Chiến thuật được dùng trong các trường hợp kịch tính “chống đối thu động” mà đúng hơn phải gọi là “không phản kháng tích cực.” Theo tờ *New York Times*, “Hôm nay những công nhân đường sắt đình công đã ngồi trên đường ray tại hơn 300 nhà ga của Nhật, chặn đứng 48 đoàn tàu khách và 144 tàu chở hàng.”<sup>17</sup>

Một trường hợp khác kịch tính hơn, cũng là của Nhật, được đăng trên cùng tờ báo: “Một cuộc tranh luận công khai đang được tổ chức tại đây tuần này về vấn đề liệu có nên gửi một “đoàn tàu ngồi cảm tử” đến vùng biển cấm xung quanh Đảo Christmas, địa điểm thử bom khinh khí của Anh sắp tới... Mục tiêu đầu tiên của cuộc hành trình sẽ là ngăn cản vụ nổ của Anh.”<sup>18</sup>

<sup>17</sup> “Những người đình công đường sắt ngồi trên đường ray” (Rail Strikers Sit in Tracks). *The New York Times* (ngày 13-5-1957), trang 14Lf. Đối sách thích hợp có lẽ là như sau: người lái tàu vặn van giảm ga cho tàu chạy chậm lại rồi nhảy từ trên tàu xuống rồi đi bộ qua ga, và nhảy lên tàu trở lại khi tàu vừa bắt kịp. Điểm yếu của vị trí anh ta trong khi anh ta đang lái tàu là anh ta không thể dừng tàu nhanh hơn đối thủ của anh ta có thể nhảy ra khỏi đường ray, đặc biệt nếu họ đã chuẩn bị để tụ tập sao cho họ không thể giải tán nhanh khỏi đường ray. Họ có thể chặn trước chiến thuật chống lại của anh ta bằng cách bịt kín đường ray và ném đi chìa khóa - nếu họ có thể thông báo trước điều này một cách thuyết phục với người lái tàu trước khi anh ta từ bỏ việc điều khiển chiếc tàu.

<sup>18</sup> “Cuộc tranh luận của Nhật Bản “cảm tử” bom nguyên tử” (Japan Debating Atomic ‘Suicide’), *The New York Times* (ngày 5-3-1957), trang 16.

## SỰ NHẬN DIỆN

Một đặc tính quan trọng của mọi trò chơi là mỗi bên biết được hệ thống giá trị của bên kia đến mức nào; nhưng vấn đề thông tin tương tự cũng xảy ra đối với việc chỉ thuần túy nhận diện nhau. Một nhân viên ngân hàng muốn cướp ngân hàng này nếu anh ta cần tìm ra một kẻ đồng lõa ở bên ngoài và một kẻ cướp ngân hàng muốn cướp ngân hàng nếu hắn ta có thể tìm ra một tay trong đồng lõa có thể có khó khăn trong việc hợp tác bởi vì họ không thể nhận diện lẫn nhau, ở đây có những hình phạt nghiêm khắc trong trường hợp một người nói ra ý định của mình cho người khác mà người đó tỏ ra không có cùng mối quan tâm như nhau. Một anh thanh niên ngại ngùng đặt vấn đề hẹn hò với một cô gái vì cô ta có thể từ chối anh ta cũng ở vào một vị trí giống như vậy. Tương tự, kẻ bắt cóc không thể nào hành động có hiệu quả nếu hắn ta không thể nào biết trước ai là người giàu ai là người nghèo; và thiểu số chống lại sự phân biệt chủng tộc ở miền Nam nước Mỹ có thể chưa bao giờ biết được đó là lớn hay nhỏ vì những hình phạt về sự tuyên bố đó.

Sự nhận diện, giống như thông tin liên lạc, không nhất thiết là có đi có lại; và hành động tự nhận diện đôi khi có thể đảo ngược được và đôi khi không. Một người có thể giành được nhiều sự nhận diện hơn anh ta mặc cả để có, một khi anh ta tỏ rõ sự quan tâm của mình về một đối tượng. Một ví dụ hay có thể tìm thấy trong tác phẩm *Measure for Measure* (Tương kế tựu kế) của Shakespeare. Angelo, đóng vai Công tước, có

một phạm nhân mà ông ta đề xuất án tử hình. Ông ta có thể tra tấn anh ta, nhưng ông ta không có động cơ để làm việc đó. Phạm nhân có một cô em gái, người đã đến xin cứu mạng cho anh trai. Angelo, nhận thấy cô em gái này xinh đẹp, nên đưa ra một sự mặc cả đáng hổ thẹn; cô em gái từ chối, Angelo dọa sẽ tra tấn người anh trai nếu cô ta không thuận khuất phục. Tới điểm này thì trò chơi được mở rộng, đơn giản bởi việc xác định nhận dạng và bởi kênh liên lạc. Lợi ích duy nhất của Angelo trong việc tra tấn người anh trai của cô gái là cái mà ông ta nhận được bằng sự đe dọa sẽ làm việc đó; một khi có ai đó mà có thể truyền thông tin đến người này, thì khả năng tra tấn là cái giá trị mà Angelo có - không phải là bản thân việc tra tấn, mà là sự đe dọa làm điều đó. Người em gái đã nhận được giá trị âm trong chuyến đi của mình; đã nói rõ ý muốn của mình và đã hiện diện để tiếp nhận thông điệp đe dọa, cô ta đã bị ép buộc phải chịu đựng cái điều mà lẽ ra cô không phải chịu nếu cô không bộc lộ nhận dạng của mình, hoặc nếu cô ta biến mất vào đám đông trước khi lời đe dọa được đưa ra.

Một trò chơi nhận diện lý thú khác được phát hiện tại ngoại thành New York vài năm trước đây. Một số người lái xe ô-tô mang theo thẻ hội viên để chứng minh với cảnh sát họ là hội viên của một câu lạc bộ; nếu lái xe có thẻ hội viên bị bắt giữ, họ chỉ việc đưa thẻ cho cảnh sát và đưa tiền hối lộ. Vai trò của các thẻ này là để nhận diện lái xe là người mà, nếu đưa hối lộ, sẽ giữ im lặng. Cái thẻ này nhận diện người lái xe như là một người mà lời hứa của anh ta có thể bảo đảm được. Nhưng cái thẻ nhận diện người lái xe chỉ sau khi anh ta bị bắt giữ; nếu cảnh sát có thể nhận diện những lái xe mang theo thẻ

bằng cách chỉ nhìn vào họ, họ sẽ tập trung việc bắt giữ vào những lái xe mang thẻ, dọa sẽ phạt nếu không đưa tiền hối lộ. Chiếc thẻ là sự nhận diện bất ngờ, tùy thuộc vào người lái xe. Một tình huống tương tự - phù hợp với các thảo luận về lời hứa cũng như sự nhận diện đã được Sutherland mô tả: "Hầu hết cảnh sát ít nhiều công bằng khi xử lý với những kẻ cắp đơn giản chỉ vì họ có lợi để làm việc đó. Họ sẽ đưa ra sự chiểu có thâm chí vì một khoản rất nhỏ mà họ sẽ không dành cho dân không chuyên nghiệp những người mà họ sẽ tống giam. Họ nhận ra rằng làm điều này là an toàn và rằng các quan chức cấp cao sẽ không biết được, khác với trường hợp nếu họ chiểu có dân nghiệp dư."<sup>19</sup>

Sự nhận diện cũng có liên quan đến sự kiện kinh tế quan trọng mà thường bị bỏ qua trong kinh tế học đương thời về sản xuất và trao đổi hàng hóa, chẳng hạn, tiềm năng to lớn của sự tàn phá đang tồn tại và đang diễn ra vì các mối đe dọa tổng tiền có thể lấy nó làm chỗ dựa. Một học sinh khỏe mạnh bình thường, tốt nghiệp trung học với học lực dưới mức trung bình một chút, phải làm việc khá vất vả để tạo ra giá trị hơn 3.000 đô-la hoặc 4.000 đô-la trong một năm; nhưng anh ta có thể phá hoại lớn hơn đó hàng trăm lần nếu anh ta quyết tâm làm chuyện đó, theo tính toán sơ bộ của tác giả. Với một sự dàn xếp mang tính thể chế trong đó anh ta tránh không phá hoại để đổi lại chỉ một phần giá trị mà anh ta có thể phá hoại, thì cậu học sinh này rõ ràng có một nghề nghiệp như một kẻ tống tiền thay vì một người thợ cơ khí hay một thư ký văn

---

<sup>19</sup> E. H. Sutherland, *The Professional* (Kẻ cắp chuyên nghiệp) (Chicago, 1954), trang 126.

phòng. May mắn là việc tống tiền thường phụ thuộc vào sự tự lộ diện và liên lạc công khai bởi bản thân kẻ tống tiền.

Tầm quan trọng của sự tự nhận diện được chứng thực bởi ý nghĩa của học thuyết cho rằng một người bị tố cáo nên được phép biết và đối mặt với kẻ tố cáo. Nó cũng được phản ánh trong cuộc điều trần bí mật trước bồi thẩm đoàn, trong những trường hợp mà các nhân chứng được nhận diện có thể bị đe dọa bởi các bị cáo tiềm năng và trong các nỗ lực nhằm giữ bí mật nhân dạng của các nhân chứng cho một vụ án cho đến khi kẻ phạm tội bị bắt. (Chiến lược của luật pháp, thực thi pháp luật và ngăn chặn tội phạm là một lĩnh vực phong phú cho việc áp dụng lý thuyết trò chơi.)

## *TRAO QUYỀN*

Một “nước đi” khác đôi khi xuất hiện là trao gửi một phần hoặc toàn bộ lợi ích của một người, hay một phần hoặc toàn bộ quyền chủ động ra quyết định, cho một kẻ đại diện, người sẽ trở thành (hoặc có thể đã là) một người tham gia trò chơi. Các hợp đồng bảo hiểm cho phép việc chia sẻ các lợi ích; công ty bảo hiểm có động cơ hành động khác hẳn so với động cơ của người được bảo hiểm và có thể đưa ra sự đe dọa hay chống lại vì lý do này. Việc yêu cầu nhiều chữ ký trên một tờ séc cũng nhằm phục vụ mục đích tương tự. Việc công ty sử dụng một đại lý chuyên đi đòi nợ là một biện pháp để đạt được mối liên lạc đơn phương với bên nợ thay vì song phương

và do đó là biện pháp để tránh phải nghe phản nàn hay đe dọa từ phía người mắc nợ. Việc cung cấp vũ khí đạn dược cho quân đội Nam Triều Tiên hoặc cho họ tiếp cận với các trại tù nhân chiến tranh để họ đơn phương thả tù binh là một biện pháp chiến thuật để tránh sử dụng quyền lực ra quyết định khó xử - khó xử bởi vì nó đòi hỏi một bên phải có đe dọa cưỡng ép hoặc răn đe hoặc tước bỏ khả năng rút lui sự đe dọa của mình, từ đó không có khả năng làm cho sự đe dọa có sức thuyết phục.

Hiệp định phòng thủ chung với Đài Loan, có thể được xem một phần như là biện pháp để chuyển quyết định phản ứng cho một người khác mà quyết tâm của họ rõ rệt hơn; và đề nghị gần đây hơn về việc trao các vũ khí hạt nhân vào tay chính phủ các nước châu Âu đã được tranh cãi công khai dựa trên luận điểm cho rằng nó sẽ tăng cường khả năng răn đe bằng cách trao sức mạnh trả đũa cho các nước mà trong những tình huống phải hành động, các nước đó được coi là ít lưỡng lự hơn so với Hoa Kỳ.

Việc sử dụng những tên ác ôn và những kẻ tàn bạo vào việc đi thu tiền từ các vụ tống tiền hoặc canh giữ các tù nhân, hoặc việc trao quyền lực cho một chỉ huy quân sự đã biết rõ động cơ là những minh họa cho một phương pháp chung để làm tăng độ tin cậy cho một kiểu phản ứng mà người đưa ra quyết định ban đầu được coi là lưỡng lự hoặc thấy không bõ một khi lời đe dọa bị thất bại. (Cũng giống như hành vi vẫn được coi là duy lý khi một người chơi duy lý hủy bỏ tính duy lý của mình trong những tình huống trò chơi nhất định, hoặc để ngăn chặn một mối đe dọa bất lợi cho mình khi mối đe dọa này chỉ tồn tại dựa trên tính duy lý của anh ta, hoặc làm cho

lời hăm dọa mà anh ta vốn không thể làm [nếu anh ta vẫn còn là một người duy lý] trở nên đáng tin cậy, thì việc người chơi lựa chọn các đối tác hay đại diện không duy lý cho mình cũng được coi là một hành vi duy lý.)

Trong ma trận trên, Hình 14 – bỏ qua con số ở trong ngoặc – nếu Hàng có nước đi thứ hai, anh ta sẽ thua ở góc dưới bên tay phải, còn Cột giành được kết quả như mong muốn. Nếu một bên thứ ba không có quyền quyết định được chọn ra để tiếp nhận, như sản phẩm phụ, phần thưởng nằm trong ngoặc, thì Hàng có thể thắng cuộc nếu có sẵn một số cách để chuyển giao không đảo ngược được nước đi của mình cho bên thứ ba. Phần thưởng cho bên thứ ba là phần thưởng anh ta giành được ở góc trên tay trái bằng nước đi thứ hai, để lại cái đầu tiên cho Hàng. Phần thưởng cho người thứ ba là phần thưởng bằng nước đi thứ hai được tài trợ bởi Hàng, còn phần thưởng của Hàng bị giảm đi tương ứng, nhưng nó vẫn bõ công cho anh ta để trao hẳn quyền đổi với một phần trong số nhiều phần thưởng của mình cho bên thứ ba, cùng với trao quyền quyết định; như hình vẽ cho thấy, anh ta vẫn nhận được giá trị tổng cộng là 3 [được 5 nhưng phải trả công cho người đại diện 2] ở góc phía trên tay trái, tương phản với giá trị bằng 1 ở dưới bên phải.

	I	II
i	3 (2)	0 (1)
ii	0 (1)	1 (0)

H. 14

## *NGƯỜI TRUNG GIAN*

Vai trò của người làm trung gian là một yếu tố khác để phân tích trong lý thuyết trò chơi. Một người làm trung gian, bất kể được áp đặt vào trò chơi bởi luật chơi ban đầu hay được những người chơi chấp thuận để tạo thuận lợi cho một kết quả hiệu nghiệm, có thể được xem là một yếu tố trong cơ cấu thông tin của trò chơi, hoặc như một bên thứ ba với cơ cấu lợi ích riêng và với sức ảnh hưởng có được thông qua việc anh ta nắm quyền kiểm soát thông tin liên lạc [giữa những người chơi còn lại]. Nhưng người làm trung gian còn có thể làm được nhiều hơn là chỉ kiểm soát thông tin qua lại - đặt ra giới hạn cho thứ tự các đề nghị, các phản đề nghị, v.v... - do anh ta có thể tạo ra các dữ liệu theo hoàn cảnh và đưa ra các đề xuất có sức mạnh. Nghĩa là, anh ta có thể gây ảnh hưởng đến kỳ vọng của các bên bằng sáng kiến của mình, theo một cách mà cả hai bên không thể nào không nhận được. Khi không có tiêu điểm rõ ràng cho thỏa thuận, anh ta có thể dùng sức mạnh của mình để tạo ra tiêu điểm bằng đề xuất hấp dẫn. Một người đi đường bắt chợt nhảy ra giữa một giao lộ để chỉ dẫn giao thông đang tắc nghẽn sẽ có được quyền phân biệt đối xử với từng chiếc xe bằng cách đưa ra các chỉ dẫn giao thông hiệu quả tới mức ngay cả những xe ít được ưu tiên nhất cũng vẫn được lợi; sự điều phối của anh ta chỉ có sức mạnh chỉ dẫn, nhưng sự phối hợp [giữa những người lái xe] đòi hỏi phải có sự chấp nhận chung [giữa những người lái xe này về] một sự chỉ dẫn nào đó. Tương tự, những người

tham gia cuộc khiêu vũ theo đội hình có thể không thích điệu nhảy nào đó do người đê xướng hô lên, nhưng chừng nào người đê xướng cầm micrô trong tay, thì không ai có thể nhảy theo điệu khác. Đường vạch trắng phân cách chạy dọc theo tâm con đường cũng như người trung gian, và có rất nhiều khả năng nó sẽ thiên về bên này hay bên kia một cách đáng kể trước khi bên bắt lợi tim được lợi thế do từ chối thẩm quyền của nó. Nguyên tắc nay được minh họa rất đúng bởi cuộc tranh cãi về tiết kiệm thời gian ban ngày; một đa số muốn kéo thời gian làm việc lên sớm hơn một giờ sẽ không thể nào thực hiện được việc đó nếu họ không giành được quyền kiểm soát việc quy định giờ theo pháp luật. Và khi có được quyền đó, thì dù một thiểu số được tổ chức tốt để chống lại việc thay đổi giờ vẫn thường không có khả năng làm cân bằng sự thay đổi giờ này bằng những nỗ lực có tổ chức để thay đổi giờ giấc họ ấn định theo đó họ ngủ dậy, ăn sáng và đi làm.

Những người làm trung gian có thể là phương tiện làm chỗ dựa để những người chơi có lý trí có thể bỏ qua một bên một số năng lực lý trí của họ. Người làm trung gian có thể hoàn thiện một số năng lực liên lạc nhất định trong khi ngăn cản một số năng lực ghi nhớ. (Về khía cạnh này anh ta thực hiện chức năng ghi chép lại mà một máy tính có thể làm.) Chẳng hạn anh ta có thể so sánh hai đề nghị do hai bên đưa ra, tuyên bố các đề nghị đó có phù hợp với nhau hay không mà không tiết lộ nội dung cụ thể của các đề nghị. Anh ta giống như một công cụ dò tìm có thể rà soát một phần thông tin được chuyển đến. Anh ta có thể tiến hành những so sánh trong một giới hạn nhất định, những điều vượt quá khả năng giới hạn nhất định, những điều vượt quá khả năng trí tuệ của

các bên tham gia, vì không bên nào có thể cam kết mình sẽ quên được điều gì đó.

Vấn đề từ chối một cách thuyết phục việc tiếp nhận kiến thức của một người bằng tay trái, trong khi cố gắng tìm kiếm kiến thức đó bằng tay phải, được minh họa rất hay bởi những cố gắng của một số cơ quan chính phủ để có được số liệu chính xác về thu nhập nhằm mục đích cho các chương trình thống kê, trong khi một số cơ quan khác của chính phủ lại cố gắng tìm kiếm số liệu giống như vậy nhằm mục đích áp đặt thuế hay trừng phạt người trốn thuế. Các chính phủ nhận thấy tầm quan trọng trong việc tìm ra cách thức bảo đảm rằng cơ quan thống kê sẽ từ chối cung cấp thông tin họ nhận được cho cơ quan thuế để có được thông tin này trước. Một trường hợp tương tự của việc dựa vào trung gian công khai là trường hợp của các công ty đưa các bí mật thương mại cho cơ quan thống kê mà cam kết hủy bỏ các số liệu riêng biệt sau khi tính toán tổng số và trung bình mà sẽ được công bố vì lợi ích của các công ty có đóng góp, hoặc các tổ chức thăm dò dư luận là những nơi giữ kín dữ liệu cá nhân về chính trị hay quan hệ tình dục có khả năng gây khó xử, chỉ công bố những số liệu tổng hợp. Việc sử dụng các nhà trung gian để chặn trước việc nhận diện xem ra là một chiến thuật quen thuộc khi một người mua hàng với giá trị lớn cho rằng một bức tranh hay quyền đi ngang qua có thể mua được với giá rẻ nếu người chủ không biết ai là người quan tâm mua.

Những người làm trung gian có thể biến thành trọng tài nếu những người tham gia trò chơi trao quyền lực dứt khoát cho họ. Nhưng các thỏa thuận phân xử phải được làm cho có thể cưỡng chế được bằng việc các bên chịu thiệt hại có thể xảy

ra, trao cho trọng tài quyền lực để trừng phạt, hay giao lại cho ông ta một phần trong hệ thống giá trị của mỗi bên. Ngược lại, ông ta phải đáng được tin cậy và có sự hứa hẹn có thể cưỡng chế được. Nhưng trong mọi trường hợp trọng tài sẽ làm tăng số lượng các biện pháp đảm bảo thực thi các lời hứa: hai người không tin cậy lẫn nhau có thể tìm một người thứ ba mà cả hai đều tin cậy, và để cho người đó giữ tiền cược.<sup>20</sup>

## *THÔNG TIN LIÊN LẠC VÀ SỰ PHÁ HỦY NÓ*

Nhiều chiến thuật trò chơi lý thú và các tình huống trò chơi phụ thuộc vào cơ cấu của thông tin liên lạc, đặc biệt sự bất đối xứng trong thông tin liên lạc và những sự lựa chọn đơn phương để khởi xướng thông tin liên lạc hoặc để phá hủy nó. Các đe dọa chẳng có ích gì nếu chúng không được thông tin tới những người mà chúng được nhắm tới. Sự tổng tiền đòi hỏi một phương tiện để chuyển các giải pháp lựa chọn đến nạn nhân. Ngay cả lời dọa, "Nín đi đừng khóc nữa, nếu không ta sẽ cho mi vài thứ để khóc," sẽ không có hiệu quả nếu đứa trẻ đã kêu gào quá to không chịu nổi rồi. (Đôi khi hình như những đứa trẻ cũng biết được điều này.) Một nhân chứng

---

<sup>20</sup> Tác giả nghe nói rằng ở các nước nơi không có truyền thống mạnh mẽ về đạo đức kinh doanh tồn tại, thì có một số ít đối tác hay giám đốc cho doanh nghiệp có thể được có ý lựa chọn từ các nền văn hóa khác nơi mà sự trung thực và công bằng giản dị được coi là đặc điểm chung hoặc nơi mà uy tín đối với họ được coi là có giá trị cao hơn nhiều.

không thể bị đe dọa để làm chứng gian nếu anh ta bị giữ cách ly ngăn anh ta có được chỉ dẫn phải nói cái gì, mặc dù anh ta có thể suy ra mối đe dọa từ việc bị cách ly.

Khi mà kết cục phụ thuộc vào sự phối hợp, việc phá hủy đúng lúc thông tin liên lạc có thể là một chiến thuật thành công. Khi hai vợ chồng tranh luận nhau qua điện thoại nên đi ăn tối ở đâu, thì bà vợ sẽ thắng nếu bà ta chỉ cần tuyên bố nơi bà ta sẽ đến và gác máy. Và tình thế nguyên trạng thường được bảo vệ bởi một người trốn tránh bàn luận về các phương án khác, thậm chí đến mức đơn giản là tắt máy trợ thính của mình.

Như đã nói ở phần đầu của chương này, hành động đám đông thường phụ thuộc vào thông tin liên lạc theo một cách mà các nhà cầm quyền có thể ngăn chặn hành động đám đông bằng cách ngăn cấm tụ tập từ ba người trở lên. Nhưng bản thân đám đông có thể hăm dọa nhà cầm quyền nếu họ có thể nhận diện ra các quan chức và liên hệ với họ. Ngay cả mối đe dọa ngầm bằng sự tẩy chay hoặc bạo lực cũng có thể được truyền đi từ đám đông bạo loạn đến cảnh sát địa phương, nếu họ biết những cảnh sát đó và là những người sẽ vẫn sống chung với họ sau khi sự kiện kết thúc. Trong trường hợp đó việc sử dụng những người ngoài có thể chặn trước hành động đe dọa chống lại chính quyền của đám đông, một phần do làm giảm cơ hội để tiến hành đe dọa sau này, nhưng một phần cũng do khó khăn của trao đổi ngầm giữa đám đông và cảnh sát. Các lực lượng cảnh sát Liên bang tại Little Rock có thể đã được miễn trừ đe dọa chỉ vì nằm bên ngoài cơ cấu trao đổi ngầm của dân chúng địa phương và rõ ràng ít quen thuộc với hệ thống giá trị của địa phương so với cảnh sát tại chỗ.

Lực lượng cảnh sát quốc gia đã rất thành công trong việc dập tắt cuộc bạo loạn chủng tộc ở Detroit năm 1943, khi cảnh sát địa phương bất lực. Việc sử dụng lực lượng người Moor, người Sikh, và người nói tiếng nước ngoài khác chống lại cuộc nổi dậy ở địa phương thành công một phần là do họ không có khả năng tiếp nhận thông tin về sự đe dọa và hứa hẹn mà kẻ thù hoặc nạn nhân có thể tìm cách truyền đi. Ngay cả việc cách ly các sĩ quan khỏi hàng ngũ các tân binh trong quân đội cũng có thể làm cho các sĩ quan kém khả năng tiếp nhận và nhận thức được các đe dọa, do đó kém khả năng chịu sự đe dọa và từ đó ngăn chặn được bản thân lời đe dọa.

Tất nhiên điều quan trọng là người đưa ra đe dọa có biết được sự đe dọa của mình không thể được nhận hay không. Vì nếu anh ta nghĩ nó có thể trong khi nó không thể, anh ta sẽ đưa ra đe dọa và không đạt được mục đích của mình, bị buộc phải thực hiện sự đe dọa với sự bất lợi về sau cho cả người đe dọa lẫn người bị đe dọa. Do đó những binh lính trong việc dẹp bạo loạn không chỉ là những người lầm mờ và không chỉ liên tục di chuyển để tránh “sự quen biết” với một số bộ phận cụ thể của đám đông. Họ còn hành xử một cách thụ động để chứng tỏ rằng không có thông điệp nào được chuyển đến. Họ không được nhìn vào mắt ai, không được đỏ mắt lên khi bị chế giễu, và họ phải hành động như thể là họ không phân biệt được giữa những người bạo loạn, mặc dù có ai đó đã tự làm mình nổi bật. Nói theo nghĩa bóng, nếu không phải nghĩa đen, họ cần đeo mặt nạ, thậm chí bộ đồng phục cũng góp phần giữ kín tung tích và bản thân điều đó gây khó khăn cho việc trao đổi thông tin qua lại.

**Việc chuyển tải bằng chứng.** “Trao đổi thông tin” không chỉ nói đến việc truyền đi các thông điệp. Để truyền đi một sự đe dọa, người ta còn phải truyền đi cùng với nó sự cam kết, và lời hứa cũng tương tự như vậy. Và để truyền đi sự cam kết đòi hỏi nhiều thứ hơn là chỉ truyền đi lời nói. Người ta phải truyền đi *bằng chứng* về sự tồn tại của cam kết. Điều này có nghĩa rằng một người chỉ có thể truyền đi sự đe dọa nếu anh ta có thể làm cho người khác nhìn thấy tận mắt điều gì đó hoặc nếu anh ta có thể tìm ra một công cụ để xác nhận những viện dẫn nào đó. Người ta có thể gửi đi một tờ séc có chữ ký bằng bưu điện, nhưng không thể chứng minh qua điện thoại rằng tờ séc có chữ ký xác thực. Người ta có thể nói rằng khẩu súng trong tay anh ta đã lên đạn nhưng không thể chứng minh được điều đó chỉ bằng lời nói. Theo quan điểm của lý thuyết trò chơi, thì hệ thống vận hành bằng khí nén Paris khác với hệ thống điện báo thông thường, và máy truyền hình khác với radio. (Một vai trò của người trung gian có thể là để chứng minh tính xác thực của các tuyên bố mà những người chơi đưa ra cho nhau. Chẳng hạn, hệ thống mã số để nhận diện có thể giúp cho người ta chuyển tiền bằng miệng qua điện thoại, người nhận tiền được bảo đảm bởi mã số phản hồi của ngân hàng rằng thực tế ngân hàng ở đầu dây bên kia bảo đảm anh ta rằng người gửi tiền đã được nhận diện bởi mã số và việc giao dịch đã được hoàn thành.) Tâm quan trọng và sự khó khăn của sự trao đổi bằng chứng được minh họa bởi đề nghị “bầu trời mở” của Tổng thống Eisenhower và các thiết bị được gợi ý khác để giải quyết tình hình bất ổn gây ra bởi nỗi lo sợ của hai bên về tấn công bất ngờ. Leo Szilard thậm chí đã chỉ ra nghịch lý trong việc người ta muốn cho các giàn điệp

nước ngoài quyền miễn trừ thay vì bị truy tố, vì họ có thể là phương tiện duy nhất mà qua đó kẻ thù nhận được bằng chứng thuyết phục về sự thật quan trọng rằng chúng ta không có kế hoạch chuẩn bị cho việc khởi xướng một cuộc tấn công bất ngờ.<sup>21</sup>

Thật lý thú khi thấy rằng bản thân nền dân chủ chính trị cũng phụ thuộc vào cơ cấu trò chơi trong đó sự trao đổi bằng chứng là không thể. Việc bỏ phiếu bí mật là gì nếu không phải là một công cụ để cướp đi quyền của cử tri bán lá phiếu của mình? Đây không phải chỉ là vấn đề bí mật mà là bí mật *bắt buộc*, tước đi quyền của cử tri. Anh ta không chỉ có thể bỏ phiếu bí mật, mà bị buộc phải làm như vậy để cho hệ thống vận hành. Anh ta phải bị từ chối mọi phương thức để chứng minh mình đã bỏ phiếu theo hướng nào. Và cái mà anh ta bị cướp đi không phải chỉ là một tài sản mà anh ta có thể đem bán. Anh ta còn bị tước đi khả năng bị đe dọa. Anh ta bị làm cho bất lực không đáp ứng được đòi hỏi của vụ tổng tiền. Sẽ không có giới hạn đối với bạo lực mà anh ta có thể bị đe dọa nếu anh ta thực sự tự do mặc cả với lá phiếu của mình, vì dù sao bạo lực đe dọa đó sẽ không được thực hiện nếu nó gây ra sự khiếp hãi đủ để thuyết phục anh ta. Nhưng khi cử tri bị bắt lực không thể chứng minh rằng anh ta chịu khuất phục sự đe dọa, thì cả anh ta và những người đe dọa anh ta đều biết rằng mọi sự trừng phạt sẽ không có liên quan gì đến cách anh ta đã thực sự bỏ phiếu. Và sự đe dọa, vì vô ích, sẽ không được sử dụng.

---

<sup>21</sup> L. Szilard, "Giải Trữ Quân Bị và Vấn Đề Hòa Bình" (Disarmament and the Problem of Peace). *Bulletin of the Atomic Scientists*, 2:297-307 (tháng 10-1955).

Một trường hợp trao đổi ngầm và bất đối xứng lý thú là trường hợp một lái xe ô-tô tại một giao điểm giao thông đông người biết rằng một cảnh sát đang chỉ dẫn giao thông. Nếu người lái xe nhìn thấy, và rõ ràng là nhìn thấy, sự chỉ dẫn của cảnh sát nhưng bỏ qua, anh ta sẽ là người không chấp hành luật, và người cảnh sát sẽ vừa có lý do và nghĩa vụ để cho anh ta một vé phạt. Nếu người lái xe tránh không nhìn vào cảnh sát, không nhìn thấy sự chỉ dẫn, và bỏ qua sự chỉ dẫn do anh ta không trông thấy, và giành quyền được ưu tiên mà anh ta không đáng có, thì anh ta sẽ chỉ bị anh cảnh sát, người có ít động cơ và không có nghĩa vụ để phạt anh ta, coi là ngu ngốc. Ngược lại, nếu thực tế người lái xe biết sự chỉ dẫn nhưng không tuân thủ, thì cảnh sát sẽ có lợi đã không nhìn thấy lái xe, nếu không anh ta buộc phải, vì uy tín của ngành, dừng công việc bận rộn đang làm và chặn lái xe kia để cho một vé phạt. Trẻ con rất giỏi trong việc lảng tránh cái nhìn nhở của cha mẹ, hiểu rằng nếu chúng nhận thấy thì cha mẹ buộc phải trừng phạt chúng về tội không nghe lời. Người lớn cũng giỏi không kém trong việc không xin phép mà họ ngờ rằng sẽ bị từ chối, hiểu rằng sự từ chối công khai sẽ là hình phạt nghiêm khắc hơn, bắt buộc người có thẩm quyền từ chối phải nhận biết sự vi phạm.<sup>22</sup>

---

<sup>22</sup> Cái mà có thể được gọi là "địa vị pháp lý" của trao đổi thông tin đã được Goffman xây dựng rất hay: "Cách xử trí khéo léo để giữ thể diện thường có tác dụng khi dựa trên thỏa thuận ngầm thông qua ngôn ngữ gợi ý - lời nói bóng gió, úp mở, các điểm dừng đúng lúc, những chuyện vui được lựa từ ngữ cẩn thận, v.v... Quy tắc của lối giao tiếp không chính thức này là người đưa ra thông điệp không nên hành động như thể là anh ta chính thức truyền đi thông điệp mà anh ta gợi ý, trong khi người tiếp nhận có quyền và nghĩa vụ để hành động như thể là họ không chính thức nhận được thông điệp chứa đựng trong sự gợi ý đó.

Tính hiệu quả của cơ cấu trao đổi thông tin phụ thuộc vào dạng thức của tính hợp lý quy cho người chơi. Điều này được minh họa bởi tình huống trò chơi gọi là “bắt con gấu bằng đuôi.” Yêu cầu tối thiểu cho một kết cục hiệu quả là con gấu có khả năng đưa ra lời hứa có thể cưỡng chế được và có bằng chứng đáng tin cậy chứng minh sự cam kết của nó, hoặc bởi hình phạt được áp dụng hoặc bởi một hành động phá hủy khả năng không tuân thủ của nó (chẳng hạn tự rút bỏ nanh vuốt của mình). Nhưng nếu con gấu có tính hợp lý hạn chế, có khả năng đưa ra những lựa chọn hợp lý và nhất quán trong số các phương án mà nó nhận biết nhưng không có năng lực để giải các trò chơi – tức là, thiếu khả năng để tự xác định một cách nội tâm sự lựa chọn mà một đối tác sẽ đưa ra – hệ thống thông tin phải giúp nó nhận được thông điệp từ đối tác của nó. Khi đó đối tác cần phải xác lập đề xuất (sự lựa chọn) cho con gấu và truyền thông tin đến cho nó, để sau đó con gấu có thể phản hồi bằng việc chấp nhận lời hứa (lúc này nó thấy rõ “đáp số” là gì) và truyền bằng chứng có căn cứ trở lại cho đối tác của nó.

---

Sự trao đổi gợi ý, do vậy, là sự trao đổi có thể từ chối được.” Ông ta đề cập đến sự tham gia “không được phê chuẩn” mà có thể xuất hiện trong sự tương tác bằng lời nói: “Một người có thể nghe lỏm người khác mà mình không quen biết; anh ta có thể nghe lỏm họ khi họ biết điều đó và khi họ lựa chọn hoặc coi như không biết anh ta đang nghe lỏm hoặc ra hiệu không chính thức rằng họ biết anh ta đang nghe lỏm họ.” Ông ta chỉ ra rằng sự bắt buộc phải đáp lại, chẳng hạn, đối với một nhận xét xúc phạm mà người đó đã vô tình nghe lỏm được sẽ phụ thuộc vào việc nghe lỏm đã có được “sự phê chuẩn” hay không (trang 224, 226).

## KẾT HỢP CÁC NƯỚC ĐI TRONG MỘT MA TRẬN TRÒ CHƠI

Người ta cho rằng, nếu một trò chơi có các nước đi tiềm năng như đe dọa, cam kết, và hứa hẹn là những cái có thể dùng trong phân tích chuẩn, thì có thể thể hiện các nước đi như vậy bằng hình thức truyền thống của các lựa chọn chiến lược, với ma trận lợi ích của trò chơi gốc được mở rộng để cho phép có những lựa chọn trong số những nước đi này.

Điểm đầu tiên cần nhận thấy là sự cam kết, sự hứa hẹn, hoặc sự đe dọa thường có thể được đặc trưng bằng một hình thức tương đương như sau: để đi một trong số những nước đi này, người chơi cắt giảm có lựa chọn - một cách có thể trông thấy và không thể đảo ngược được - *một số* lợi ích của mình trong ma trận. Đây chính là điều mà nước đi nhắm tới.<sup>23</sup> Chúng ta cũng có thể nói rằng người ta công khai lựa chọn một chiến lược từ trước để đối phó lại sự lựa chọn của người khác. Nhưng nó không chỉ đòi hỏi sự lựa chọn. Người chơi còn phải chịu hình phạt do không theo đuổi chiến lược phản ứng cụ thể về sau mà anh ta đã lựa chọn từ trước. Và để chịu hình phạt do không đi theo một chiến lược là sự tương đương toán học với việc trừ đi một khoản hình phạt

---

<sup>23</sup> Daniel Ellsberg, một số công trình của ông trong lĩnh vực chiến lược chứa trong các bài giảng được nêu ở Chương 1, một cách độc lập, sử dụng đúng công thức xác lập sự đe dọa hay cam kết này, tức là, sự cắt giảm có chọn lọc một số lợi ích của ai đó trong ma trận chiến lược.

từ phần thưởng của mình trong tất cả các ô mà không tương ứng với chiến lược đã chọn.<sup>24</sup>

Cụ thể, ở Hình 15 A, Hàng sẽ tự cam kết đối với ii bằng cách lấy các phần thưởng của mình trong hàng thứ nhất trừ đi một khoản đủ lớn – 5 trong ví dụ được nêu – để làm cho ii trở thành chiến lược trội, tức là chiến lược mà anh ta sẽ đi theo bất kể người chơi kia chọn cột nào. Kết quả sẽ là ma trận sửa đổi được hiển thị trên Hình 15 B. (Cam kết bản thân đối với i với hình phạt bằng 5 sẽ có ma trận trên Hình 15 C.) Nay giờ liệu chúng ta có thể thiết lập một ma trận lớn hơn đại diện không chỉ cho những lựa chọn thực sự của các hàng

<sup>24</sup> Các đe dọa, hứa hẹn, và cam kết vô điều kiện đã được minh họa; một "hàm phản ứng" chung hơn được minh họa trong ma trận kèm theo. Nếu Hàng có thể gắn các hình phạt thích hợp vào sự lựa chọn của mình tại mọi ô ngoài các ô có dấu sao, anh ta sẽ dành cho Cột một bài toán tối đa hóa đơn giản mà Cột có thể giải bằng cách lựa chọn chiến lược thứ ba của mình. Hàng đã "thắng" hầu hết ô ưu tiên của mình. Cụ thể là, anh ta đã bảo đảm cho mình một ô có lợi nhất trong số các ô để lại cho Cột phần thưởng không thấp hơn giá trị "tối thiểu hóa tối đa" của anh ta. Đây là sự khai quát hơn chiến thuật mà, đối với các lựa chọn đơn giản hai chiều hay ba chiều, có thể được nhận diện là "sự cam kết," "đe dọa," "hứa hẹn," hay sự kết hợp của chúng. (Sự khai quát hóa hơn nữa có thể bao gồm các chiến lược ngẫu nhiên; vấn đề này được trình bày ở Chương 7.)

	I	II	III	IV	V
i	6	10	2	9	7
ii	1	11	10	2	10
iii	9	12	4	0	1
iv	20	15	6	1	17
v	8	2	16	18	14
vi	*	2	10	7	4
vii	5	*	8	7	20
viii	6	1	10	5	*

	I	A	II	
i	2	5	1	0
ii	0	1	5	2

	I	B	II	
i	-3	5	-4	0
ii	0	1	5	2

	I	C	II	
i	2	5	1	0
ii	-5	1	0	2

H. 15

và cột trong trò chơi gõc, như ở trên Hình 15 A, mà còn là các chiến lược *cam kết*, *đe dọa*, *hứa hẹn*, v.v... hay không? Chắc chắn là như vậy, một khi chúng ta đã xác định nước đi nào có sẵn và thứ tự trong đó chúng được thực hiện. Thủ lấy trò chơi đơn giản trong đó Hàng có khả năng cam kết rõ ràng từ trước, và Cột có nước đi đầu tiên trong trò chơi gõc, tức là chọn cột của mình trước khi Hàng đưa ra lựa chọn cuối cùng hàng của mình.

Ngay từ đầu Hàng, là người đi thứ hai, đã có bốn chiến lược có sẵn. Anh ta lấy i bất kể điều gì xảy ra; anh ta có thể lấy ii bất chấp điều gì xảy ra; anh ta có thể chơi i đối với cột I và ii đối với cột II; hoặc anh ta có thể chơi ii đối với cột I và i đối với cột II. Tính đến khả năng cam kết, bây giờ anh ta có sự lựa chọn cam kết *đầu tiên* cho mình, và đối với từng sự lựa chọn đầu tiên này anh ta có thể gắn bất kỳ chiến lược nào trong bốn chiến lược như đã nhắc đến đối với nước đi sau cùng của anh ta. Ví dụ, anh ta có thể tự cam kết ii và chơi ii bất kể điều gì xảy ra; anh ta có thể tự cam kết ii và chơi i bất kể điều gì xảy ra; anh ta có thể cam kết ii và chơi i đối với cột I, ii đối với cột II; hoặc anh ta có thể cam kết ii và chơi ii đối

với cột I, i đối với cột II. Tính cộng lại, anh ta có 12 tổ hợp chiến lược có thể xảy ra.

Cột có tám tổ hợp chiến lược có thể xảy ra: đối với mỗi một trong ba tình huống dự phòng anh ta có một trong hai nước đi, các nước đi là I và II, các tình huống dự phòng là sự cam kết của Hàng đối với i, sự cam kết của Hàng đối với ii, và sự không cam kết của Hàng.

Nếu chúng ta đặt các chiến lược này vào hình thức ma trận, chúng ta sẽ có Hình 16. Ma trận  $12 \times 8$  của Hình 16 đại diện cho trò chơi ngầm ("không hợp tác") mà tương ứng với các quyết định cá nhân của người chơi về *chơi cách nào* đối với trò chơi gốc. Tám chiến lược có thể xảy ra có sẵn đối với Cột, chẳng hạn, có thể được coi như tám *tập hợp các chỉ dẫn đầy đủ* mà anh ta có thể trao cho một đại diện là người sẽ chơi trò chơi gốc cho anh ta - tức là, chơi trò chơi mà anh ta chọn một trong hai cột, tùy thuộc vào hiện Hàng có cam kết trước hay không và bằng cách nào. Sẽ không có mắt mát gì cho các bên khi họ phải chơi trò chơi mở rộng một cách ngầm, vì cái mà đáng lẽ là *sự thích nghi* của một người chơi đối với các nước đi trước của người khác thì bây giờ hoàn toàn được đưa vào việc xác lập các chiến lược cụ thể trong phiên bản mở rộng của trò chơi. Chúng là những chiến lược phản ứng hoặc thích nghi.

Điều này được thể hiện trong việc đánh dấu trên Hình 16. Giống như trước, những sự lựa chọn của Cột trong trò chơi ban đầu gồm hai nước đi được gắn tên là I và II. Những lựa chọn của Hàng là i và ii. Ngoài ra, ký hiệu "2" sẽ chỉ sự cam kết của Hàng đối với hàng ii, "1" là sự cam kết đối với hàng

	I	II	III	IV	V	VI	VII	VIII
	0-I 1-I 2-I	0-I 1-I 2-II	0-I 1-II 2-I	0-I 1-II 2-II	0-II 1-I 2-I	0-II 1-I 2-II	0-II 1-II 2-I	0-II 1-II 2-II
I	0; I-i, II-i	8 5	2 5	2 5	2 5	1 0	1 0	1 0
II	0; I-i, II-ii	0 1	0 1	0 1	0 1	5 2	5 2	5 2
III	0; I-i, II-ii	8 5	2 5	2 5	2 5	5 2	5 2	5 2
IV	0; I-i, II-i	0 1	0 1	0 1	0 1	1 0	1 0	1 0
V	1; I-i, II-i	8 5	2 5	1 0	1 0	5 2	5 2	1 0
VI	1; I-i, II-ii	-5 1	-5 1	0 2	0 2	-5 1	-5 1	0 2
VII	1; I-i, II-ii	8 5	2 5	0 2	0 2	5 2	5 2	0 2
VIII	1; I-i, II-i	-5 1	-5 1	1 0	1 0	-5 1	-5 1	0 1
IX	2; I-i, II-i	-5 5	-6 0	-5 5	-4 0	-5 5	-4 0	-5 5
X	2; I-i, II-ii	0 1	5 -2	0 1	5 -2	0 1	5 -2	0 1
XI	2; I-i, II-ii	-5 5	5 2	-5 5	5 2	-5 5	-5 5	5 2
XII	2; I-i, II-i	0 1	-4 0	0 1	-4 0	0 1	-4 0	0 1

H. 16

i, và "0" là quyết định không tự cam kết. Trong trò chơi mở rộng, mỗi "chiến lược" đơn nhất đối với Cột giờ đây được chỉ định bởi ba cặp ký hiệu, như là 0-I, 1-II, 2-I, mà sẽ có nghĩa là "chọn cột I nếu anh ta không tự cam kết, cột II nếu anh ta cam kết đối với hàng 1, và cột I nếu anh ta cam kết hàng 2." Đối với Hàng, một chiến lược bao gồm quyết định về 0, 1, hoặc 2, cộng với một cặp ký hiệu chỉ ra anh ta sẽ phản ứng như thế nào đối với từng lựa chọn của Cột. Chẳng hạn, 1; I-i, II-i có nghĩa là "Cam kết đối với hàng i, rồi chọn hàng i bất

kể Cột chọn gì." Biết được các phần thưởng của trò chơi gốc, Hình 15 A, những người chơi có thể nhận diện các phần thưởng trong trò chơi mở rộng của Hình 16. Chúng ta có thể tưởng tượng tương Hàng và Cột, thay vì gặp nhau để chơi trò chơi gốc, sẽ gửi các đại diện của mình để chơi cho mình, mỗi đại diện được chỉ dẫn đầy đủ đối với tất cả tình huống đã dự tính (nghĩa là được cho từng chiến lược cụ thể trong trò chơi mở rộng). Để xác định những chỉ dẫn nào được đưa ra, Hàng và Cột xem xét ma trận trên Hình 16. Tức là, họ sẽ chơi trò chơi ngầm trong ma trận đó, để cho các đại diện của họ chỉ đóng vai trò người giao liên.

"Đáp số" của trò chơi ngầm mở rộng này là gì? Hay, nói đúng hơn, liệu chúng ta có thể nhận diện một đáp số rõ ràng cho trò chơi gốc hay không? Và, nếu vậy, nó thể hiện như thế nào trong ma trận mở rộng? Trò chơi gốc rõ ràng có một đáp số dành cho những người chơi hợp lý. (A) Nếu Hàng cam kết đối với hàng i, với hình phạt bằng 5 nếu phá vỡ cam kết của mình, Cột có thể thấy rằng hàng i sẽ được chọn, bất kể anh ta chọn cột nào. Cột chọn cái ô ưu tiên của mình trong hàng trên, tức là ô nằm trên bên trái, i,I. Và Hàng biết rằng, nếu anh ta tự cam kết đối với hàng i, anh ta sẽ có phần thưởng trong ô phía trên bên trái, tức là 2. (B) Nếu, thay vào đó, Hàng lại cam kết hàng ii (lấy phần thưởng trong hàng i trừ đi 5), Cột chọn II thay vì I, và Hàng biết rằng anh ta sẽ nhận được 5. Cuối cùng, (C) nếu Hàng không cam kết, Cột biết rằng Hàng sẽ chọn hàng phần thưởng cao nhất trong cột đã chọn. Tức là nếu Cột chọn I, Hàng chọn i, và Cột nhận được 5. Nếu Cột chọn II, Hàng chọn ii, và Cột nhận được 2. Cột thích I hơn, điều này sẽ để lại cho Hàng phần thưởng bằng 2, và Hàng có

thể dự kiến được điều này. Như vậy kết quả tốt nhất của Hàng là cam kết với hàng ii. Đây là “đáp số” rõ rệt. Nó có phần thưởng [<sub>5</sub><sup>2</sup>], và nó tương ứng với chiến lược 2, I-ii, II-ii cho Hàng, và đối với tất cả bốn chiến lược chứa 2-II cho Cột. (Điều mà Cột có thể đã làm trong các tình huống 0 và 1 không bị ảnh hưởng gì, một khi Hàng đã đi nước đầu tiên.) Những chiến lược này được đánh dấu sao trong các ô ở Hình 16, hàng x. (Trong thực tế, nước đi đầu tiên của Hàng là sự lựa chọn để chơi một trong ba trò chơi hai nước đi khác nhau A, B, và C như trên Hình 15, trong đó anh ta chọn nước đi thứ hai.)

Chúng ta mô tả các ô như thế nào, hay các cặp chiến lược, mà đại diện cho đáp số trên Hình 16? Chúng xác lập một đáp số thuộc dạng được gọi là *đáp số theo nghĩa yêu hoàn toàn*.<sup>25</sup> Chúng ta có thể tìm được nó, trong khuôn khổ của ma trận mở rộng, với một quá trình xóa bỏ các hàng và các chiến lược “bị lấn át”. Một hàng bị lấn át bởi hàng khác nếu mọi phần thưởng đối với Hàng trong hàng trội ít nhất cũng bằng phần thưởng tương ứng trong hàng bị lấn át và ít nhất có một phần thưởng là tốt hơn. Áp dụng tiêu chí này, hàng thứ nhất bị lấn át bởi hàng thứ ba, và chúng ta gạch bỏ nó đi. (Lập luận có thể là Hàng có thể loại bỏ một cách an toàn chiến lược được đại diện bởi hàng thứ nhất, vì hàng thứ ba ít nhất cũng tốt ngang bằng trong mọi tình huống và thực sự tốt hơn trong một số tình huống.) Hàng thứ hai, hàng thứ tư cũng như vậy, cũng như tất cả hàng còn lại, trừ hàng thứ 10. Cả hàng thứ ba và hàng thứ 10, không hàng nào trội hơn hàng kia, do đó lúc này chúng ta vẫn giữ cả hai. So sánh các cột, không cột

---

<sup>25</sup> So sánh với Luce và Raiffa, trang 106-109.

nào trội hơn các cột khác. Nhưng, sau khi loại trừ tất cả các hàng trừ hàng thứ ba và thứ 10 (với lập luận rằng Hàng cuối cùng sẽ không chọn những hàng này), Cột có thể chỉ cần so sánh các ô thứ ba và thứ 10 trong các cột. Bây giờ rõ ràng là cột thứ hai trội hơn cột thứ nhất, thứ ba, thứ năm, và thứ bảy. Sau khi gạch bỏ những cột bị lấn át trong tập hợp rút gọn các hàng, chúng ta có thể nhìn lần nữa vào các hàng iii và x. Lúc đầu, không hàng nào trội hơn hàng kia. Nhưng với việc *các cột* thứ nhất, thứ ba, thứ năm và thứ bảy không còn, hàng thứ 10 sẽ trội hơn hàng thứ ba. Gạch bỏ hàng thứ ba, chúng ta còn lại một hàng duy nhất, hàng x, hàng mà giao với bốn cột. Các phần thưởng là giống nhau trong bốn giao điểm, nói lên rằng việc Cột chơi chiến lược nào trong bốn chiến lược đó không dẫn đến khác biệt gì về lợi ích cho anh ta, chừng nào mà Hàng chơi hàng thứ 10. (Tức là, một khi Hàng đã cam kết bản thân đối với hàng thứ hai của ma trận gốc 2 x 2, Hình 15 A, như Cột kỳ vọng anh ta làm, thì sẽ không có gì khác biệt trong việc Cột giao cho đại diện của mình chỉ dẫn nào về hai trường hợp đã không xảy ra.)<sup>26</sup>

---

<sup>26</sup> Đáng lưu ý rằng thứ tự mà chúng ta loại bỏ các hàng và các cột mà dù tiêu chuẩn loại bỏ có thể ảnh hưởng đến hình dạng của “đáp số.” Trong thủ tục được nêu ra trong bài, trước tiên chúng ta loại bỏ tất cả các hàng trừ hàng ba và 10; sau đó chúng ta nhận thấy rằng các cột I, III, V, và VII, dù tiêu chuẩn để loại bỏ, và chúng ta loại bỏ chúng; vào lúc này, hàng iii bị lấn át, và nó bị loại bỏ; và chúng ta còn lại hàng x giao với bốn cột mà đưa ra các phần thưởng đồng nhất trong hàng đó. Nhưng chúng ta có thể đã lưu ý, khi chúng ta loại bỏ bốn cột, rằng hai cột nữa cũng có thể bị loại bỏ tại giai đoạn đó, đó là các cột VI và VIII, mà sẽ cho thấy các phần thưởng thấp hơn đối với Cột, trong hàng iii, so với các cột II và IV. Nói cách khác, tại điểm đó trong quá trình, hàng iii và các cột VI và VIII đủ tiêu chuẩn để loại bỏ; nhưng nếu chúng ta tùy ý chọn trước hết loại bỏ hàng iii và sau đó chuyển đến các cột, hai cột đang xem xét

Như thế, đây là cách mà một đáp số cho *trò chơi có nước đi nối tiếp* ban đầu thể hiện trong trò chơi tinh ("không có nước đi," hoặc lựa chọn ngầm đồng thời). Nó là một đáp số đạt được bằng cách loại bỏ các chiến lược bị lẩn át, với tiêu chí lẩn át áp dụng chỉ cho những chiến lược chưa bị loại bỏ trong từng giai đoạn. Đây dường như là hình thức chung của đáp số trong trò chơi ngầm mở rộng tương ứng với trò chơi có nước đi nối tiếp khi mà cái sau có một đáp số xác định. Việc

---

sẽ không còn bị lẩn át nữa. Như thế, theo nghĩa nào đó, nội dung của "đáp số" của chúng ta phụ thuộc vào một lựa chọn thủ tục tùy nghi; tuy vậy, bất kể chúng ta còn lại với hai ô có các phần thưởng giống nhau hay bốn ô có phần thưởng giống nhau phụ thuộc vào sự lựa chọn tùy nghi đó. Mặc dù vậy, các phần thưởng là giống nhau trong mỗi trường hợp. Lý do có thể là tại một giai đoạn Cột nhận thấy anh ta không cần lý giải thêm, rằng Hàng có một lựa chọn đã được xác định rõ ràng làm cho việc liệu Cột có thu hẹp quyết định của mình hay không là không có tác dụng gì, nhưng điểm chính xác mà tại đó anh nhận ra điều này, và những cột nào còn lại mà không bị loại bỏ khi anh ta nhận ra, sẽ phụ thuộc ở mức độ nào đó vào nhiều con đường khác nhau mà anh ta theo đuổi trong quá trình lập luận của mình. (Nếu có các chi phí thông tin liên lạc trong việc thu hẹp sự lựa chọn chiến lược của anh ta, Cột có thể chỉ ưu tiên chọn chiến lược 2-II, để lại sự lựa chọn chưa được xác định tương ứng với chiến lược 0 và 1 của Hàng. Lấy trường hợp ngược lại, nếu có những rủi ro chiến lược của Hàng bị ghi chép sai sót hoặc thông tin không đúng, hoặc lựa chọn không sáng suốt, Cột sẽ giảm rủi ro bằng cách chỉ định 0-I. Trong trường hợp sau, thực ra, anh ta xem hàng III không phải là không thể mặc dù nó bị lẩn át bởi hàng X. Và nếu, để xử lý tiếp, anh ta nghĩ ngờ rằng trọng tài có xu hướng lắng nghe "hàng V" khi các hàng khác thực sự được chọn, anh ta có thể thu hẹp tiếp sự lựa chọn của anh ta xuống 0-I, 1-I, 2-II, "đáp số" sẽ là giao điểm của hàng X và Cột II, vì giao điểm của V và IV là thấp hơn của V và II và cho anh ta lý do để tiếp tục hoàn thiện sự lựa chọn của mình. Nói chung, bằng cách gắn các rủi ro về sai sót các loại hoặc các chi phí khác nhau của các phương pháp khác nhau để xác định một chiến lược, một bài toán khá phong phú hơn được hình thành, và là bài toán có thể dẫn đến các kết luận khác. Những bài toán được xem xét ở Chương 7 và 9, bao gồm một số hình thức nhất định của hành vi ngẫu nhiên, sai sót, hoặc thông tin sai, có thể đem lại kết quả loại này.)

loại bỏ các hàng và các cột có thể được coi trùng với quá trình tính toán trước *nước đi cuối cùng* hợp lý trong tất cả các tập hợp các nước đi có thể xảy ra trước đó, sau đó, biết rằng nước đi cuối cùng nào sẽ theo sau từng nước đi áp trước nước cuối cùng, tính toán nước đi áp trước nước đi cuối cùng tốt nhất cho tất cả các tập hợp các nước đi có thể trước đó và cứ như thế ngược trở lại nước đi đầu tiên tốt nhất của trò chơi.

Trong khi nó có ích và thỏa mãn hiểu biết trong việc nghiên cứu xem bằng cách nào các chiến thuật như là đe dọa, cam kết, và hứa hẹn có thể được hấp thụ vào “siêu trò chơi” trừu tượng (trò chơi ở “dạng chính tắc”), cần nhấn mạnh rằng chúng ta không học tập được gì về các chiến thuật này bằng cách nghiên cứu các trò chơi mà đã ở dạng chính tắc. Các đối tượng nghiên cứu của chúng ta, như là các chiến thuật cùng với các cơ cấu trao đổi thông tin và cưỡng chế mà chúng phụ thuộc vào, và việc sắp đặt các nước đi theo thời gian, tất cả đã biến mất vào lúc trò chơi ở dạng chính tắc. Điều chúng ta muốn là một lý thuyết hệ thống hóa sự nghiên cứu về các thành phần phổ biến khác nhau tạo nên các cấu trúc nước đi của trò chơi, mà một mô hình quá trừu tượng sẽ bỏ qua các thành phần đó.<sup>27</sup>

---

<sup>27</sup> Cần nói thêm rằng, việc ép một trò chơi cụ thể thành dạng ma trận siêu trò chơi nói chung không phải là một kỹ thuật khả thi để phân tích; số lượng hàng và cột (tức là, số lượng các chiến lược nước đi nối tiếp) trở nên vô cùng lớn, ngay cả cho các trò chơi đơn giản. Để minh họa, chúng ta xem xét ma trận  $3 \times 3$ , với Cột sẽ đi trước; cộng thêm một cơ hội trước đó cho Hàng để tự cam kết đối với bất kỳ chiến lược phản ứng được xác định một phần hay toàn bộ nào; cuối cùng, để nghiên cứu “sự phòng thủ” chống lại các đe dọa, cho phép Cột một cơ hội trước đó để cam kết lựa chọn cột. Tức là, Cột có thể cam kết bản thân trước một cách vô điều kiện nếu anh ta muốn, Hàng sau đó có thể cam kết bản thân có điều kiện theo bất cứ cách nào anh ta muốn, rồi sau đó Cột

Tuy nhiên, sự thể hiện bằng ma trận của một trò chơi có nước đi nối tiếp giúp nhấn mạnh rằng “tính xác định” chuẩn của các trò chơi được giải quyết bằng các nước đi chiến thuật không làm nó đi lệch khỏi tính cốt lõi của trò chơi về chiến lược. Sự đe dọa “chiến thắng” và quyết định một kết cục chỉ bởi vì nó khiến người chơi kia lựa chọn cái có lợi cho mình. Người chơi kia giữ lại sự tự do lựa chọn ban đầu của mình, và sự lựa chọn của anh ta vẫn phụ thuộc vào dự đoán của anh ta về sự lựa chọn cuối cùng của người đe dọa. Như vậy, lựa chọn đầu tiên của người đe dọa - đe dọa hay là không - sẽ phụ thuộc vào điều mà anh ta kỳ vọng ở người bị đe dọa sẽ kỳ vọng người đe dọa sẽ làm. Tính chất kỳ vọng tương hỗ của trò chơi vẫn giữ nguyên. Sự đe dọa, giống như sự cam kết vô điều kiện hoặc giống như khái niệm rộng hơn của “hàm phản ứng”, khi có sẵn nhiều sự lựa chọn cho hành động, có hiệu lực bởi việc hạn chế các kỳ vọng của người chơi khác thông qua sự vận dụng những động cơ của chính mình.

---

chọn một cột và cuối cùng Hàng chọn một hàng. Chúng ta hãy dừng làm phức tạp trò chơi bằng cách giới hạn các quy mô của hình phạt hoặc bằng cách đưa vào sự bất định hay bất kỳ hệ thống thông tin liên lạc không hoàn thiện nào. Trò chơi “đơn giản” này, mà không phải quá khó để phân tích dưới dạng mở rộng của nó, hóa ra có nhiều hơn một “googol” (số 1 và một trăm số không đi theo -  $10^{100}$ ) số các cột.

## *NGHỊCH LÝ CỦA LỢI THẾ CHIẾN LUỘC*

Tất nhiên, sẽ là một nguyên tắc hệ quả rằng nếu ma trận lợi ích ngay từ đầu đã cho thấy các giá trị dành cho một trong các người chơi bị giảm đi theo cùng một khuôn mẫu như trong một ma trận trong đó anh ta sẽ có tình làm giảm nó tại nước đi đem lại thắng lợi, anh ta chỉ đơn giản thắng mà không cần có nước đi công khai. (Đây là điểm mà dưới hình thức sơ đồ, được minh họa ở đoạn cuối của Chương 2, và được nói tới như là một ví dụ trừu tượng của nguyên tắc mà, trong mặc cả, điểm yếu có thể là sức mạnh.) Có lẽ không có một nguyên tắc nào của lý thuyết trò chơi hiển thị nổi bật trò chơi có động cơ hỗn hợp như nguyên tắc rằng việc làm xấu đi một số hoặc thậm chí tất cả kết cục tiềm tàng cho một người chơi cụ thể và việc không có sự cải thiện nào ở các kết cục đó có thể là, một cách rõ ràng - thậm chí kịch tính - lợi thế cho người chơi bị bất lợi như vậy. Nó giải thích vì sao một hình phạt nghiêm khắc và chắc chắn được thực hiện đối với việc *đưa tiền* cho kẻ tổng tiền có thể bảo vệ cho nạn nhân tiềm năng, bằng cách nào mà việc đốt cháy các cây cầu sau lưng mình khi đối mặt với kẻ thù có thể làm mất tinh thần kẻ địch và khiến chúng rút lui, hoặc vì sao một phụ nữ, ở thời đại xưa kia, dám thách đố bên tìm kiếm bằng cách liều lĩnh đặt vật tìm kiếm trong ngực mình.<sup>28</sup>

---

<sup>28</sup> Nó cũng giải thích vì sao một "lời hứa" tránh sự lựa chọn mà sẽ gây hại cho bên kia có thể không được bên kia hoan nghênh. Một lời hứa mà

Trong chiến tranh Triều Tiên người ta đưa tin không chính thức rằng khi Bộ Tài chính phong tỏa các nguồn tài chính của Trung Quốc, thì mọi người cũng biết rằng một số tài sản không phải của Trung Quốc cũng bị phong tỏa như một biện pháp để tránh cho các chủ sở hữu khỏi bị đe dọa cưỡng ép đối với thân nhân của họ vẫn còn ở Trung Quốc. Hoàn toàn giống như thế, đối với các chủ sở hữu sống tại Hoa Kỳ, chính các hình phạt đối với việc chuyển tiền cho Trung Quốc sẽ nâng cao khả năng của họ để chống lại sự cưỡng ép. Chủ ý đặt tài sản của mình dưới dạng làm cho việc trốn tránh luật pháp trở

*cho phép* anh ta thoải mái lựa chọn có thể đảm bảo cho chúng ta rằng anh ta sẽ thực hiện nó, do đó chúng ta có thể đưa vào điều đó và đưa ra một số lựa chọn trước bất lợi cho anh ta. Tương tự như vậy, *việc cộng thêm* các giá trị một cách chọn lọc vào các phần thưởng của người khác có thể hoàn toàn làm xấu đi vị thế của anh ta - nếu chúng ta có phương tiện để thực hiện việc cộng thêm đó. Trong ma trận kèm theo, giả sử Hàng có nước đi đầu tiên, Hàng có thể "thắng" - anh ta có thể nhận được 7 và Cột chịu thiệt - nếu anh ta đơn phương bảo đảm sẽ bù lại cho Cột trong trường hợp có kết quả tại i,II, sự đèn bù xuất phát từ chiến thắng của chính anh ta. Nếu anh ta hứa sẽ trả 2 cho Cột trong trường hợp đó, anh ta sẽ nhận được 8, Cột nhận được 3. Nếu ngược lại, khi không có lời hứa đèn bù, Hàng không thể chọn i, và kết quả là tại ii,I với phần thưởng bằng 1 và 10 tương ứng. Rõ ràng Cột muốn rằng Hàng không thể cam kết bản thân để cho "quà." (Nếu kẻ tổng tiền không thể hạ mức đòi hỏi xuống tới một mức đòi hỏi, mà khi cộng với mức phạt đối với việc trả tiền cho kẻ tổng tiền, thấp hơn thiệt hại mà hắn đe dọa, hắn có thể dễ dàng trả khoản tiền phạt của nạn nhân. Điều này bảo đảm cho sự đáp ứng từ nạn nhân của hắn. Do đó, sự đe dọa được thực hiện với bất lợi cho nạn nhân).

0	2	10	1	
1	10	2	0	
				trở thành

0	2	8	3	
1	10	2	0	

nên khó khăn hơn, hoặc vận động hành lang để có hình phạt nặng hơn đối với việc chuyển trái phép tiền của mình, hoặc thậm chí tạm thời nhận mình là kẻ có cảm tình với Cộng sản để cho tài sản của mình bị phong tỏa có thể là một chiến thuật bắt buộc cho các nạn nhân tiềm năng, để ngăn chặn các đe dọa cưỡng ép từ trước.

Một nguyên tắc tương tự được phản ánh trong Điều 26 của hiệp định hòa bình Nhật Bản, cho Hoa Kỳ một số quyền nhất định nếu việc chuyển nhượng về sau lãnh thổ Nhật Bản cho các cường quốc khác có lợi hơn. Khi có tin người Nhật chịu áp lực của người Nga về các chuyển nhượng lãnh thổ vào năm 1956, Bộ trưởng Ngoại giao Hoa Kỳ John Foster Dulles đã mô tả một cách nhấn mạnh điều khoản đó tại cuộc họp báo và nói rằng ông ta vừa mới “nhắc nhở người Nhật về sự tồn tại của các điều khoản đó.”<sup>29</sup> Ý định rõ ràng là để củng cố sự phản kháng của Nhật Bản. Và có thể coi rằng bằng cách “nhắc nhở” người Nga về điều khoản đó thông qua trung gian báo chí, Dulles giúp đem lại cho người Nhật câu mặc cả quen thuộc, “Nếu tôi làm điều đó cho anh, tôi phải làm điều đó với tất cả người khác.” Đó là, theo thuật ngữ đã dùng từ trước, một “sự cam kết” được bảo đảm bằng một hình phạt bồi thường cho Hoa Kỳ. (Điều nghịch lý là Hoa Kỳ không thể cho Nhật Bản lợi ích của mánh lới mặc cả này trừ phi Hoa Kỳ có động cơ rõ ràng trong việc tận dụng quyền đòi lãnh thổ của mình nếu chiến thuật thất bại.)<sup>30</sup>

<sup>29</sup> Bản ghi lời bình luận của Bộ trưởng Ngoại giao Hoa Kỳ Dulles tại buổi họp báo, *The New York Times* (ngày 29-8-1956), trang 4.

<sup>30</sup> Việc lập trường của một bên có thể bị làm cho suy yếu một cách tai hại bởi các quyền lực pháp lý mới được gọi ý một cách sâu sắc bởi một

## **“NHỮNG NƯỚC ĐI CHIẾN LƯỢC”**

Nếu thực chất của một trò chơi về chiến lược là sự phụ thuộc của sự lựa chọn hành động đúng của mỗi bên vào cái điều anh ta kỳ vọng bên kia sẽ hành động, thì có thể sẽ có ích khi định nghĩa một “nước đi chiến lược” như sau: Một nước đi chiến lược là nước đi có ảnh hưởng đến sự lựa chọn của người khác, theo một cách có lợi cho mình, bằng việc gây ảnh hưởng đến các kỳ vọng của người khác về cách mà mình sẽ hành động. Người ta hạn chế sự lựa chọn của đối tác của mình bằng cách hạn chế hành vi của chính mình. Mục đích là thiết lập cho chính mình và truyền đi một cách thuyết phục đến bên kia một phương thức hành vi (bao gồm các phản ứng có điều kiện đối với hành vi của bên kia) mà sẽ để lại cho bên kia một bài toán tối đa hóa đơn giản mà đáp số của nó đối với anh ta là tối ưu cho bản thân mình, và loại bỏ khả năng làm điều tương tự của bên kia.

Có lẽ không có sự tương phản nào nổi bật hơn trong việc so sánh trò chơi động có cơ hỗn hợp và trò chơi xung đột thuần túy (có tổng lợi ích bằng không) hơn tầm quan trọng của việc để cho chiến lược của một bên được bên kia phát hiện

---

trong các lập luận được nêu ra để chống lại việc hợp pháp hóa cái chết không đau đớn, cho phép những bệnh nhân không thể cứu được có quyền yêu cầu được kết liễu: “Sẽ có ... tác động gì đối với những người già ốm yếu không chữa được, những người đã nghĩ rằng những người xung quanh muốn kết liễu họ?” (John Beavan, “Quyền Sống-và Chết của Bệnh Nhân” (The Patient's Right to Live-and Die), *The New York Times Magazine* (ngày 9-8-1959), trang 14, 21-22.)

và thấu hiểu. Khó có cái gì nói đúng tinh thần của trò chơi có tổng lợi ích bằng không bằng tầm quan trọng của việc “không bị phát hiện” và của việc sử dụng một phương thức ra quyết định được bảo vệ chống lại dự đoán suy diễn bởi đối phương.<sup>31</sup> Khó có cái gì đại diện cho hành vi chiến lược trong trò chơi có động cơ hỗn hợp tốt hơn là lợi thế của việc có thể áp dụng một phương thức hành vi mà bên kia sẽ coi đó là điều dĩ nhiên.

Tất nhiên, có thể là một lợi thế trong trò chơi có tổng lợi ích bằng không khi làm cho đối thủ tin chắc vào một lối chơi cụ thể của mình, nhưng chỉ khi nếu niềm tin đó là sai. Trong trò chơi có động cơ hỗn hợp, một bên có lợi trong việc truyền đi sự thật về hành vi của mình – nếu, quả thực, anh ta thành công trong việc hạn chế hành vi của chính mình theo hướng mà, khi được bên kia đoán đúng, sẽ chiến thắng.

Nghịch lý khác của các trò chơi động cơ hỗn hợp là việc thực sự không hiểu biết có thể là một lợi thế cho một người chơi nếu nó được đối phương nhận ra và tính đến. Nghịch lý này, mà có thể xuất hiện hoặc trong bài toán phối hợp hoặc trong việc miễn trừ khỏi bị đe dọa, không có gì tương ứng trong các trò chơi có tổng lợi ích bằng không. Và tương tự, trong trò chơi có tổng lợi ích bằng không giữa những người chơi hợp lý với thông tin đầy đủ thì có thể sẽ không bao giờ là một lợi thế khi phải đi trước (để chơi “trò chơi thiểu số” theo ngôn từ của von Neumann và Morgenstern). Còn trong trò chơi hỗn hợp thì chắc chắn là có thể.

---

<sup>31</sup> Về điểm này, von Neumann và Morgenstern nói (trang 147): “Chúng tôi đã đặt những xem xét về nguy cơ để cho chiến lược của một bên bị phát hiện bởi đối phương vào một vị trí trung tâm tuyệt đối.”

## 6

# LÝ THUYẾT TRÒ CHƠI VÀ NGHIÊN CỨU THỰC NGHIỆM

Phần thảo luận trên gợi mở một số kết luận về một phương pháp thích hợp để nghiên cứu các trò chơi mặc cả. Thứ nhất là không nên để cấu trúc toán học của hàm lợi ích đè nặng lên công việc nghiên cứu. Kết luận thứ hai, hơi chung chung hơn một chút, là trừu tượng hóa quá mức sẽ gây ra một số hiểm họa: (thí dụ) chúng ta sẽ làm thay đổi tính chất của trò chơi nếu ta sửa đổi một cách quá mức các chi tiết tình huống của trò chơi, hoặc khi ta loại trừ các giả định gây phức tạp [cho phân tích] chẳng hạn như việc những người chơi không biết chắc chắn về các hệ giá trị của nhau. Thông thường thì chính chi tiết tình huống có thể giúp cho những người tham gia cuộc chơi khám phá ra một kết cuộc ổn định, hoặc ít nhất là kết cuộc không có hại cho nhau. Xét theo ví dụ nêu ở phần trước, khả năng của Holmes và Moriarty cùng xuống tàu ở một nhà ga có thể tùy thuộc vào sự tồn tại của một (số) yếu tố tình huống hơn là phụ thuộc vào cấu trúc hình thái của trò chơi. Đó có thể là một sự kiện hay hiện tượng trên xe lửa, dưới nhà

ga, một cái gì đó mà họ cùng biết từ trước, hay cùng nghe được qua loa phóng thanh khi xe lửa dừng lại; và mặc dù khó có thể khái quát hóa thành các nhận định khoa học rằng cái gì sẽ dẫn đến khả năng phối hợp thành công giữa hai người, chúng ta phải thừa nhận rằng những yếu tố quyết định kết cục của trò chơi như vậy thường bị các phân tích trừu tượng ở mức cao coi là chi tiết không liên quan.

Kết luận thứ ba - đặc biệt thích hợp trong trường hợp thông tin không hoàn hảo, khi người chơi không biết chắc về hệ giá trị của nhau, hoặc không biết chắc về các chọn lựa chiến lược, và đặc biệt là khi kết cục chỉ có thể đạt tới bằng một chuỗi nước đi hoặc chuỗi thao tác - là một phần thiết yếu của nghiên cứu về các trò-chơi-có-động-cơ-hỗn-hợp phải dựa vào thực nghiệm. Nói thế không có nghĩa câu hỏi người ta thật sự hành động như thế nào trong các trò chơi có động cơ hỗn hợp, đặc biệt là các trò chơi quá phức tạp để nắm bắt thành thạo, là một câu hỏi thực nghiệm. Mà nó là một khẳng định mạnh mẽ hơn: rằng các nguyên tắc quy định việc chơi *thành công*, các nguyên tắc *chiến lược*, các định đề lý thuyết *chuẩn tắc*, không thể chỉ rút ra bằng cách phân tích thuần túy dựa trên các tính toán thực hiện trước khi trò chơi diễn ra.

Trong một trò chơi tổng-lợi-ích-bằng-không, nhà phân tích chỉ thật sự phải đương đầu với một trung tâm nhận thức, một nguồn quyết định duy nhất. Quả thật, có hai người tham gia trò chơi, mỗi người với nhận thức của riêng mình; nhưng chiến lược tối-đa-tối-thiểu đã biến tình huống ban đầu thành một tình huống chỉ gồm hai quyết định hoàn toàn đơn phương. Giữa hai người không cần phải biết về nhau; không cần có sự tâm đầu ý hợp; không cần phải trao đổi những lời gợi ý;

không phải so sánh ẩn tượng, hình ảnh hoặc sự cảm thông. Không có yếu tố nhận thức xã hội ở trong đó. Nhưng trong trò chơi có động cơ hỗn hợp, hai hoặc nhiều trung tâm nhận thức phải lệ thuộc lẫn nhau một cách đặc biệt cẩn thiết. Cần phải có trao đổi thông tin; ít nhất là một chút nhận biết phải được trao đổi giữa những người chơi. Thông thường cần phải có một hành động xã hội nào đó, cho dù dưới hình thức thô sơ hay ngầm hiểu; và cả hai người chơi ở một mức độ nào đó sẽ phải lệ thuộc vào mức độ nhận thức xã hội và tương tác của họ. Ngay cả khi có hai người hoàn toàn tách biệt chơi với nhau trong sự yên lặng tuyệt đối, thậm chí không nhận diện được nhau, thì họ cũng phải ngầm có được một mức độ đồng cảm với nhau.

Do đó, không có cách nào để một nhà phân tích có thể tái hiện toàn bộ quy trình ra quyết định, cho dù theo phương pháp nội quan hay phương pháp tiên đề. Không có cách nào để xây dựng một mô hình tương tác giữa hai hay nhiều đơn vị ra quyết định, với hành vi và những kỳ vọng của những đơn vị ra quyết định đó xuất phát từ suy luận thuần túy. Một nhà phân tích có thể suy luận ra các quyết định của một đầu óc duy lý chỉ khi biết được các chuẩn mực quy định các quyết định đó; nhưng bằng phương pháp phân tích thuần chính thống thì không thể suy luận ra những gì tương tác giữa hai trung tâm nhận thức. Cần phải có ít nhất hai người để kiểm chứng các yếu tố này. (Hai nhà phân tích có thể làm được, nhưng chỉ có thể được bằng cách sử dụng chính họ như những đối tượng trong một cuộc thí nghiệm.) Nhận một ẩn ý về cản bản hoàn toàn khác với giải mã một thông điệp chính thức hoặc giải quyết một vấn đề toán học; nó bao gồm phát hiện

thông điệp được cài đặt trong một bối cảnh cụ thể bởi một người khác khi anh này nghĩ rằng mình và người nhận thông điệp có cùng một số ẩn tượng hay mối liên kết nào đó. Nếu không có bằng chứng thực nghiệm, chúng ta không thể chỉ dựa trên phép diễn dịch hình thức để kết luận rằng ẩn ý nào sẽ được hiểu [và ẩn ý nào không được hiểu] trong một trò chơi-kết-quả-khác-không; cũng tựa như chẳng thể kết luận chỉ dựa trên phép diễn dịch hình thức rằng một chuyện tiểu lâm nào đó sẽ buồn cười [hay không buồn cười].

Để minh họa, hãy xem xét trường hợp hai người cùng nhìn vào một vết mực, liệu cả hai có nhận ra cùng một hình ảnh hay cùng một gợi ý ở trong đó không nếu mỗi người đang cố gắng và biết rằng người kia cũng đang cố gắng sắp đặt để có cùng một hình ảnh hay một gợi ý? Câu trả lời cho câu hỏi này chỉ có được bằng cách thử nghiệm. Nhưng, nếu họ có thể, thì họ đã làm được điều mà một lý thuyết trò chơi *thuần hình thức* không thể xem xét phân tích; họ có thể làm *tốt hơn* mức mà một lý thuyết trò chơi diễn dịch thuần túy có thể tính đến. Và, nếu họ có thể làm tốt hơn nữa - nếu họ có thể vượt qua những giới hạn của một lý thuyết trò chơi thuần hình thức - thì ngay cả một lý thuyết chuẩn tắc, nguyên tắc và chiến lược cũng không thể chỉ dựa trên nền tảng phân tích thuần hình thức. Chúng ta không thể xây dựng một lý thuyết mô tả hay một lý thuyết nguyên tắc trên cơ sở giả thiết rằng có những quy trình tri thức nào đó mà những người chơi có lý trí *không* thể hiểu và sử dụng được, và các quy trình đó thuộc loại “nhận một ẩn ý”. Câu hỏi liệu người chơi có lý trí, khi hành động một mình hay hành động tập thể, có thể làm tốt hơn mức mà một lý thuyết trò chơi thuần hình thức dự báo hay không, và do đó có thể bỏ qua các

nguyên tắc chiến lược được do lý thuyết này quy định hay không, là một câu hỏi thực nghiệm.<sup>1</sup>

<sup>1</sup> Một ví dụ thực nghiệm tốt về khía cạnh thông tin-và -tiếp nhận thông tin trong chiến lược trò chơi là thí nghiệm được trình bày bởi M. M. Flood. Ông đã cho những người tham gia chơi một trò chơi tổng lợi ích khác không  $2 \times 2$  (có hai người chơi, và trong một lần chơi mỗi người có 2 nước đi) và họ được chơi trò này 100 lần. Điểm đặc biệt trong ma trận kết quả của trò này là những người tham gia chỉ có thể thắng nếu 2 người phối hợp để kết quả rơi vào đúng một số ô nào đó trong mỗi lần chơi [thường là 2 trong 4 ô, một ô có lợi hơn cho người này, và ô còn lại có lợi hơn cho người kia] Tuy nhiên, để phân phối phần thắng trong chuỗi 100 lần chơi, họ phải thỏa thuận [ngầm] được với nhau một phương thức thay đổi từ một ô thắng sang ô thắng còn lại, khi phần thưởng trong mỗi ô thắng chỉ thiên vị một trong hai người. Phương pháp mặc cả duy nhất để tìm ra cách chia phần hợp lý [giữa hai người] qua việc thay đổi ô kết quả là thông qua nước đi mà họ lựa chọn tại mỗi lần chơi. Giai đoạn "thông tin" này - và bắt cứ giai đoạn nào tiếp theo sau đó khi một người chơi phá ngang không chịu đi theo kiểu chơi đã thỏa thuận [ngầm] để một mình hưởng lợi nhiều hơn và bị người kia trả đũa [ở các lần chơi sau] - [là các giai đoạn mà] hai người gộp lại là thiệt thòi [hơn là các giai đoạn mà họ hợp tác thành công], bởi vì một chọn lựa không có sự phối hợp đồng nghĩa với việc mất đi một cơ hội kiếm tiền. M. M. Flood, "Một vài trò chơi thực nghiệm," (*Some Experimental Games*), *Management Science*, 5:5-26 (tháng 10-1958).

Câu hỏi làm cách nào để truyền đạt một đề nghị một cách hiệu quả và làm cách nào diễn giải đề nghị của đối phương ẩn trong cách anh ta chơi rõ ràng là tùy thuộc vào việc hai bên cùng hiểu biết về cách chơi chung - một khả năng mà cả hai cùng biết để hoàn thành chuỗi chơi khi mà một phần của chuỗi này đã được bộc lộ [từ những lần chơi đã thực hiện] - không phải là không giống như quy trình diễn ra trong các thực nghiệm của các nhà tâm lý học thuộc phái Gestalt đã được đề cập trong một phần chú thích trước. Trong khi một lý thuyết thuần hình thức về trao đổi thông tin có thể cho ta một số tiêu chuẩn tối thiểu nào đó về "tính hiệu quả" trong việc trao đổi thông tin mà những người chơi có lý trí phải đạt được, thì câu hỏi liệu những người chơi có thể đạt được kết quả tốt hơn nữa không vẫn là câu hỏi mang tính chất thực nghiệm. Người ta nắm bắt được các ẩn ý tới đâu, những loại ẩn ý nào dễ dẫn tới thành công nhất, là những vấn đề thực nghiệm trong nhận thức xã hội, và có lẽ thích hợp với nghiên cứu thực nghiệm. (Vấn đề tương tự cũng xuất hiện khi hai người tham gia vào một vụ đấu giá nhận ra rằng họ đang cùng mất tiền vì cạnh tranh ra giá với nhau, và

Cũng cần phải nhấn mạnh rằng lý do loại suy luận này không xuất hiện trong trò chơi tổng lợi ích bằng không là vì bất cứ một sự tương tác xã hội nào như thế cũng không thể mang lại lợi ích cho cả hai người chơi cùng một lúc, và vì ít nhất một trong những người chơi có lý trí sẽ có cả động cơ lẫn khả năng để phá hủy tất cả các dạng trao đổi thông tin xã hội. Nhưng trong một trò chơi tổng-lợi-ích-khác-không, khi có sự bất định ban đầu về việc: kết quả nào sẽ hiệu quả nhất trong số các kết quả có thể xảy ra, và liệu có cần có sự phối hợp [giữa những người chơi] để đạt tới kết quả hiệu quả nhất đó hay không, thì một người chơi duy lý không thể rút lui khỏi quá trình xã hội nhằm tự vệ; anh ta không thể tắt máy trợ thính để tránh khỏi bị ức chế do những gì anh ta nghe, nếu sự hoàn toàn tắt tiếng của chiếc radio khiến cho việc hợp tác hiệu quả trở thành bất khả thi. Anh ta cũng không thể không mở một lá thư một cách duy lý khi nó được chuyển đến, bởi vì người khác có thể đã suy luận rằng anh ta sẽ mở thư đó và đã có những bước đi đối phó.

Lúc này, một câu hỏi khác xuất hiện, liêu con đường của lý thuyết trò chơi có liên tục đâm nhánh bao trùm toàn bộ lĩnh vực tâm lý học xã hội, hay nó sẽ dẫn đến một khu vực giới hạn hơn đặc biệt thích hợp với lý thuyết trò chơi. Liệu có thể khám phá ra những định đề tổng quát về hành vi hợp tác trong các trò chơi có động-cơ-hỗn-hợp thông qua thực nghiệm và quan sát hay không, và những định đề đó có dẫn tới

---

cố gắng, khi không đưa ra bất cứ chứng cứ hiển nhiên nào về việc thông đồng, sắp xếp một số hình thức cùng nhau giảm nhẹ mức độ cạnh tranh trong đấu giá để cùng nhau giảm bớt mất mát tiền bạc đồng thời phân phối các khoản lợi cũng như các cơ hội giữa họ.)

những hiểu biết ứng dụng rộng rãi được trong thế giới các tình huống mặc cả không? Mặc dù không có gì bảo đảm về kết quả thành công, nhưng chắc chắn có một số lĩnh vực đầy hứa hẹn cho các nghiên cứu [tiếp theo]; và ngay cả nếu chúng ta không thể khám phá ra những định đề tổng quát, ít nhất chúng ta cũng có thể bác bỏ về mặt thực nghiệm một số cái mà nhiều người vẫn tin tưởng. Rõ ràng là lý thuyết trò chơi hiện nay cực kỳ kém phát triển xét về khía cạnh thực nghiệm.

Hãy xem xét một trò chơi tương tự như trò chơi đã trình bày ở phần trước, bao hàm việc di chuyển các tấm thẻ trên một tấm bản đồ, hay môn cờ vua chỉnh sửa để biến đổi thành trò chơi kết quả khác không. Những trò này có thể được dùng để thể hiện trò chơi trong “chiến tranh giới hạn”; cả hai bên tham gia cùng có lợi bằng cách tránh những chiến lược hủy diệt cả hai. Đây là một trò chơi trong đó khả năng của hai người chơi trong việc tránh gây tổn thất lẫn nhau phụ thuộc rất nhiều vào sự tồn tại của *những phương tiện* giúp phối hợp thành công các ý định [của hai người chơi] được tạo ra bởi những chi tiết tình cờ của trò chơi, bởi những chuyện như hiện trạng của bản đồ hay tấm bảng chơi, những cái tên gợi ý của những quân cờ, truyền thống hay tiền lệ liên quan đến trò chơi, và viễn cảnh hay bối cảnh ý nghĩa được người chơi nhận thức trước khi trò chơi bắt đầu. Đó là một trò chơi đủ phức tạp [mà sự thành công của nó] đòi cả hai bên phải thấu hiểu nước đi cũng như phải có cách truyền đạt ý định của hai người chơi một cách hữu hiệu. Nếu chúng ta giả thiết trong chốc lát rằng khó khăn kỹ thuật trong việc tạo ra một trò chơi thuộc loại đó đã được giải quyết, thì cũng đáng xem xét tiếp xem cần phải tìm hiểu thêm những vấn đề gì và cần phải kiểm

nghiệm những giả thuyết nào. Sau đây là một câu hỏi thuộc loại đó: nhìn chung, có phải rõ ràng là những người tham gia trò chơi dễ đạt tới một giải pháp hiệu quả, nghĩa là một giải pháp không gây tổn thất cho cả hai, hơn khi (a) sự liên lạc đầy đủ hay gần đầy đủ được cho phép, (b) hoàn toàn không hoặc gần hoàn toàn không có sự liên lạc, ngoại trừ những gợi ý chuyển tải từ việc di chuyển (những nước cờ), hoặc (c) sự liên lạc là bất đối xứng, trong đó một người chơi có khả năng truyền đạt thông điệp nhiều hơn mức mà anh ta nhận được? Không có bảo đảm chắc chắn nào rằng một câu trả lời đơn nhất và khả dụng rộng rãi sẽ xuất hiện; tuy nhiên, việc phát hiện ra một số định đề đáng trong trường hợp tổng quát về vai trò của trao đổi thông tin là hoàn toàn có thể. Ý nghĩa to lớn của câu hỏi này được chứng thực bởi một vài trong số những tranh luận gần đây về việc liệu có khả năng kiềm chế chiến tranh ở mức giới hạn hay không nếu có sự liên lạc thông tin giữa hai bên, hoặc nếu có những tuyên bố đơn phương ngay từ đầu của bên này hay bên kia, hoặc nếu hầu như không có sự liên lạc công khai giữa hai bên giao tranh.<sup>2</sup>

---

<sup>2</sup> Để ngăn ngừa bất cứ hình thức hiểu lầm nào: tác giả không có ý muốn nói là chiến tranh giới hạn có thể được mô phỏng trong phòng thí nghiệm hay cho rằng các kết quả thực nghiệm liên quan đến quy trình giới hạn này có thể được chuyển tải trực tiếp vào thế giới thực bên ngoài. Các thực nghiệm về mô hình đã được mô tả cần được dì chung dưới tiêu đề "nghiên cứu cơ bản." Và nó chủ yếu chỉ liên quan đến khía cạnh cảm quan và truyền thông của vấn đề, chứ không phải khía cạnh động cơ - ngoại trừ ở mức độ các động cơ tác động đến nhận thức xã hội. Tuy nhiên, khả năng những kết quả nghiên cứu này có ứng dụng trực tiếp trong thực tế được nâng lên do gần đây phần lớn các lý thuyết nghiên cứu về vai trò của trao đổi thông tin trong chiến tranh giới hạn, hay nghiên cứu về các thể loại giới hạn được thấy nhiều nhất, đều có vẻ dựa trên cái mà người ta mô tả là các trò chơi thực nghiệm ẩn.

Một chuỗi các câu hỏi khác, cũng phù hợp với những vấn đề của chiến tranh giới hạn, ở phạm vi quốc tế hay một mức độ khác, sẽ là liệu một kết quả ổn định, hiệu quả sẽ dễ xuất hiện hơn khi những hàm nghĩa của trò chơi - tên và ý nghĩa được công khai gắn liền với các nước đi, những quân cờ và vật chủ trên bàn cờ - trở nên quen thuộc và dễ nhận biết hoặc khi chúng hoàn toàn mới lạ, không quen thuộc, và có lẽ không tạo ra được những gợi ý tương tự nơi cả hai người chơi. Nói về trò chơi ở một mức độ rộng lớn đặc biệt, thì liệu đó có phải là các bên tham gia có thể kìm giữ cuộc chiến ở mức giới hạn tại Đông Nam Á, sử dụng các loại vũ khí quy ước và nguyên tử, hay là trong một trận chiến chống lại một kẻ thù không biết mặt trên mặt trăng, sử dụng vũ khí vi trùng kỳ lạ? Đây là những câu hỏi quan trọng; chúng nằm ở chính tâm điểm của lý thuyết trò chơi; và chúng là những câu hỏi không thể có được câu trả lời tự tin mà không có chứng cứ thực nghiệm. Và không có gì để bàn cãi rằng những người chơi có lý trí có năng lực tri thức để vượt lên trên những chi tiết này của cuộc chơi và phớt lờ chúng; tính chất quan trọng của các chi tiết là ở chỗ chúng có thể cực kỳ hữu ích cho cả hai bên chơi và rằng những người chơi có lý trí biết rằng họ có thể bị lệ thuộc vào việc sử dụng những chi tiết này như những gợi ý trong quá trình đạt đến một giải pháp chung.

Một kết cuộc ổn định, hiệu quả sẽ dễ xuất hiện hơn giữa hai người chơi có cùng tính khí và nền tảng văn hóa, hay là giữa hai người chơi hoàn toàn khác biệt? Một kết cuộc ổn định, có hiệu quả sẽ dễ xuất hiện hơn giữa hai người chơi có kinh nghiệm thực hành, hay là giữa hai tay mơ, hay là giữa một tay lão luyện và một tay mơ; và trong cặp đấu cuối, ai sẽ có lợi thế hơn?

Trong một trò chơi thuộc loại này, các nước đi đầu tiên có tầm quan trọng như thế nào? Nếu những chuỗi hành vi ổn định, nghĩa là “luật chơi,” không được phát hiện sớm từ đầu, thì rốt cuộc liệu chúng có được phát hiện ra hay không? Liệu một cuộc chơi thành công cho cả hai phía sẽ dễ xuất hiện hơn nếu triết lý chung của mỗi bên bắt đầu bằng các quy tắc “chặt chẽ” hay bằng các loại vũ khí và nguồn lực cực kỳ “giới hạn,” chỉ được nới lỏng đôi chút khi thời cơ đòi hỏi, hay là nếu mỗi bên tham gia tự đưa ra cho mình những giới hạn rộng hơn ngay từ đầu để tránh phải nói dàn từng bước trong quá trình chơi?

Một “nhà trung gian” sẽ có mức ảnh hưởng như thế nào trên một trò chơi thuộc loại này, và những loại vai trò trung gian nào có hiệu quả nhất? Nếu nhà trung gian cũng có phần mình trong kết cuộc cuối cùng thì điều đó giúp ích hay cản trở cả hai bên tham gia trò chơi? Ở mức độ nào một nhà trung gian có thể phân biệt đối xử với một trong hai bên tham gia mà vẫn có thể tăng cường khả năng đưa đến một kết quả ổn định và có hiệu quả?

Sẽ là thú vị trong một trò chơi thuộc loại này khi bên tham gia đôi lúc đánh giá chính họ lẫn các đối tác của họ về những ván đòn như bên nào đang chơi hung hăng hơn hay hợp tác hơn, và “các quy tắc” nào mà một bên nghĩ rằng đang có hiệu lực và nghĩ rằng bên kia cũng nghĩ rằng đang có hiệu lực; bên nào “đang thắng” theo một ý nghĩa hai chiều (nhớ rằng rằng việc thiếu hiểu biết về hệ giá trị thực của nhau khiến cho việc đánh giá này luôn là một ván đòn mang tính suy đoán); khi nào thì cuộc chơi đã tiến đến một bước ngoặt “quan trọng,” hoặc khi nào thì một “sự đổi mới” trong chiến thuật được đưa

ra, hoặc khi nào thì một nước đi đặc thù của bên kia được giải thích là một “đòn trả đũa” hay là một sáng kiến mới.

Bởi vì một “điều luật trả thù” chủ yếu là *mang tính đạo đức* về mặt bản chất; bởi vì những ràng buộc cùng được [các bên] nhìn nhận trong bất cứ dạng nào của “chiến tranh giới hạn” chủ yếu được dựa trên một yếu tố mà về mặt *tâm lý* và về mặt xã hội rất giống với *truyền thống*; và bởi vì hệ giá trị đạo đức và truyền thống thường chẳng liên quan gì tới trò chơi đang có (chẳng hạn như, sự trả đũa hạt nhân theo cấp độ giữa Liên Xô và Hoa Kỳ trong khi cuộc chiến hạt nhân giới hạn được sử dụng ở châu Âu, hay việc đánh bom một số trường tiểu học trong một khu vực mà gần đây không có vấn đề bạo lực chủng tộc, hay việc đưa các hình thức cạnh tranh phi giá cả mới mẻ vào áp dụng trong một ngành công nghiệp cụ thể nào đó), cho nên có vẻ như là phần thực nghiệm trong lý thuyết trò chơi sẽ bao gồm cả phần công việc thí nghiệm như trong thí nghiệm của Muzafer Sherif. Ông này phát hiện ra rằng khi không có quy chuẩn nào tồn tại để đưa ra các phán quyết trong phòng thí nghiệm, thì các quy chuẩn sẽ được tạo ra bởi các chủ thể; và khi các quy chuẩn được tạo ra cho hai bên cùng có mặt trong một quy trình, thì việc phát triển quy chuẩn của mỗi bên sẽ có tác động ảnh hưởng đối với bên kia. Có một quá trình học hỏi chân thật về *những giá trị*; trong khi hình thành hệ giá trị cho riêng mình, mỗi bên sẽ điều chỉnh hệ giá trị của mình sao cho thích ứng với hệ thống của bên kia. Khi nguồn cung cấp các chuẩn mực “khách quan” có sẵn không còn đủ khả năng tạo ra một chuỗi hoàn chỉnh các quy tắc, nghĩa là, khi trò chơi trở nên “vô hạn định,” thì các quy chuẩn thuộc một thể loại nào đó phải được phát triển, được các bên

cùng nhận thức và chấp nhận; các mô hình hành động và đáp trả phải được chính đáng hóa.<sup>3</sup> Trong một hình thức hợp tác gần như là vô thức, các bên đối đầu phải đạt tới cùng một định nghĩa chung về việc thế nào là một động thái cách tân, một động thái thách thức hay khẳng định, hay một cử chỉ hợp tác, và họ phải phát triển một quy chuẩn chung nào đó liên quan đến hình thức trả đũa phù hợp với tội ác khi xảy ra việc vi phạm các quy tắc.<sup>4</sup>

---

<sup>3</sup> Một ví dụ đặc sắc về việc tạo ra các quy chuẩn trong thực hành - và cũng là ví dụ gợi ý rằng quy trình tạo quy chuẩn cần phải được phân tích - chính là việc thừa nhận rộng rãi trong các cuộc thương thảo về giải trừ quân bị hồi năm 1957 rằng bất cứ vùng thanh sát nào cuối cùng được thống nhất cũng sẽ phải được lựa chọn từ trong một dãy các vùng có hình miếng bánh táo với đỉnh nằm ở Bắc Cực.

<sup>4</sup> Với tư cách là một chuyên gia về lý thuyết trò chơi, bạn có thể hy vọng rằng người ta có thể vẽ một đường phân định rõ ràng giữa tâm lý học thực nghiệm thuộc lý thuyết trò chơi và phần còn lại của môn tâm lý học xã hội; [dù gi] nó vẫn được coi là một lý thuyết *chiến lược*, chứ không phải toàn bộ lãnh địa của hành vi xung đột. Nhưng vẫn không rõ ràng là phải bắt đầu vẽ trước đường phân chia từ đâu. Ví dụ như "sự thù nghịch" dường như có thể là một tố chất thuộc về tâm lý hay tính cách được kim giữ ngoài phạm vi lý thuyết trò chơi, nhưng nếu sự thù nghịch của một người chơi trong trò chơi là một hạn chế đáng kể đối với khả năng của anh ta trong việc tiếp nhận ý định của đối thủ, thì nó trở thành một phần của "cấu trúc trao đổi thông tin." Một thí nghiệm của Deutsch thích hợp trong trường hợp này. Ông ta để cho từng cặp hai người chơi tham gia các trò chơi có tổng lợi ích khác không (theo hình thức ma trận) theo từng chuỗi hai bước chơi, trò chơi cho phép lựa chọn cả cách chơi "hợp tác" và "bất hợp tác." Những người chơi theo hình thức bất hợp tác với một người chơi hợp tác có một cơ hội, trong bước chơi thứ hai, đáp trả lại đề nghị hợp tác ngầm. Nhưng, "khi kỳ vọng của họ về sự lựa chọn của người kia không được khẳng định, họ có xu hướng giải thích sự lựa chọn của đối thủ như một biểu hiện của sự thờ ơ, hoặc như việc thiếu hiểu biết cản bản về cách "nên" chơi như thế nào... Trong nhóm này, việc hiểu biết về lựa chọn của đối thủ, do ý nghĩa được gắn kết với lựa chọn đó, có xu hướng tăng cường những tình cảm tiêu cực trước đó liên quan đến những ý định của người kia."

Một “kịch bản” có thể, chẳng hạn như, xác định rằng một trong các bên là “kẻ gây hấn”; nó có thể là các kết quả của những nước đi trước đó trong cùng một trò chơi bởi những người chơi khác; nó có thể là một câu chuyện nền khẳng định dạng một hình thức phân định địa lý nào đó tuân theo tình trạng ban đầu; hoặc nó có thể gắn một tuyên bố đạo đức nào đó của một người chơi với một phần cụ thể nào đó trên bàn cờ. Những dữ liệu nền này sẽ không có ảnh hưởng nào đối với cấu trúc logic hay cấu trúc toán học của trò chơi; chúng sẽ được sử dụng không ngoài sức mạnh nào khác ngoài khả năng thăm dò gợi ý. Một lần nữa, người ta có thể đặt lại ván cờ sao cho trong nước đi đầu tiên nó phù hợp với cách nó đang ở giữa ván cờ mà trước đó hai người đã chơi, và xem xét liệu kết quả có bị ảnh hưởng bởi việc thông báo cho những người chơi biết việc sắp xếp những bước đi đầu tiên trong ván cờ trước đó như thế nào không. Nếu những người chơi có xu hướng phát triển “những quy chuẩn” dựa trên cấu hình tinh của cục chơi như họ nhận định về nó ngay từ đầu, thì rất có thể các quy chuẩn đó sẽ bị bóp méo nếu ta cung cấp, thông qua một cách thức hoàn toàn “không mang tính áp đặt”, một câu chuyện nền giúp chỉ dẫn – một cách gợi ý – một điểm khởi đầu khác dựa trên một giả thuyết khác nào đó.<sup>5</sup>

---

Xem Morton Deutsch, *Conditions Affecting Cooperation* (Điều kiện ảnh hưởng đến sự hợp tác), Trung tâm Nghiên cứu Quan hệ Nhân sự, New York University, 1957. (Một bài báo dựa trên chuyên khảo này, không bao gồm điểm được trích dẫn ở đây, có tựa đề “Niềm tin và sự Nghi ngờ,” (Trust and Suspicion) được đăng tải trên *The Journal of Conflict Resolution*, 2:265-279 [tháng 12-1958].)

<sup>5</sup> Các câu hỏi về thuế thu nhập mô tả trong Chương 3 (các trang 62-65) cho thấy sức mạnh của sự gợi ý.

Có lẽ sẽ thú vị khi xem xét liệu mỗi người chơi có thể thật sự nhận biết được khi người kia đang “kiểm tra” sự quyết đoán của anh ta, “đe dọa” anh ta, và vân vân; và có lẽ có thể nghiên cứu quy trình theo đó từng bước đổi đầu cù thể được khoác lên một tầm quan trọng tượng trưng, chẳng hạn như mỗi người chơi nhận ra rằng anh ta đang thiết lập một vai trò và uy tín thông qua cách anh ta hành xử tại mỗi thời điểm cù thể trong ván cờ.

Một chiêu kích khác của cuộc chơi có vẻ đáng được nghiên cứu chính là ý nghĩa của *tiệm tiến* (incrementalism) gắn liền với những nước đi và hệ thống giá trị. Ví dụ, hãy xem xét một trò chơi bao gồm việc di chuyển các quân cờ trên một bàn cờ hay việc di chuyển binh lính trên trận địa. Nếu những người chơi phải chơi theo lượt, mỗi lần di chuyển một quân một ô, trò chơi sẽ tiến triển chậm theo các mức gia tăng nhỏ; tình hình trên bàn cờ có thể thay đổi trong quá trình chơi, nhưng điều đó được thực hiện bởi một chuỗi những thay đổi nhỏ mà có thể quan sát được, đánh giá được, và điều chỉnh theo được, với nhiều thời gian cho phép các sai lầm cá nhân hay các sai lầm chung - vốn có tác dụng phá hủy giá trị của cả hai bên - có thể được quan sát, điều chỉnh, và tránh được trong bước chơi kế tiếp. Nếu có sự liên lạc, sẽ có thời gian cho những người chơi mặc cả bằng ngôn ngữ trực tiếp và tránh những nước đi tàn phá đối với cả hai bên. Nhưng thay vào đó giả sử rằng [các bên] có thể di chuyển nhiều quân cờ cùng một lúc theo bất cứ hướng nào và bất cứ khoảng cách bất định nào, và rằng các quy tắc chơi khiến cho kết quả của một hành vi thù địch có sức tàn phá ghê gớm với một bên hoặc cả hai bên. Lúc này trò chơi không còn mang nhiều tính chất tiệm tiến; sự

việc có thể diễn ra bất ngờ. Có thể có xu hướng muốn tấn công bất ngờ. Trong khi một bên chơi có thể nắm bắt được tình hình ở vào một thời điểm cụ thể, anh ta không thể dự đoán kết quả trước nhiều hơn một hoặc hai bước đi. Có ít cơ hội để phát triển một tạm ước, hoặc truyền thống tin tưởng, hoặc các vai trò khống chế và quy phục của những người chơi, bởi vì tốc độ của trò chơi khiến cho tình hình diễn tiến quá nhanh, trước khi người ta có thời gian để rút ra kinh nghiệm hoặc rút tỉa được sự hiểu biết cần thiết. Nhưng liệu một trò chơi tiệm tiến hơn sẽ giúp các bên chơi dễ phối hợp thành công hơn hay không, hay nó lại dẫn đến một cách chơi liều lĩnh hơn? Hay là điều này lệ thuộc vào loại người chơi và loại gợi ý mà chúng ta đưa vào trong chính trò chơi? Liệu yếu tố then chốt có phải tính tiệm tiến của *các nước đi* trong trò chơi, hay là tính tiệm tiến trong các hệ giá trị của người chơi (ở đây là sự gia tăng trong hệ thống tính điểm)? Hay liệu những vấn đề này có thể trở nên tương hợp, bù đắp cho nhau, để nếu đưa tính tiệm tiến vào phía này thì có thể bù đắp cho việc thiếu tính tiệm tiến ở phía kia?? Tính chất thích đáng của những câu hỏi này được kiểm nghiệm bởi cuộc tranh luận về vai trò của vũ khí hạt nhân trong chiến tranh giới hạn, ý nghĩa của ước muốn thực hiện cuộc tấn công phủ đầu trong một tình huống phụ thuộc vào sự ngăn chặn lẫn nhau, và nhiều đề nghị khác nhau nhằm giảm bớt tốc độ của chiến tranh hiện đại và cô lập nó về mặt địa lý, cùng với sự bất đồng về việc liệu có thể tồn tại một sự việc như một cuộc chiến giới hạn trên lục địa Tây Âu hay không. Tính tiệm tiến có thể được dùng để so sánh trong các phân tích hình thức, một khi các chuẩn mực thực nghiệm cần thiết

đã được xác định thông qua thí nghiệm hoặc quan sát.<sup>6</sup>

Các câu hỏi này đều dành cho các trò chơi hai người, ngoại trừ vai trò có thể của người trung gian. Những trò chơi tương tự có thể được chơi bởi ba hoặc nhiều người hơn, mỗi người theo cách riêng mình; và tác giả phỏng đoán rằng - ít nhất trong số những người chơi "thành công" - nhiều kết quả thực nghiệm sẽ dễ xuất hiện hơn khi càng có nhiều người chơi. Nói ở mức chung hơn, hình thức phối hợp được dùng trong việc hình thành các nhóm hay liên minh có thể trở thành các chủ đề của nghiên cứu thực nghiệm. Khác với lối tiếp cận đơn giản và đối xứng đôi khi vẫn được dùng trong lý thuyết trò chơi để nghiên cứu việc hình thành các liên minh, có thể sẽ thú vị hơn nếu trong nghiên cứu về việc hình thành nhóm, ta đưa ra giả định về một vài yếu tố bất đối xứng nào đó, các tiền lệ, trật tự nước đi, các cấu trúc thông tin không hoàn hảo, và các chi tiết theo nghĩa rộng khác. Chắc chắn ảnh hưởng đối với việc hình thành các liên minh trong các bối cảnh thông tin bất đối xứng, và do đó bất hoàn hảo, thường dẫn tới việc phải nghiên cứu thực nghiệm một cách có hệ thống.<sup>7</sup>

<sup>6</sup> "Chiến tranh giới hạn không chỉ phải tìm cách ngăn chặn hình thức bạo lực nhất của nó; mà còn phải tìm cách làm chậm lại nhịp độ của chiến tranh hiện đại, nếu không, tốc độ chóng mặt của các hoạt động quân sự nối tiếp nhau sẽ cản trở việc thiết lập mối quan hệ giữa các mục tiêu chính trị và quân sự. Nếu mối quan hệ này bị mất đi, bắt cứ cuộc chiến nào cũng có thể phát triển qua các giai đoạn không thể tiên liệu trước để dẫn đến một cuộc chiến tổng lực." (Henry A. Kissinger, *Nuclear Weapons and Foreign Policy* (Vũ khí Hạt nhân và Chính sách đối ngoại), [New York, 1957].)

<sup>7</sup> Alex Bavelas đã mô tả một thí nghiệm về sự phối hợp thuận tay, trong đó mỗi một người trong số năm người chơi riêng biệt phải di chuyển qua lại các quân cờ hình học cho đến khi họ đạt được sự phân phối các quân cờ cho phép hình thành năm ô vuông riêng biệt. Các quân cờ được

---

cắt sao cho nhiều hình vuông "sai" có thể được hình thành, nghĩa là các hình vuông sử dụng việc kết hợp các quân cờ khiến cho không thể xếp thêm bốn hình vuông khác với những quân cờ còn lại. Ông quan tâm đến điều xảy ra khi "những thành công" giả tạo này diễn ra. "Đối với một cá nhân sau khi đã hoàn tất được một hình vuông thì một cách dễ hiểu là thật khó khăn khi phá bỏ hình vuông đó. Anh ta chỉ có thể hành động 'xa rời mục tiêu' với sự thoái mái khi điều đó ở một mức độ nào đó lệ thuộc vào nhận thức của anh ta về cái toàn cảnh. Về mặt này, mô hình trao đổi thông tin có lẽ có những ảnh hưởng được xác định rõ ràng... Các bước chơi ban đầu... đã bộc lộ... ra rằng các lực bắt buộc chống lại việc tái cấu trúc là rất lớn, và rằng, với một sức hạn chế trao đổi thông tin đáng kể, thi khó mà có được một giải pháp" ("Các mô hình truyền thông trong các nhóm phân theo công tác" (Communication Patterns in Task-oriented Groups), trong D. Cartwright và A. F. Zander, *Group Dynamics* (Sự năng động nhóm) [Evanston, 1953], trang 493). Một số công trình mang tính gợi ý cao, đặc biệt là về "nhận thức thiên vị về những gì bất biến," được tường thuật bởi Charles E. Osgood, *Journal of Conflict Resolution*, 3:304-05 (tháng 12-1959). truyền thông, thi khó mà có được một giải pháp" ("Các mô hình truyền thông trong các nhóm phân theo công tác" (Communication Patterns in Task-oriented Groups), trong D. Cartwright và A. F. Zander, *Group Dynamics* (Sự năng động nhóm) [Evanston, 1953], trang 493). Một số công trình mang tính gợi ý cao, đặc biệt là về "nhận thức thiên vị về những gì bất biến", được tường thuật bởi Charles E. Osgood, *Journal of Conflict Resolution*, 3:304-05 (tháng 12-1959).

## **PHẦN III**

# **CHIẾN LƯỢC CÓ THÀNH PHẦN NGẪU NHIÊN**



## NGẦU NHIÊN HÓA HÚA HẸN VÀ ĐE DỌA

Trong lý thuyết về trò chơi xung đột thuần túy (các trò chơi có tổng lợi ích bằng không), các chiến lược ngẫu nhiên hóa đóng một vai trò trung tâm. Không quá lời khi nói rằng những tiềm năng của hành vi ngẫu nhiên hóa chiếm phần lớn sự quan tâm trong lý thuyết trò chơi trong suốt hơn một thập niên rưỡi qua.<sup>1</sup> Cốt lõi của tính chất ngẫu nhiên hóa trong một trò chơi có tổng lợi ích bằng không có hai người chơi là việc ngăn chặn không cho đối phương nắm bắt được cách chơi riêng của một người - để ngăn ngừa việc dự đoán trước mang tính suy diễn của anh ta về việc người kia có thể quyết định

---

<sup>1</sup> John von Neumann, khi phát biểu về "định lý căn bản về sự tồn tại của các chiến lược tốt", nghĩa là định lý cho rằng tất cả các trò chơi có tổng lợi ích bằng không với một con số giới hạn các chiến lược thuần túy có một cặp cân bằng tối thiểu hóa tối đa - tối da hóa tối thiểu ("đáp số") nếu các chiến lược hỗn hợp được cho phép sử dụng, nói: "Ở phạm vi mà tôi biết được, không thể có lý thuyết trò chơi trên những cơ sở này nếu không có định lý đó... Trong suốt quá trình tìm hiểu tôi đã nghĩ rằng chẳng có gì đáng để xuất bản cho đến khi 'định lý tối thiểu hóa tối đa' được chứng minh." ("Communication on the Borel Notes", *Econometrica*, 21:124-125 [tháng 1-1953]).

như thế nào, và để bảo vệ chính mình khỏi hành vi làm lộ tẩy bất thường mà kẻ thù có thể nhận ra hoặc để tránh cho một bên khỏi xu hướng khinh suất trong việc chọn lựa khiến đối phương có thể đoán ra. Tuy nhiên, trong các trò chơi hỗn hợp giữa xung đột và lợi ích chung, tính chất ngẫu nhiên không đóng một vai trò trọng tâm như thế, và vai trò mà nó thực sự nắm giữ thì khá khác biệt.<sup>2</sup>

Việc ngẫu nhiên hóa trong lý thuyết của các trò chơi (có tổng lợi ích khác không) này không liên quan chủ yếu đến việc ngăn chặn không để chiến lược của một bên bị bên kia dự đoán được. Trong những trò chơi này, như đã lưu ý ở trên, một bên thường quan tâm nhiều hơn đến việc làm cho bên kia dự đoán về cách chơi của mình, và dự đoán nó một cách chính xác, hơn là ngụy trang chiến lược của mình.

---

<sup>2</sup> Thay vào đó, người ta có thể giải thích các chiến lược hỗn hợp trong các trò chơi có tổng lợi ích bằng không như một phương tiện để đưa tính liên tục của các chiến lược vào một trò chơi chiến-lược-riêng-rẽ mà không có điểm tựa là một chiến lược thuần túy, từ đó biến nó thành một trò chơi có một điểm tựa. Trong cách giải thích này thì vai trò của các chiến lược hỗn hợp trong các trò chơi có tổng lợi ích bằng không không khác nhiều so với vai trò của chúng trong các trò chơi có tổng lợi ích khác không. Người ta có thể tung một đồng xu để đối thủ của mình không thể đoán chắc được là sấp hay ngửa; hoặc người ta có thể tung một đồng xu để “trung bình hóa” sấp và ngửa, nhằm tạo ra (theo một ý nghĩa có giá trị được kỳ vọng) một chiến lược nửa vời giữa sấp và ngửa. Cả hai cách giải thích đều hữu ích. Nếu cách thứ hai có vẻ phức tạp hơn chút ít, thì cách thứ nhất có thể nắm bắt được tinh thần của vấn đề bởi vì nó tự bộc lộ với người chơi. Và cách thứ nhất khiến chúng ta nhớ ra rằng vấn đề, ngay cả trong trường hợp ngẫu nhiên hóa, vẫn là nhằm ngăn chặn khả năng phán đoán của đối phương về sự chọn lựa chiến lược thật sự của một người chơi, và rằng cơ cấu chọn lựa, các quy trình ghi nhận và thông tin sự chọn lựa, và bắt cứ sự chuẩn bị trước nào cần phải có theo đòi hỏi của quá trình ngẫu nhiên, đều phải làm sao để không thể tiếp cận đối với hệ thống thông tin của người đó.

Dĩ nhiên là có những thành tố kết quả bằng không cài đặt trong một trò chơi lớn hơn. Trong chiến tranh giới hạn, người chơi có thể quan tâm đến việc trao đổi thông tin hơn là ngụy trang các giới hạn đưa ra để bên kia quan sát, nhưng trong phạm vi những giới hạn đó có thể cho xuất kích máy bay của mình theo một cách ngẫu nhiên nhằm làm giảm thiểu thông tin tình báo chiến thuật của kẻ thù.<sup>3</sup> Một lần nữa, các thông tin có thể được trao đổi, hoặc các thỏa thuận được áp dụng trên cơ sở mẫu, trong đó không bên nào có thể cung cấp cho bên kia đầy đủ thông tin. Chẳng hạn như các hiệp ước kiểm soát vũ khí có thể được giám sát bởi kỹ thuật lấy mẫu cho phép mỗi bên có thông tin vừa đủ để biết về lực lượng của đối phương bộc lộ sự tuân thủ hay không tuân thủ mà không phải cung cấp quá nhiều khiến cho khả năng tiến hành một cuộc tấn công bất ngờ thành công được nâng cao hơn nhiều.

Nhưng vai trò chính của tính ngẫu nhiên theo nghĩa truyền thống về các trò chơi có tổng lợi ích khác không lại là một vai

---

<sup>3</sup> Trong các trường hợp cụ thể có thể có một nghịch lý khó chịu nằm sẵn trong một sự lựa chọn cách thức bí mật hay cách thức bộc lộ. Nếu để minh chứng rằng một bên cam kết đe dọa, hay một bên trong thực tế có khả năng thực hiện sự đe dọa, thì bên đó phải thể hiện chứng cứ về cam kết hay năng lực của mình cho bên kia, chứng cứ có thể thuộc loại nhất thiết đưa ra thông tin có ích cho bên thứ hai trong việc đối phó với mối đe dọa. Để chứng minh cho kẻ thù biết rằng ta có vũ khí tiềm năng có thể vượt qua được hàng phòng thủ của kẻ thù, chúng ta có thể phải biểu lộ vũ khí đó hoặc một vài khía cạnh của vũ khí đó, hoặc cung cấp kiến thức kỹ thuật để chứng minh rằng vũ khí đó khả thi. Làm như thế sẽ giúp cho bên kia rất nhiều trong việc chuẩn bị một lực lượng phòng thủ chống lại vũ khí đó. Nếu, để chứng minh chúng ta sẽ tiến hành một cuộc chiến tranh cục bộ trong một khu vực không rõ ràng, thì điều cần thiết là phải bố trí quân đội ở nơi đó trước, và kẻ thù sẽ có điều kiện lợi thế biết trước được địa điểm chính xác thay vì phải chuẩn bị ở mọi hướng.

trò khác. Nó là một công cụ khiến cho các mục tiêu không thể phân chia trở nên có thể phân chia, hay những mục tiêu bất tương ứng trở nên thuần nhất. “Các giá trị kỳ vọng” của chúng có thể được phân chia bởi tính may rủi ngẫu nhiên trong khi bản thân mục tiêu lại không thể. Chúng ta tung đồng tiền để xem ai giành được mục tiêu, và chơi “được gấp đôi hay mất hết” khi chúng ta không thể tạo ra sự thay đổi. Chúng ta có thể phân chia nghĩa vụ công dân một cách bình đẳng bằng cách lựa chọn những người đăng ký quân sự thông qua rút thăm may rủi, khi mà chúng ta muốn chỉ một phần trong số những người đủ điều kiện tham gia phục vụ một thời gian dài thay vì huy động tất cả mọi người phục vụ một thời gian ngắn.

Trong vai trò này, tính ngẫu nhiên rõ ràng là có liên quan đến sự hứa hẹn. Nếu những hứa hẹn ưu đãi duy nhất có được lớn hơn mức cần thiết nhưng không thể phân chia được, thi một hình thức rút thăm may rủi đưa ra một khả năng cụ thể của việc chuyển giao ưu đãi có thể làm giảm bớt quy mô giá trị kỳ vọng của lời hứa và làm giảm bớt chi phí cho người đưa ra lời hứa đó. Một đề nghị giúp một ai đó một khoản ưu đãi lớn theo hình thức may rủi ngẫu nhiên thì cũng tương đương với việc chắc chắn giúp một khoản ưu đãi nhỏ. (Có thể có một lợi thế bổ sung là tính may rủi ngẫu nhiên có tương quan với nhu cầu của anh ta).

Nhưng trong khía cạnh này, sự hứa hẹn khác với sự đe dọa. Khác biệt là ở chỗ sự hứa hẹn sẽ khiến phải trả giá khi nó thành công, còn sự đe dọa sẽ khiến phải trả giá khi nó thất bại. Một sự đe dọa thành công là sự đe dọa không được thực

hiện. Nếu tôi hứa nhiều hơn mức tôi cần như là một cách hấp dẫn đối tượng, và nếu lời hứa thành công, tôi phải trả nhiều hơn mức tôi cần. Nhưng một mối đe dọa “quá lớn” thì có khả năng là vô dụng thay vì tốn kém. Nếu tôi đe dọa cho nổ tung cả hai chúng ta thành từng mảnh khi mà chỉ cần đe dọa gây khó chịu thôi là đủ, bạn vẫn sẽ tuân thủ; bởi vì tôi sẽ không gây khó chịu mà cũng không tàn sát, và sai lầm không gây ra tốn kém nào. Nếu tất cả những gì tôi có là một trái lựu đạn để cho nổ giữa chúng ta và ước gì nó là một trái lựu đạn cay, tôi có thể giảm cấp trái lựu đạn xuống “mức” còn là một trái bom hơi cay bằng cách đe dọa một cơ hội phần trăm thích hợp nào đó rằng trái bom sẽ phát nổ, khiến cả hai cùng chết nếu bạn không tuân thủ. Nhưng nhu cầu thực hiện điều này thì không rõ ràng như trong trường hợp một lời hứa, khi mà bất cứ sự thái quá nào trong giá trị được hứa hẹn đều là sự mất mát.

Quy mô của sự đe dọa có thể là một vấn đề nếu nó đòi hỏi chi phí phải có cho điều gì đó đi kèm để *tạo ra* sự đe dọa và nếu tạo ra sự đe dọa lớn hơn thì tốn kém hơn tạo ra những đe dọa nhỏ. Nếu sự đe dọa hơi cay là đủ, thì tôi không cần đe dọa cho nổ, và nếu các trái bom hơi cay ít tốn kém hơn những thiết bị nổ, và nếu tôi phải trưng bày trái bom để cho sự đe dọa thuyết phục, thì cách tốt hơn là đe dọa bằng hơi cay rẻ tiền hơn. Nhưng lựu đạn có thể rẻ hơn, và do vậy động cơ đã đi theo hướng khác. Đối với nhiều sự đe dọa lý thú thì chi phí lớn nhất là nguy cơ phải thực hiện lời đe dọa, và “chi phí” thông thường hơn thì không phải là yếu tố kiểm soát.

## NGUY CƠ THẤT BẠI

Tuy nhiên, nguy cơ thất bại thật sự tạo ra động cơ để chọn các hình thức đe dọa dung hòa hơn là cực đoan. Nếu mỗi đe dọa duy nhất có thể đưa ra là một hành động khủng khiếp nào đó, người ta có thể có xu hướng làm giảm nhẹ nó bằng cách gắn nó với một phương tiện phân chia may rủi - bằng cách đe dọa một *khả năng cụ thể* nào đó mà sẽ được tiến hành trừ phi sắp tới có sự tuân thủ, chứ không phải cam kết chắc chắn rằng biện pháp trừng phạt gây thiệt hại cho cả hai bên sẽ được áp dụng.

	I	II
i	0	1
ii	1	0
	0	$-x$
	0	$-y$

H. 17

Để minh họa, hãy xem xét ma trận trong Hình 17, trong đó Cột có sự lựa chọn trước, tiếp theo là Hàng, nhưng trong đó Hàng có sự chọn lựa trong việc đưa ra đe dọa trước để hạn chế sự chọn lựa của Cột (xem  $X$  và  $Y$  là những số dương). Ở một điều kiện, chiến lược của Hàng rõ ràng là để đe dọa hàng ii nếu Cột chọn cột II. Nếu Hàng không đưa ra đe dọa, Cột chọn II vì biết rằng Hàng sau đó sẽ chọn i. Với mỗi đe dọa - và giả sử rằng Hàng cam kết thực hiện đe dọa đó và Cột biết

điều đó – việc lựa chọn II đưa đến các kết cuộc không hấp dẫn đối với cả hai, và Cột có thể được kỳ vọng sẽ chọn I.

Điều kiện là Hàng hoàn toàn chắc chắn rằng không có gì đi chệch đường! Có thể nó hoàn toàn phán đoán sai những bước đi cuối cùng của Cột. Một cách lựa chọn, Hàng có thể cam kết thực hiện đe dọa của mình nhưng không thành công trong việc truyền đạt đe dọa đó một cách thuyết phục cho Cột, khiến cho Cột mắc sai lầm không biết đến mối đe dọa, đẩy cả hai xuống ô bên phải ở bên dưới. Trong khi đó, về phía mình Cột có thể sắp xếp một cam kết trước thông qua việc chọn II, và không thành công trong việc truyền đạt cam kết đó một cách chính xác cho Hàng kịp thời để Hàng có thể cân nhắc tính toán, hoặc Cột có thể gặp phải tình huống mất năng lực mà Hàng không biết, làm triệt tiêu khả năng xảy ra của I. Trong trường hợp đó, cam kết của Hàng sẽ chỉ đảm bảo rằng kết cuộc tồi tệ sẽ xảy ra cho cả hai bên. Sự thất bại cho dù vì bất cứ lý do gì đi nữa sẽ khiến cho có thể xảy ra khả năng là sự đe dọa sẽ thất bại. Nếu chúng ta đưa điều đó ra xem xét, thì chúng ta có thể có một lý do để cho Hàng mong ước rằng phải chi các “thiệt hại” lợi ích trong ô bên phải ở dưới không quá khó ưa như chúng thực sự diễn ra.

Nếu Hàng bị ràng buộc vào các chiến lược “thuần túy” – nếu nó phải cụ thể hóa sự đe dọa hay cam kết của mình mà không xét đến sai lầm hay cơ hội – nó chẳng thể làm gì ngoài việc mong ước rằng những con số trong ô bên phải ở dưới không quá khó chịu. Nhưng nếu nó có thể ngẫu nhiên hóa sự đe dọa của mình thì thực ra nó có thể “làm giảm cấp độ” để giảm bớt chi phí cao của thất bại. Nếu, chẳng hạn như, nó có thể không chọn lựa hàng ii trong trường hợp cột II được lựa

chọn, nhưng có cơ hội 50-50 giữa i và ii trong trường hợp đó, nó vẫn hy vọng có thể đe dọa Cột phải lựa chọn I trong khi làm giảm bớt tính chất nghiêm trọng của nguy cơ thất bại.

Chúng ta có thể đề cập cụ thể hơn. Xem  $P$  là khả năng sự đe dọa sẽ thất bại vì bất cứ lý do gì. (Đối với mục đích hiện tại của chúng ta, đây là một khả năng “độc lập”, không phụ thuộc vào chiến lược của Hàng). Hãy để cho Hàng đe dọa sử dụng ii với khả năng tương đương với  $\pi$ , trong trường hợp Cột chọn II. Nói cách khác, nếu Cột không tuân thủ thì có khả năng xảy ra  $\pi$  là Hàng sẽ lựa chọn ii để gây ra tình trạng khó chịu cho cả hai bên, và khả năng  $(I - \pi)$  mà nó sẽ lựa chọn I để tạo ra sự dễ chịu cho hai bên. Vậy thì Hàng sẽ lựa chọn giá trị nào của  $\pi$ ?

Trước hết,  $\pi$  phải lớn như thế nào để làm cho mỗi đe dọa có hiệu quả thật sự, nghĩa là làm cho nó hiệu quả trên cơ sở giả thiết rằng nó *không* thất bại vì bất cứ lý do độc lập gì bao hàm trong  $P$ ? Đây là một vấn đề thuộc sự lựa chọn của Cột khi nó phải đối mặt với nguy cơ  $\pi$ . Nếu Cột chọn I thì nó sẽ có được 0. Nếu chọn II, thì kỳ vọng của nó là một trung bình cân bằng giữa I và  $-X$ , với tác động lần lượt của  $(I - \pi)$  và  $\pi$ . Nếu mức trung bình này nhỏ hơn 0, nó sẽ có động cơ chọn I – tùy thuộc vào khả năng độc lập,  $P$ , mà vì lý do này hay lý do khác nó sẽ chọn II mặc dù động cơ hiển nhiên của nó hướng về I. Điều kiện cho một mối đe dọa hiệu quả như sau.<sup>4</sup>

---

<sup>4</sup> Bởi vì việc phân tích phụ thuộc duy nhất vào so sánh về các *sự khác biệt* giữa giá trị tuyệt đối của các lợi ích dành cho từng người trong hai người chơi, không có ép buộc việc sử dụng một thước đo đặt lợi ích ưu tiên của anh ta ở +1 và lợi ích ưu tiên kế tiếp ở 0. Do vậy, việc giải thích đầy đủ, của cách giải thích  $I/(I + X)$ , là: tỷ trọng giữa (1) sự khác biệt

$$0 > (I - \pi) - \pi X,$$

$$\pi > \frac{I}{I + X}$$

Thứ hai, giả sử rằng bất cứ sự đe dọa nào với  $\pi$  ở trên mức sàn được thiết lập bằng công thức trên sẽ thành công hay thất bại với các khả năng lần lượt là  $(I - P)$  và  $P$ . Nếu sự đe dọa thành công, phần thưởng của Hàng là  $+ I$ . Nếu thất bại, kỳ vọng của nó là một trung bình cân bằng giữa  $o$  và  $-Y$ , với tác động lần lượt là  $(I - \pi)$  và  $\pi$ . Thê thì, giá trị kỳ vọng của kết quả khi mối đe dọa là đủ lớn để tạo ra hiệu quả sẽ được tính theo công thức

$$(I - P) + P(o - \pi Y) = I - P - P\pi Y.$$

Giá trị này rõ ràng là cao hơn, cái thấp hơn là giá trị của  $\pi$ . Hàng do đó nên sắp xếp giá trị thấp nhất của  $\pi$  ở mức mà

---

giữa các lợi ích bên phải phía trên và bên trái phía trên của Cột với (2) những sự khác biệt giữa (a) thu hoạch phía trên bên phải và phía trên bên trái của nó và (b) các thu hoạch phía dưới bên phải và phái trên bên trái. Tính chất đơn giản của công thức, do đó, phản ánh lợi thế đã giành được từ tiện ích phân hạn này. Chỉ cần một thông số để xác định tính chất của các mối quan hệ giữa ba cách đánh giá này. (Trong một vấn đề tiếp theo sau bao hàm ô bên dưới phía trái, tất cả bốn lợi ích đều thích đáng và sẽ cần phải có một thông số thứ hai. Tuy nhiên, trường hợp đó có thể được đơn giản hóa hơn nữa nếu lợi ích bên dưới phía trái có thể được xem tương đương với một trong những lợi ích còn lại và vẫn minh họa được cho chủ điểm; chúng ta có được sự hiểu biết ít hoàn chỉnh hơn nhưng nhiều I và o hơn qua cách đó). Về việc giải thích những con số này, xin xem A. A. Alchian, "Ý nghĩa của việc đo lường giá trị sử dụng" (The Meaning of Utility Measurement), *American Economic Review*, 43: 26-50 (tháng 3-1953), hay Luce và Raiffa, trang 12-38.

nó có thể để đáp ứng điều kiện đầu tiên. Đối với một mối đe dọa đáng kể nào đó - có một giá trị kỳ vọng lớn hơn zero, là điều mà Hàng có thể kỳ vọng từ ma trận đặc biệt này nếu nó không đưa ra sự đe dọa nào - một giá trị  $\pi$  phải được sắp xếp sao cho đáp ứng được điều kiện

$$I - P - P\pi Y > 0$$

hoặc  $\frac{I - P}{P} \cdot \frac{I}{Y} > \pi$

Như vậy, phạm vi hiệu quả của  $\pi$  trong ví dụ này được tính theo

$$\frac{I - P}{P} \cdot \frac{I}{Y} > \pi \frac{I}{I + X}$$

Và không có mối đe dọa nào cần đưa ra nếu không tồn tại khoảng trống giữa hai giới hạn này, nếu

hay  $\frac{I - P}{PY} < \frac{I}{I + X}$

$$\frac{P}{I - P} > \frac{X + I}{Y}$$

Chỉ có một mối đe dọa “rất nhỏ” - một mối đe dọa với  $\pi$  nhỏ hơn  $I - P$  - đáng được đưa ra nếu:

hay  $\frac{I - P}{PY} < 1$       hay  $\frac{P}{I - P} > \frac{I}{Y}$

Sau đây là một trường hợp trong đó mối đe dọa có thể chia nhỏ được có giá trị hơn một mối đe dọa bất di bất dịch, và trong đó mối đe dọa bất di bất dịch là không đáng đưa ra

trong khi mối đe dọa có thể phân nhỏ được là đáng. Cuộc tranh cãi xoay quanh nguy cơ thất bại, một nguy cơ đã được giả thiết là độc lập với quy mô của chính  $\pi$ . Đây là một giả thiết tương đối đặc biệt. Nếu chúng ta giải thích  $P$  là khả năng chúng ta đánh giá sai về kẻ thù của mình và cường điệu mong muốn của kẻ thù muốn tránh ô bên dưới phía bên phải, thì giả thiết của chúng ta ngũ ý một sự phân phối lợi ích nước đôi. Điều đó ngũ ý rằng chúng ta có hoặc một người mà lợi ích của anh ta được thể hiện một cách tương xứng qua những con số trong ma trận của chúng ta, hoặc một người mà lợi ích của anh ta rất khác biệt đến mức không có mối đe dọa liên quan - trong phạm vi các giá trị lên đến mức  $\pi = 1$  - có thể can ngăn được anh ta. Nếu thay vào đó chúng ta giả thiết rằng tỷ lệ các cột lợi ích trong các ô bên trên và bên dưới phía bên phải thể hiện một hình thức phân phối tần suất theo hình chuông, và rằng đối phương cụ thể của chúng ta được rút ra một cách ngẫu nhiên, thì khả năng sự đe dọa của chúng ta thành công hay không sẽ thay đổi một cách trực tiếp theo giá trị của chính  $\pi$ . Khả năng một tên trộm được chọn ra ngẫu nhiên từ giới trộm cắp sẽ bị ngăn chặn bởi khả năng cụ thể nào đó của sự bắt giữ và sự kết tội, về mặt giả thiết sẽ thay đổi một cách trực tiếp theo khả năng thứ hai; mô hình đơn giản phân tích ở trên xem xét những tên trộm như có thể phân chia được thành hai nhóm - những đối tượng ăn cắp tiền và chắc chắn bị ngăn chặn theo những con số trong ma trận, và những đối tượng ăn cắp tiền cho vui và nằm bên ngoài khả năng tác động của bất cứ sự đe dọa nào ở mức độ thể hiện trong ô bên dưới phía bên phải. Mặt khác, nếu khả năng thất bại của chúng ta phản ánh sự thất bại trong việc liên lạc với kẻ thù chẳng

hạn, có thể có lý do tốt hơn để giả thiết rằng khả năng thất bại là độc lập với mối đe dọa được truyền đạt.

Một điểm thú vị đáng ghi nhận là việc gắn kết khả năng thực hiện mối đe dọa, theo như mô hình trên, cũng đồng nghĩa đáng kể với việc hạ cấp mức độ đe dọa một cách trực tiếp hơn. Để hiểu điều này, hãy xem  $X$  trong ô phía dưới phía bên phải như một khoản phạt đánh trên cả Hàng lắn Cột, hay số roi sẽ đánh hoặc số ngày ngồi tù mà cả hai sẽ phải chịu nếu mối đe dọa được thực hiện. Nếu  $X$  là số tiền đô-la tối đa, số roi hay số ngày ngồi tù mà Hàng có thể dùng để đe dọa, chúng ta hãy giải thích  $\pi$  như cách mà Hàng mô tả một tỷ lệ nào đó của mức phạt đối đãi sẽ được áp dụng; nếu  $\pi$  được đặt ở mức 0,5 chẳng hạn, thì cả Hàng lắn Cột cùng nhận được chính xác một nửa mức phạt tối đa. Nếu chúng ta giải thích ma trận theo cách này, và đặt câu hỏi là giá trị nào của  $\pi$  sẽ đưa ra mức đe dọa tối đa từ quan điểm của Hàng, chúng ta sẽ có cùng một cách phân tích và có cùng một kết luận như trước, nghĩa là  $\pi$  sẽ nhỏ ở mức có thể tùy thuộc vào một giá trị cực tiểu tương đương với  $I/(I+X)$ . Như vậy chúng ta có thể giải thích  $\pi$  hoặc là khả năng thực hiện mối đe dọa, hoặc là mức độ mà mối đe dọa chắc chắn sẽ được thực hiện. Bởi vì cả hai hình thức cùng đi đến một kết quả, và vì chúng ta có thể giải thích  $\pi$  theo cả hai cách, nên có lẽ là công bằng khi nói rằng *trong trường hợp này* vai trò của tinh ngẫu nhiên hóa chính là việc nó tạo ra khả năng phân chia một mối đe dọa mà lẽ ra là quá lớn và không thể phân chia, và khả năng tạo ra một mối đe dọa “nhỏ hơn” so với mối đe dọa có sẵn. (Tuy nhiên cũng cần lưu ý ở đây là việc giảm bớt một mối đe dọa

bằng cách giảm khả năng thực hiện mối đe dọa cũng sẽ làm giảm giá trị của kết cuộc một cách tương ứng cho cả hai bên tham gia, trong khi một hình thức giảm bớt trực tiếp về mức độ có thể không bị giới hạn bởi những thay đổi tương ứng trong giá trị hay tính khả dụng cho cả hai bên).<sup>5</sup>

## *RỦI RO CỦA VIỆC THỰC HIỆN VÔ Ý*

Có một yếu tố “chi phí” khác có khả năng làm cho một mối đe dọa được giảm bớt. Đó là nguy cơ xuất phát từ việc một bên thực hiện mối đe dọa một cách khinh suất, ngay cả khi kẻ thù vẫn tuân thủ (hoặc lẽ ra đã tuân thủ nếu mối đe dọa đã không xảy ra do bất cẩn). Khẩu súng đe dọa một tên trộm hoặc một nạn nhân bị bắt giữ có thể bắt ngờ cướp cò trước khi anh ta có một cơ hội tuân thủ. Con chó đe dọa cắn những người xâm nhập có thể cắn một người nào đó không xâm nhập.

Giả sử một người xin đi quá giang chĩa súng vào tài xế và tài xế đe dọa sẽ cho cả hai cùng chết trừ phi người quá giang vứt khẩu súng qua cửa sổ. Người tài xế thực hiện sự đe dọa của mình bằng cách nhấn ga tăng hết tốc lực và tạo ra một

---

<sup>5</sup> Tính ngẫu nhiên cũng có thể liên quan mật thiết với chính việc dàn xếp mối đe dọa, hoặc bao hàm trong quy trình quyết định mà người đe dọa có thể mong muốn thực hiện hay không. Do đó việc giải thích tính ngẫu nhiên chỉ như một phương tiện điều chỉnh quy mô của mối đe dọa chỉ có thể áp dụng trong một số trường hợp.

nguy cơ hiển nhiên về khả năng xảy ra một tai nạn chết người, thì vẫn có một nguy cơ tai nạn sẽ xảy ra trước khi người đi quá giang có cơ hội hiểu ra mối đe dọa và tuân thủ. Trong trường hợp này, nguy cơ của việc thực hiện do sự cố cũng chính là một phần quan trọng của mối đe dọa. Cách duy nhất một bên có thể đưa ra đe dọa là bắt đầu thực hiện nó. Cho đến khi người tài xế tăng tốc, người đi quá giang vẫn không có lý do gì để tin anh ta; một khi tài xế thực sự tăng tốc, cần có một khoảng thời gian tối thiểu để cho người đi quá giang tuân thủ và tài xế giảm bớt tốc độ. Do vậy có một khoảng giữa, cho dù có thể nó ngắn đến đâu, để cho nguy cơ tồn tại; nguy cơ gây ra bởi tốc độ cao do vậy phải là một nguy cơ đủ nhỏ phù hợp với sự chịu đựng của người tài xế trong giai đoạn khoảng giữa này. Nếu thay vào đó chiếc xe hoàn toàn an toàn ở mọi tốc độ dưới 60 dặm một giờ nhưng sẽ trượt ra khỏi đường ở đúng vận tốc 60 dặm và không có các cấp độ nào ở giữa mà có thể tạo ra một mối nguy cơ nhẹ về tai nạn, người tài xế có thể không có động cơ để tạo ra một tốc độ nguy hiểm và người đi quá giang sẽ biết điều đó và sẽ không phản ứng với lời đe dọa suông về tốc độ cao. Chính khả năng xảy ra của một "mối đe dọa chia nhỏ được", một mối đe dọa có kèm theo nguy cơ nhưng không có sự chắc chắn về tai nạn chết người, tạo ra cho người tài xế điều kiện nào đó để tiến hành; nhưng để thực hiện khả năng đó anh ta cũng phải chịu đựng mối đe dọa trong một khoảng giới hạn nào đó.

Nếu trong những tình huống thuộc loại này chúng ta giả thiết - đúng như trong trường hợp của người đi quá giang - rằng nguy cơ của việc thực hiện do khinh suất tương ứng về mặt tỷ lệ với khả năng, tức là  $\pi$ , mà một bên sẽ thực hiện đe

dọa nếu đổi phương không tuân thủ - nếu khuynh hướng con chó giữ nhà cắn người qua đường vô tội tương ứng với khuynh hướng tấn công những người đột nhập vào - thì sẽ có một công thức không khác biệt nhiều với công thức đã đạt được ở trên. Sử dụng cùng một ma trận như ở trên (lần này bỏ qua khả năng cho rằng một mối đe dọa có tiềm năng thực hiện sẽ thất bại) và giả sử  $a\pi$  là khả năng của việc thực hiện do khinh suất, giá trị tối thiểu của  $\pi$  cũng sẽ như trước. Giá trị kỳ vọng của kết quả đối với Hàng, mà phải lớn hơn 0 nếu Hàng thực hiện mối đe dọa, sẽ có ở phía bên trái của công thức.

$$(I - a\pi) - a\pi P > 0$$

hoặc  $\frac{I}{a(I+Y)} > \pi > \frac{I}{I+X}$

Mối đe dọa lớn nhất một lần nữa chính là mối đe dọa vừa đủ vượt qua giới hạn dưới; có một giới hạn trên đối với  $\pi$  mà có thể nhỏ hơn  $I$ ; và tùy thuộc vào các giá trị tương ứng của  $X$  và  $Y$  và thông số "chi phí"  $a$ , ta có thể hoặc không thể tìm ra một giá trị có lời cho  $\pi$ .

### *CAM KẾT NGẪU NHIÊN*

Sau khi đã tìm ra cơ sở hợp lý cho "đe dọa chia nhỏ được", chúng ta có thể đặt câu hỏi liệu chiến thuật "cam kết vô điều kiện" cũng có thể là chiến thuật mà trong một số trường hợp có thể làm giảm nhẹ một cách có lợi tính chắc chắn hay không.

Như đã nêu trong các Chương 3 và 5,<sup>6</sup> một cam kết thuần túy - nghĩa là một cam kết rõ ràng đối với một chiến lược thuần túy - tương đương với "nước đi đầu tiên" trong một trò chơi hai-người, hai-bước trong đó một người sẽ phải bằng cách nào đó là bên đi sau; đó là một phương tiện để đạt được mức tương đương với đi trước. Chúng ta sẽ phải giảm nhẹ cách giải thích đó nếu chúng ta giả thiết rằng Hàng, bên sẽ đi sau trong trò chơi nhưng có sự lựa chọn để cam kết trước, cam kết một cơ hội 50-50 chọn hàng i hoặc ii. Để thực hiện điều này một bên phải giành quyền đi sau, để khai thác quyền cam kết trước. Nếu một bên thực sự phải đi trước, bằng một chọn lựa cụ thể, thì khả năng của một cam kết ngẫu nhiên sẽ mất đi. (Cam kết ngẫu nhiên tương đương với một "nước đi đầu tiên" xác lập bởi một công cụ ngẫu nhiên với những lợi thế do một bên đưa ra, với những lợi thế nhưng không phải là nước đi thực sự mà bên kia biết được cho đến khi bên này thực sự di chuyển).

Cùng một ma trận lợi ích đó (hình 1) có thể được sử dụng để minh họa cho tình huống này nếu chúng ta thay đổi các luật lệ của trò chơi để cho phép Hàng có cam kết *vô điều kiện* trước khi Cột có lựa chọn nhưng không cho phép Hàng thực hiện việc lựa chọn của mình tùy thuộc vào lựa chọn của Cột. Một cam kết chắc chắn với ii dẫn đến việc lựa chọn cột I nhưng bị lãng phí bởi vì ô bên dưới phía trái - mà Hàng cam kết lúc này - không chứa phần thưởng. Vấn đề của Hàng là nó cần hàng ii để khiến cho Cột chọn I, nhưng nó lại cần hàng i để thu lợi từ I. Một sự thỏa hiệp có thể đạt được bằng một

---

<sup>6</sup> Các trang 78, 177.

cam kết ngẫu nhiên – một cam kết đối với một lựa chọn ngẫu nhiên. Nếu Hàng cam kết tung đồng tiền (cơ hội 50-50) để chọn i hoặc ii sau khi Cột đã lựa chọn, thì Cột sẽ chọn I chừng nào X vẫn lớn hơn I.<sup>7</sup> Trong trường hợp đó Hàng có được một giá trị kỳ vọng là 0,5. Nếu Hàng đặt  $\pi$  (khả năng nó lựa chọn ii) ở mức ngay trên  $I/(I+X)$  thì nó sẽ đạt được giá trị kỳ vọng lớn nhất phù hợp với việc lựa chọn I của Cột. (Nếu phần thưởng của Cột trong ô bên dưới phía trái khác không, chẳng hạn 0,5 hoặc -0,5, công thức cho giá trị cực đại của  $\pi$  sẽ khác đi đôi chút). Nếu phần thưởng của Hàng trong ô bên dưới phía trái là -I, không có cam kết nào về một cơ hội có được ii lớn hơn 50% đáp ứng được. Và nếu phần thưởng đó là -X hoặc thấp hơn nữa, sẽ không có khả năng việc kết hợp giữa i và ii sẽ có kết quả; bất cứ sự kết hợp nào với  $\pi$  đủ lớn để dẫn đến cột I sẽ là quá lớn để mang lại cho Hàng một giá trị kỳ vọng dương.

Có một cơ sở lý luận khác đối với cam kết chia nhỏ được. Trong trường hợp vừa được thảo luận, chính mong muốn của Hàng chọn ô bên trên trong I đã dẫn nó đến việc tối thiểu hóa giá trị của  $\pi$ . Trong hình 18, chính động cơ của Cột đòi hỏi cơ hội nào đó đối với hàng i, nghĩa là một giá trị chia nhỏ của  $\pi$ . Trong trường hợp này, một cam kết chắc chắn đối với hàng ii khiến cho Cột chọn II; một cam kết chắc chắn đối với i khiến cho Cột chọn I; hoàn toàn không có cam kết nào để cho Cột thích chọn II; một mối đe dọa phải chọn i trừ phi Cột chọn I

<sup>7</sup> Nghĩa là chừng nào mà phần thưởng của Cột trong ô bên dưới phía bên phải kém hơn phần thưởng của anh ta trong ô bên trên phía trái ở mức độ mà phần thưởng trong ô bên trên phía bên phải vượt quá phía trên bên trái. Xem chú dẫn trước về cấp độ phần thưởng.

sẽ không có hiệu lực trừ phi Hàng hứa sẽ tránh chọn ii. Trong tất cả các trường hợp “chiến lược thuận túy” này, Hàng kết thúc với kết quả 2. Tuy nhiên, nó có thể thực hiện tốt hơn đôi chút với một cam kết hỗn hợp. Nó có thể, bởi vì cả nó và Cột đều bị hấp dẫn vào cột I, chỉ bất đồng với nhau về sự lựa chọn của Hàng trong cột đó.

	I	II
i	4 2	1 1
ii	0 3	2 2

H. 18

Nếu Hàng dành cho Cột một cơ hội 50-50 giữa các hàng i và ii, Cột có được một giá trị kỳ vọng là 2 trong cột đầu tiên, và 1,5 trong cột thứ hai, và sẽ chọn cột đầu tiên. Điều này để lại cho Hàng một giá trị kỳ vọng là 2,5. Bởi vì Hàng thích chọn ii, nó muốn có được khả năng cao nhất của hàng ngang đó phù hợp với nhu cầu cung cấp cho Cột sự mong muốn chọn cột I. Nghĩa là, nó muốn giá trị lớn nhất của  $\pi$  theo đó (thể hiện trong ma trận)

$$4(I - \pi) > (I - \pi) + 2\pi$$

hoặc

$$3/5 > \pi$$

Cam kết hỗn hợp đặc biệt này có thể được gọi là *sự kết hợp* của mối đe dọa chia nhỏ được với lời hứa chia nhỏ được. Hàng, trong thực tế, “đe dọa” một khả năng tương đối cao sẽ áp dụng i trong trường hợp II được lựa chọn và cũng hứa điều đó nếu I được lựa chọn.

Hàng có thể thậm chí làm tốt hơn nữa nếu có thể làm cho  $\pi$  trở thành *có điều kiện* đối với sự chọn lựa của Cột. Bất cứ khả năng hiện thực nào ở mức 0,75 đối với hàng ii, có điều kiện đối với sự chọn lựa cột I, là một mức đủ khiến cho chắc chắn rằng Hàng sẽ tra đưa sự chọn lựa cột II bằng hàng i. Nhưng nếu nó bị giới hạn trong việc làm cho sự đe dọa của mình không tệ hơn mức lời hứa được làm tốt - nếu nó phải gắn kết cùng một khả năng thực hiện vào cả hai sự đe dọa lẫn lời hứa - thì giới hạn trên cho một giá trị hiệu lực của  $\pi$  là 0,6, với một giá trị kỳ vọng cho Hàng là 2,6 (và là 1,6 đối với Cột). Với một  $\pi$  riêng dành cho sự hứa hẹn, thì giới hạn trên sẽ là 0,75 cho một lợi ích kỳ vọng là 2,75 (và chỉ là 1,0 cho Cột).

## ĐE DỌA CÓ YẾU TỐ NGÃU NHIÊN

Điển hình là các đe dọa chiến lược trong đó hành động trừng phạt - nếu đe dọa thất bại và phải được thực hiện - vừa đau đớn vừa tổn kém cho cả hai bên. Mục đích là ngăn chặn trước, chứ không phải trả thù sau. Đưa ra một mối đe dọa đáng tin bao hàm việc chứng minh rằng một bên sẽ phải thực hiện lời đe dọa đó, hoặc tạo ra các động cơ cho chính bên đó hoặc chịu những hình phạt mà khiến cho một bên rõ ràng mong muốn. Mục tiêu rõ ràng của việc đưa lính Mỹ tới châu Âu như một "đội quân tiền tuyến" là nhằm thuyết phục phía Nga rằng chiến tranh ở châu Âu sẽ lôi Mỹ vào cuộc cho dù phía Nga có nghĩ rằng Mỹ muốn tham gia hay không - và việc không thực hiện cam kết về mặt thực tế là không thể.

Như một quy luật, một bên phải đe dọa rằng anh ta sẽ hành động, chứ không phải là anh ta có thể hành động, nếu đe dọa thất bại. Nói rằng một bên có thể hành động cũng có nghĩa nói rằng bên đó có thể không hành động, và nói điều này là thú nhận rằng một bên đã nắm quyền ra quyết định - điều mà bên này không cam kết. Chỉ nói rằng một bên có

thể thực hiện đe dọa, chứ không nói rằng bên đó chắc chắn sẽ thực hiện, thì đồng nghĩa với việc mời gọi đối phương suy đoán là liệu bên đó có mong muốn trừng phạt chính mình lẫn đối phương hay là sẽ bỏ qua cơ hội. Hơn nữa, nếu một bên nói rằng anh ta có thể - chứ không phải anh ta sẽ - và đối phương không đầu hàng trước lời đe dọa, và người đe dọa chọn lựa cách không thực hiện lời đe dọa, anh ta chỉ cung cấp niềm tin của đối phương rằng khi anh ta có một chọn lựa rõ ràng giữa hành động hay thoái lui thì anh ta sẽ chọn thoái lui (tự an ủi rằng mình không bị coi là kẻ không giữ lời bởi vì chưa bao giờ nói rằng chắc chắn sẽ hành động).

Dù vậy có những mối đe dọa thuộc loại này mà vẫn có hiệu quả bất chấp kẽ hở đó. Tuy nhiên, chúng chỉ có tác dụng khi thông qua một quy trình phức tạp hơn một cấp so với một cam kết khẳng định là chắc chắn sẽ thực hiện. Hơn nữa, chúng có thể xảy ra do khinh suất và dẫn đến hành vi không chủ ý. Ví lý do này chúng ít có cơ may được nhận ra và hiểu rõ.

Cốt lõi của các đe dọa này là, mặc dù một bên có thể có hoặc không thực hiện đe dọa nếu bên bị đe dọa không tuân thủ, thì *quyết định cuối cùng không hoàn toàn nằm trong sự kiểm soát của bên đe dọa*. Mỗi đe dọa không hoàn toàn thuộc loại “Tôi có thể có hoặc có thể không, tùy thuộc cách tôi lựa chọn”, mà chưa đựng một yếu tố thuộc loại “tôi có thể có hoặc có thể không, và ngay cả chính tôi cũng không hoàn toàn chắc chắn.”

Vậy thì yếu tố không chắc chắn trong quyết định xuất phát từ đâu? Nó phải xuất hiện từ đâu đó nằm ngoài kiểm soát của người đe dọa. Dù chúng ta gọi nó là “may rủi”, tai nạn, tác

động bên thứ ba, sự không hoàn chỉnh trong cơ chế ra quyết định, hay chỉ là các quy trình mà chúng ta không hiểu hết, thì nó vẫn là một thành tố trong tình huống mà cả bên đe dọa lẫn bên bị đe dọa không thể hoàn toàn kiểm soát được. Một ví dụ là mối đe dọa của một cuộc chiến tranh khinh suất.

## *ĐE DỌA CỦA CHIẾN TRANH VÔ Ý*

Ý tưởng rằng cuộc chiến tranh toàn diện có thể được khởi động một cách thiếu thận trọng - thông qua một sự cố nào đó, một báo động giả, hay lỗi về cơ chế; thông qua sự hoảng sợ của ai đó, sự điên cuồng hay độc ác; thông qua việc hiểu sai các ý định của kẻ thù hay việc hiểu chính xác sự hiểu sai của kẻ thù về chúng ta - không phải là một ý tưởng hấp dẫn. Theo nguyên tắc thông thường thì người ta muốn giữ giả định đó ở mức tối thiểu; và trong những tình huống đặc biệt khi cảng thẳng gia tăng và các lực lượng chiến lược được đặt trong tình trạng báo động cao nhất; khi động cơ phản ứng nhanh được tăng cường bởi ý tưởng rằng bên kia sẽ tấn công trước, thì điều có vẻ đặc biệt quan trọng là việc chống lại quyết định bối đồng, chống lại những sai lầm trong phán đoán, và các mâu hành vi đáng ngờ hay mơ hồ. Đường như xét về lý do con người lẩn cơ học, thì khả năng xảy ra một cuộc chiến tranh vô ý tăng theo khung hoảng.

Đe dọa - nếu chúng ta gọi cơ chế hành vi ngẫu nhiên này

như vậy - có một số đặc điểm thú vị. Nó có thể tồn tại dù chúng ta có nhận ra hay không. Ngay cả những người đã nghi ngờ rằng liệu có phải đe dọa trả đũa trên quy mô lớn của Hoa Kỳ là một công cụ răn đe tiềm năng đối với sự gây hấn *nhỏ* trong nhiều năm qua, nhưng cảm thấy khó hiểu khi phía Nga đã không có những hành động xấu hơn mức mà họ đã thực hiện, cũng có thể nhận thấy rằng đe dọa mà Hoa Kỳ đưa ra được hỗ trợ bởi một đe dọa tiềm ẩn bổ sung, đó là Hoa Kỳ có thể bị kích động bởi những hành vi của Liên Xô. Hơn nữa, ngay cả nếu Hoa Kỳ không muốn chịu đựng dù chỉ là một khả năng nhỏ xảy ra một cuộc chiến do sơ suất, và không muốn sử dụng cơ chế này một cách có chủ đích, thì "mối đe dọa" đang được đề cập có thể là một sản phẩm phụ của các hành động khác mà Hoa Kỳ có động cơ mạnh mẽ để thực hiện. Hoa Kỳ có thể gấp phải đe dọa này cho dù muốn hoặc không khi Hoa Kỳ (và phía Nga) áp dụng những biện pháp để phòng tương ứng với một cuộc khủng hoảng. Hiểu được điều này, phía Nga có thể đưa nguy cơ vào trong cân nhắc của mình. Cuối cùng, mối đe dọa không bị giảm bớt mức độ đáng tin ngay cả nếu phía Nga hoàn thành mục đích của họ mà không châm ngòi cuộc chiến. Nếu phía Nga tính toán rằng cơ hội xảy ra một cuộc chiến tranh vô ý trong một tháng cụ thể nào đó là từ rất nhỏ đến nhỏ vừa phải nếu họ tạo ra một cuộc khủng hoảng, và họ vẫn tiếp tục kế hoạch của họ, và không có cuộc chiến lớn nào xảy ra, thì họ cũng chẳng có lý do gì để giả thiết rằng việc lặp lại sẽ ít rủi ro hơn, không khác gì mấy so với việc một người thoát chết trong một trò chơi bóp cò một lần súng ru-lô lại dám khẳng định rằng khẩu súng không nguy hiểm chút nào.

## *CHIẾN TRANH HẠN CHẾ NHƯ MỘT NGUỒN PHÁT SINH RỦI RO*

Chiến tranh hạn chế như một công cụ ngăn chặn đối với sự gây hấn cũng đòi hỏi phải được giải thích là hành động nâng cao *khả năng xảy ra* một cuộc chiến lớn hơn. Nếu chúng ta hỏi các lực lượng phương Tây ở châu Âu sẽ làm cách nào để ngăn chặn một cuộc tấn công của phía Nga hoặc kháng cự thế nào nếu nó xảy ra, câu trả lời thường là một chuỗi các *quyết định*. Trong trường hợp tấn công ở một quy mô vừa phải, chúng ta có thể đưa ra quyết định tiến hành một cuộc chiến tranh hạn chế; đó sẽ không phải là một quyết định đi đến hủy diệt lẫn nhau. Nếu chúng ta có thể chống lại phía Nga trên quy mô nhỏ, thì hoặc họ sẽ phải từ bỏ ý tưởng của mình hoặc họ sẽ tăng thêm một bước về quy mô của cuộc chiến. Ở một điểm nào đó có một bước nhảy không liên tục từ chiến tranh hạn chế sang chiến tranh toàn diện, và chúng ta hy vọng đương đầu với *họ* bằng sự lựa chọn đó. Nếu đây không phải là chuỗi điển hình các quyết định đã được hình dung, thì có lẽ ít nhất nó cũng điển hình ở một khía cạnh: nó bao hàm các quyết định *có chủ đích* - các quyết định thực hiện một hành động hoặc từ bỏ một hành động, khởi phát cuộc chiến hay không, tăng cường quy mô ác liệt hay không, đáp trả một thách thức hay không.

Nhưng một cách giải thích khác có thể áp dụng với chiến tranh hạn chế. Nguy cơ của chiến tranh tổng lực thì hầu như

chắc chắn sẽ gia tăng theo sau việc nổ ra một cuộc chiến tranh hạn chế; nó hầu như chắc chắn sẽ được gia tăng do việc mở rộng cuộc chiến hạn chế. Do vậy, đe dọa tham gia vào một cuộc chiến tranh hạn chế có hai phần. Một là đe dọa gây ra tổn kém trực tiếp trên đối phương, từ hao tổn sinh mạng, chi phí, đến mất lãnh thổ, mất thể diện hoặc bất cứ điều gì khác. Hai là đe dọa đưa bên kia, cùng với bên này, đổi mặt với một nguy cơ cao hơn của một cuộc chiến tranh toàn diện.<sup>1</sup>

Ở đây một lần nữa là đe dọa về một cuộc chiến tổng lực có thể xảy ra, chứ không phải là chắc chắn sẽ xảy ra, nếu bên kia tham gia vào một số hành động nào đó. Như vậy, một lần nữa, việc nó có xảy ra hay không không phải là vấn đề hoàn toàn được kiểm soát bởi bên đe dọa. Chỉ có điều một cuộc chiến tổng lực sẽ xảy ra như thế nào - đâu chính là điểm xuất phát của sai lầm, khởi phát hay sự hiểu lầm - thì không ai biết chắc. Dù cho đó là bất cứ điều gì khiến cho một cuộc tranh chiến hạn chế giữa hai siêu cường trở thành một nguy cơ, thì đó là một nguy cơ rất thật mà cả hai bên không thể hóa giải được cho dù bên này lẫn bên kia đều muốn. Quyết định cuối cùng, hay hành động then chốt đưa đến một quy trình không thể đảo ngược, không phải là điều gì nhất thiết được kỳ vọng cùng thực hiện một cách có chủ đích. "May rủi" giúp quyết định liệu cuộc chiến toàn diện có xảy ra hay không, với những sai sót là vấn đề suy đoán dựa trên bản chất của cuộc chiến tranh hạn chế cũng như bối cảnh mà nó xảy ra.

---

<sup>1</sup> Glenn H. Snyder cùng nhấn mạnh quan điểm này trong "Ngăn chặn bằng Phủ nhận và Trừng phạt" (Deterrence by Denial and Punishment) (Chuyên khảo Nghiên cứu Số I: Trung tâm Nghiên cứu Quốc tế Đại học Princeton, ngày 2-1-1959) trang 12, 29.

Tại sao người ta lại đe dọa bằng một cuộc chiến tranh hạn chế thay vì một cuộc chiến tổng lực nhằm ngăn chặn một cuộc tấn công? Trước tiên, đe dọa bằng một cuộc chiến tranh hạn chế - theo phân tích này - đồng nghĩa với việc đe dọa bằng một nguy cơ xảy ra cuộc chiến tổng lực, chứ không phải chắc chắn có cuộc chiến tổng lực; do vậy nó là một đe dọa giảm nhẹ hơn so với đe dọa trả đũa quy mô lớn và thích hợp hơn cho tình huống nào đó. Thứ hai, trong trường hợp kẻ thù đánh giá sai những ý định hay những cam kết của chúng ta, nó có lợi thế như là một giai đoạn trung gian: chúng ta có thể *tham gia* vào một cuộc chiến tranh hạn chế, tạo ra đúng cái nguy cơ cho cả hai bên mà chúng ta đe dọa sẽ tạo ra, mà từ đó không phải chịu đựng một cuộc chiến tổng lực mà cả hai bên cùng phải trả giá do đánh giá sai của kẻ thù. Thay vào đó chúng ta trả một giá nhẹ hơn cho nguy cơ của một cuộc chiến tổng lực, một nguy cơ mà kẻ thù có thể làm giảm bớt bằng cách rút quân hay giải quyết.

Thứ ba, trong trường hợp kẻ thù bất hợp lý hoặc hiếu chiến, hoặc khi chúng ta đánh giá sai động cơ hay cam kết của đối phương, hoặc trong trường hợp hành động gây hấn của kẻ thù đã đi quá đà không còn ngăn chặn được, hoặc khi hành động của kẻ thù đang được thực hiện bằng những đội quân bù nhìn hay vệ tinh nằm ngoài khả năng kiểm soát tức thời của kẻ thù, thì việc đe dọa bằng một nguy cơ sẽ mang tính cẩn trọng cao hơn so với đe dọa bằng một sự chắc chắn. Nếu chúng ta đe dọa sử dụng một cuộc chiến tổng lực, nghĩ rằng chưa quá trễ để ngăn chặn kẻ thù, và khi nó thực sự như thế (quá trễ), chúng ta hoặc sẽ phải tiến hành thực hiện lời đe dọa, hoặc để cho lời đe dọa của chúng ta mất độ tin cậy.

Nhưng nếu chúng ta có thể đe dọa kẻ thù bằng một tỷ lệ 1 phần 20 nguy cơ xảy ra chiến tranh tổng lực trong trường hợp kẻ thù vẫn lẩn tới, và nếu kẻ thù tiếp tục lẩn tới, chúng ta có thể giữ hơi vì có đến 19-trên-1 cơ hội để chúng ta thoát ra mà không phải tiến hành một cuộc chiến toàn diện. Dĩ nhiên, nếu chúng ta giảm mức độ nguy cơ cho chính mình, chúng ta cũng giảm mức độ nguy cơ cho kẻ thù; đặt quá nhiều sự an toàn trong một mối đe dọa có thể làm giảm cấp độ mối đe dọa đó. Nhưng trong những trường hợp có nguy cơ chúng ta hoàn toàn phán đoán sai cam kết của kẻ thù đối với một hành động, hoặc hoàn toàn phán đoán sai khả năng của kẻ thù trong việc kiểm soát các cơ sở, các đồng minh, các tư lệnh của họ, thì một nguy cơ vừa phải hơn có thể ngăn chặn được bất cứ điều gì vẫn còn nằm trong quyền kiểm soát của kẻ thù.

Nếu chúng ta sử dụng cách giải thích này vào chiến tranh hạn chế, chúng ta cũng có thể sử dụng một cách giải thích tương tự cho việc mở rộng chiến tranh hay đe dọa mở rộng chiến tranh. Đe dọa đưa các loại vũ khí mới vào sử dụng trong một cuộc chiến tranh hạn chế, theo cách lập luận này, không được đánh giá đơn thuần theo lợi thế quân sự hay chính trị tức thời, mà còn theo nguy cơ của một mối đe dọa có chủ đích về một cuộc chiến rộng lớn hơn mà nó bộc lộ ra. Cũng như một cuộc chiến tranh hạn chế vừa phải có thể làm tăng đáng kể khả năng xảy ra một cuộc chiến lớn trong vòng 30 ngày tới, thì việc chuyển đổi từ vũ khí quy ước sang các loại vũ khí mới lạ có thể nâng cao khả năng xảy ra đó theo một mức khác.

Chúng ta đang được dẫn dắt theo cách này để đi đến một cách giải thích mới về “đội quân tiền tuyến” (trip wire). Cách

so sánh về các lực lượng quân sự chiến tranh hạn chế của Mỹ ở châu Âu, theo lập luận này, không phải là một sợi dây kích nổ mà chắc chắn sẽ kích nổ một cuộc chiến tranh tổng lực nếu nó hoạt động tốt và sẽ hoàn toàn thất bại nếu nó không hoạt động. Điều chúng ta có là một chuỗi theo cấp độ các sợi dây kích nổ, mỗi một dây gắn với một cơ chế may rủi, với *khả năng* kích nổ gia tăng hàng ngày khi kẻ thù di chuyển từ đường dây này đến đường dây khác. Điểm đặc biệt của cách so sánh này, cần phải nhấn mạnh ở đây, là cho dù sợi dây kích nổ có kích hoạt một cuộc chiến tranh tổng lực hay không thì - ít nhất ở một mức độ nào đó - nó nằm ngoài khả năng kiểm soát của chúng ta, và phía Nga biết điều đó.

Cách giải thích tương tự cũng có thể đúng với Kim Môn (một hòn đảo thuộc Đài Loan, nằm rất gần với Trung Quốc). Người ta có thể tranh luận rằng phía Trung Quốc và Nga bị ngăn chặn bởi viễn cảnh của một cuộc chiến tranh lớn, chứ không phải viễn cảnh của việc thua trong một cuộc chiến tranh hạn chế hay thắng một cuộc chiến tranh hạn chế với giá quá đắt. Ngay cả khi họ được thuyết phục để tin tưởng rằng Mỹ sẽ rất thận trọng và khéo léo để giữ cuộc chiến ở mức giới hạn, và phía họ cũng được chuẩn bị để có nước đi rất thận trọng và khéo léo, họ có thể đơn giản cảm thấy rằng quy trình để tiến tới các cuộc chiến mở rộng và mở rộng hơn nữa không phải là một quy trình mà họ hay Mỹ có thể hoàn toàn hiểu biết hay tiên liệu được, và rằng nguy cơ, mặc dù nhỏ về mặt số học, là đáng lưu ý.

## *HÀNH VI CÓ TÍNH RỦI RO TRONG CHIẾN TRANH HẠN CHẾ*

Nếu một trong những chức năng của chiến tranh hạn chế là phơi bày nguy cơ về một cuộc chiến tổng lực, nhằm gây khó khăn cho kẻ thù và khiến cho việc theo đuổi các mục tiêu giới hạn của kẻ thù trở nên rủi ro ở mức không chấp nhận được đối với chính họ, thì các quy tắc hành vi thông thường trong chiến tranh hạn chế cần phải được điều chỉnh lại. Mục đích tối hậu có thể không phải là *đảm bảo* rằng nó sẽ dừng ở mức giới hạn, mà là giữ nguy cơ của một cuộc chiến tranh tổng lực trong các phạm vi vừa phải ở mức *trên zero*. Ít nhất đây có thể là chiến lược cho một bên đang ở trong nguy cơ “thua” trong cuộc chiến tranh hạn chế. Nếu ít có khả năng xảy ra tình trạng các bước tiến gây hấn của kẻ thù có thể được kiềm chế bởi sức đề kháng giới hạn cục bộ, thì có thể có nhiều lý do hơn để tạo ra nguy cơ cho cả hai bên một cách có chủ đích. (Khả năng khác, nếu bên gây hấn càng có thể thiết kế các bước tiến của mình sao cho ngay cả sự đề kháng cục bộ cũng có vẻ có nguy cơ bùng nổ, thì sự đề kháng cục bộ sẽ càng có vẻ kém hấp dẫn).

Chú ý nâng cao nguy cơ về một cuộc chiến tổng lực do vậy là một chiến thuật phù hợp với bối cảnh của chiến tranh hạn chế. Dĩ nhiên, người ta không thể nâng cao nguy cơ chỉ bằng cách nói miệng. Người ta không thể chỉ thông báo với kẻ thù rằng diễn tiến ngày hôm qua mới chỉ là hai phần trăm của

việc chuẩn bị tiến tới một cuộc chiến tổng lực, và hôm nay sẽ là bảy phần trăm, và họ cần phải cảnh giác. Người ta cần phải có hành động mà - giả sử anh ta và kẻ thù của mình cùng quan tâm và cẩn trọng như nhau trong việc duy trì cuộc chiến ở mức giới hạn - khiến cho mọi người cảm thấy rằng cuộc chiến sẽ nằm trong tầm kiểm soát kém chắc chắn đi một chút.

Ý tưởng rất đơn giản ở đây là một cuộc chiến tranh hạn chế có thể vượt khỏi tầm khống chế theo từng bước. Tại một thời điểm nào đó người ta có một nhận thức hay ý tưởng về mức độ nó đã "vượt tầm kiểm soát" như thế nào. Và các hành động khác nhau - sáng kiến thay đổi, hành động vi phạm các giới hạn, biểu hiện "vô trách nhiệm", hành động thách thức hay khẳng định, việc áp dụng thái độ đe dọa chiến lược, việc sử dụng các liên quân và cộng sự khó kiểm soát, các chiến thuật đánh lừa hay quấy rối, đưa vào sử dụng những vũ khí mới, mở rộng quy mô quân đội hay vùng chiến sự - có xu hướng nâng cao phán đoán của hầu hết mọi người về một tình hình "vượt ngoài tầm kiểm soát" như thế nào. Chia sẻ một nguy cơ gia tăng như thế với kẻ thù có thể tạo động cơ từ bỏ mạnh mẽ cho kẻ thù. Cách ưa thích hơn là người ta tạo ra nguy cơ được chia sẻ bằng các cuộc tập trận hay các cam kết không đảo ngược, do đó chỉ có sự rút quân của kẻ thù mới làm dịu tình hình; ngược lại nó có thể trở thành một cuộc chiến cân não.

## *TRẢ ĐÜA VÀ QUÁY RỐI*

Cuộc chiến tranh cục bộ hạn chế không phải là bối cảnh duy nhất trong đó hành vi tạo nguy cơ có chủ đích có thể được sử dụng như một loại đe dọa. Giữa những đe dọa về trả đũa quy mô lớn và chiến tranh hạn chế tồn tại một khả năng trả đũa kém-quy-mô-lớn, đó là trả đũa theo cấp độ. Có rất ít các công trình phân tích về trả đũa giới hạn đã được xuất bản.<sup>2</sup>

Chúng ta sẽ giải thích như thế nào về một hành động kịch tính, chẳng hạn như trả đũa hạt nhân hạn chế trên lãnh thổ kẻ thù? Như trong cuộc chiến tranh hạn chế, có thể lại có hai phần thuộc về "chi phí" đối với một kẻ thù. Một là chi phí trực tiếp: sinh mạng, sự tàn phá, sự sỉ nhục, hoặc bất cứ điều gì tương tự. Thứ hai là nguy cơ về một cuộc chiến tổng lực. Không ai hoàn toàn biết rõ điều gì sẽ xảy ra nếu một quốc gia cho phát nổ vũ khí hạt nhân trong quốc gia kẻ thù. Nếu hành động đó được nhìn nhận như là một hành động đơn lẻ, có giới hạn về mặt ý định, không phải là một phần của một cuộc tấn công quy mô, mà cũng không phải là một phần của cuộc đánh lén chống lại khả năng trả đũa của bên kia, thì nạn nhân có thể nhận thấy là không khôn ngoan khi tung ra một cuộc chiến tổng lực nhằm đáp trả lại nỗi đau và sự sỉ nhục. Nhưng ngay cả khi không đáp trả lại như thế, thì phía nạn

---

<sup>2</sup> Một cuộc thảo luận nghiêm túc mới đây là do Morton A. Kaplan, trong "Chiến lược trả đũa hạn chế" (The Strategy of Limited Retaliation) (Bản Ghi nhớ Chính sách 19 của Trung Tâm Nghiên cứu Quốc tế; Princeton, ngày 9-4-1959).

nhân vẫn có xu hướng làm một điều gì đó mà đến lượt mình gây ra các hậu quả có thể cuối cùng tiến đến ranh giới của một cuộc chiến tổng lực. Nếu sự đáp trả chỉ đơn giản là tấn công giáng trả trong một cách thức tương tự, thì quá trình này có thể giảm dần, hoặc bùng nổ. Do đó, ngay cả khi mỗi bên mong muốn hành động một cách cẩn trọng, thì việc không có khả năng hiểu rõ mỗi bên sẽ phản ứng như thế nào có thể mang lại một quy trình động mà cuối cùng sẽ bùng nổ thành một cuộc chiến tổng lực.

Điều đó vẫn có thể khó xảy ra. Ở đây chúng ta đang đề cập đến một hành động mà *có thể có hoặc có thể không* dẫn tới một cuộc chiến tranh tổng lực, kết quả cuối cùng *không* nằm trong sự kiểm soát hoàn toàn của các bên tham gia, và khả năng dẫn đến một cuộc chiến tranh tổng lực là một ván đe thuộc phạm vi phán đoán. Đề cập đến những khả năng này không nhất thiết nghĩa là gợi ý chúng, mà để cho biết những khả năng đó nên được giải thích như thế nào. Sự trừng phạt áp đặt trên nạn nhân - một sự trừng phạt mà bên đe dọa cùng phải chia sẻ với nạn nhân - chính là một mức gia tăng về khả năng xảy ra một cuộc chiến toàn diện có thể nhận biết.

## *HÀNH VI CÓ TÍNH RỦI RO VÀ ĐE DỌA “BẮT BUỘC”*

Có một sự khác biệt điển hình giữa đe dọa nhằm khiến cho kẻ thù phải *làm* gì đó (hoặc *ngưng làm* gì đó) và đe dọa nhằm giữ chân kẻ thù không khởi động một điều gì đó. Nét khác biệt

chính nằm ở tính thời điểm, nằm ở chỗ ai sẽ có nước đi đầu tiên, nằm ở chỗ sáng kiến của ai sẽ được đưa vào thử nghiệm. Để ngăn chặn bước tiến của kẻ thù bằng một đe dọa thì có thể là đủ nếu đốt cây cầu phía sau tôi khi tôi đối mặt với kẻ thù. Để ép buộc kẻ thù rút lui bằng đe dọa, tôi cần phải cam kết tiến lên phía trước, và điều này đòi hỏi phải phóng hỏa đám cỏ phía sau tôi khi ngọn gió thổi về hướng quân thù. Tôi có thể chặn chiếc xe hơi của bạn trên đường bằng cách đặt chiếc xe của tôi trên đường của bạn; đe dọa ngăn chặn của tôi mang tính thụ động, quyết định cho dung xe thuộc về bạn. Tuy nhiên, nếu bạn thấy tôi ở trên đường và đe dọa sẽ dụng xe trừ phi tôi di chuyển đi, thì bạn sẽ không có được lợi thế như vậy; quyết định dung xe vẫn thuộc về bạn, còn tôi lại thích biện pháp ngăn chặn. Bạn phải dàn xếp để phải dung xe nếu tôi không di chuyển, và đó là một cấp độ phức tạp hơn.

Một mối đe dọa mang tính chất ép buộc hơn là ngăn chặn, do đó, thường có hình thái áp dụng hình phạt *cho tới khi* phía bên kia hành động, hơn là *nếu* phía bên kia hành động. Điều này là như thế bởi vì thông thường cách duy nhất để trở nên có cam kết cụ thể đối với một hành động là khởi xướng nó. Khởi xướng một nỗi đau đều đặn, ngay cả người đe dọa cũng phải chịu nỗi đau, có thể tạo ra ý nghĩa như một mối đe dọa, đặc biệt khi người đe dọa có thể khởi xướng nó theo một chiều hướng không thể đảo ngược đến nỗi chỉ có sự tuân thủ của người kia mới có thể làm giảm nhẹ nỗi đau mà cả hai cùng chịu đựng. Nhưng khởi xướng theo cách không thể đảo ngược một thảm họa, nếu bên khởi xướng cũng phải chịu đựng, là một điều không hay. Tuy nhiên, khởi xướng không thể đảo ngược một *nguy cơ* vừa phải về một thảm họa mà hai bên

cùng phải chịu, nếu sự tuân thủ của bên kia là có thể diễn ra trong một khoảng thời gian đủ ngắn để giữ nguy cơ nằm trong các giới hạn chịu đựng, thì có thể là một phương cách làm giảm cấp độ đe dọa xuống một mức mà bên này sẵn sàng triển khai. Đặt kẻ thù (và chính mình) trước nguy cơ một phần trăm của một thảm họa khủng khiếp sẽ xảy ra sau mỗi tuần lễ mà kẻ thù không chịu tuân thủ thì ít nhiều cũng tương tự như đặt kẻ thù (và chính mình) trước một tỷ lệ tổn thất hàng tuần tương đương với một phần trăm của thảm họa. (Các từ “ít nhiều” và “tương đương” có thể giải thích rất linh hoạt ở đây.<sup>3</sup>

“Lắc lư con thuyền” là một ví dụ tốt. Nếu tôi nói, “hãy chèo thuyền, nếu không tôi sẽ làm lật nhào con thuyền và cả hai cùng chết đuối”, thì bạn sẽ nói rằng bạn không tin tôi. Nhưng nếu tôi làm lắc lư con thuyền đến mức nó có thể bị lật nhào, bạn sẽ bị ấn tượng hơn. Nếu tôi không thể áp dụng sự đau đớn gần đến mức cả hai cùng chết, thì “một ít” cái chết, dưới hình thức một khả năng nhỏ rằng con thuyền sẽ lật nhào, là gần tương đương. Nhưng, để cho nó có hiệu quả, tôi phải thật sự làm cho con thuyền lâm vào hiểm nghèo; chỉ nói suông rằng tôi có thể làm lật nhào cả hai là không thuyết phục.

Một cách lý tưởng, để minh họa cho mục đích này, có lẽ tôi cần có một chiếc hộp đen nhỏ có chứa một bánh quay ru-lét và

---

<sup>3</sup> Khởi xướng một hành động rủi ro, nếu một bên không thể khởi xướng nó theo chiều hướng không thể đảo ngược, không nhất thiết phải “chiến thắng” trước đối thủ: bên kia có thể vẫn hy vọng, bằng cách hành động kiên quyết, khiến cho người khởi xướng phải thối lui. Một bên vẫn phải thắng trong “cuộc chiến cân não” nếu kẻ thù chọn cách chơi kéo dài thời gian. Nhưng ít nhất thì tình huống đối xứng này sẽ thay thế một tình huống trong đó tính bất đối xứng có lợi cho kẻ thù là bên chiến thắng mặc nhiên nếu không có bên nào hành động.

một thiết bị sẽ phát nổ và bằng một cách nào đó chắc chắn sẽ kích hoạt một cuộc chiến toàn diện. Thế rồi tôi đặt chiếc hộp nhỏ này xuống, và nói với người Nga rằng tôi đã khởi động nó để mỗi ngày bánh ru-lét sẽ quay một lần với một xác suất cho trước (được đánh số cụ thể và phía Nga biết rõ) rằng, vào bất cứ ngày nào đó, chiếc hộp nhỏ sẽ kích động một cuộc chiến toàn diện. Tôi nói với họ - *biểu thị cho họ thấy* - rằng chiếc hộp nhỏ vẫn sẽ tiếp tục hoạt động cho đến khi các yêu cầu của tôi được tuân thủ và rằng *không có cách nào để tôi có thể ngăn chặn nó*. Lưu ý tôi không nhấn mạnh rằng tôi sẽ quyết định về một cuộc chiến toàn diện, hay khởi xướng nó một cách có chủ đích, nếu như chiếc hộp quay trúng con số quyết định. Tôi giao lại hết cho chiếc hộp việc nó sẽ *tự động* nhấn chìm cả hai trong chiến tranh nếu con số kết hợp đúng (hay sai) xuất hiện vào bất cứ ngày nào.<sup>4</sup>

Trên cơ sở rằng, ngay cả khi kẻ thù tuân thủ, có nguy cơ nào đó chiếc hộp sẽ châm ngòi chiến tranh trước khi kẻ thù có cơ hội trấn tĩnh lại và thực hiện thương lượng, thì vẫn có một lợi thế trong việc làm cho việc chiếc hộp sẽ phát nổ vào một ngày nào đó trở nên ít chắc chắn hơn. Trong việc ngăn chặn thông thường - khi mà không có chuyện gì xảy ra trừ phi kẻ thù hành động ngược lại đề nghị của chúng ta - đe dọa quá nhiều có thể trở nên quá loãng nhưng không làm cho mình tự thua cuộc; trong trường hợp hiện tại - nơi mà đe dọa bắt đầu triển khai ở một tốc độ ổn định theo thời gian ngay

---

<sup>4</sup> Chiến thuật này có thể ít rủi ro hơn nếu cơ chế mang tính tự động nhiều hơn; càng riang tính tự động, thì kẻ thù càng ít có động cơ kiểm tra các ý định của tôi trong một cuộc chiến cân não, như thế kéo dài thời gian của nguy cơ.

khi chúng ta cam kết thực hiện nó - một đe dọa quá lớn có thể làm hỏng mục đích của nó. Trong tình huống này, một đe dọa nhỏ đảm bảo thực thi được không chỉ là biện pháp thay thế khả dĩ cho một đe dọa lớn chắc chắn; nó là một phương cách thay thế ưu việt hơn và cần thiết.

Hãy lấy một ví dụ. Một nước châu Âu, sau khi có được một lực lượng hạt nhân trả đũa ở quy mô khiêm tốn, yêu cầu Nga rút ra khỏi Hungary nếu không nước nọ sẽ gây ra tổn thất khủng khiếp cho nước Nga. Phía Nga phớt lờ mối đe dọa, bởi vì cách đó không thuyết phục khi mà nước đe dọa lại khiến chính họ *phải làm* điều gì đó mang tính chất tự sát như thế. Nhưng cách khác, nước này đe dọa bắn một quả tên lửa mỗi ngày qua Liên Xô, với một đầu đạn hạt nhân và một thiết bị ngẫu nhiên sẽ kích nổ nơi nào đó trên không phận nước Nga nếu nó không bị bắn hạ. Phía Nga nói họ không tin là nước nọ sẽ thực hiện điều đó; và nước nọ thực hiện. Phía Nga phản kháng và đe dọa, một ngày trôi qua, và nước nọ lại thực hiện. Có thể là một đầu đạn bay ngang qua và phát nổ, có thể là một vài đầu đạn phát nổ, có thể không có đầu đạn nào phát nổ; nếu một vài đầu đạn phát nổ, có thể chúng phát nổ bên trên các thành phố, có thể bên trên vùng nông thôn đông dân cư, có thể trên những vùng hoang vắng. Nước nọ vẫn tiếp tục duy trì.

Vậy nước nọ đang làm gì? Yếu tố cốt lõi nước nọ đang thực hiện - bên cạnh việc gây tổn thất và làm mất mặt Nga - là gây ra nguy cơ đau đớn khiến cho cả nước này lẫn Nga (và cả phần còn lại của thế giới) sẽ phải tham gia vào một cuộc chiến tổng lực trong tương lai gần, một cuộc chiến mà cả nước này lẫn Nga đều không muốn. Nước này trên thực tế đang nói rằng "Nếu các người không rút ra khỏi Hungary, chúng tôi có

thể khiến cho một cuộc chiến tổng lực xảy ra." Vậy thì người Nga phải rút ra khi nào? Họ rút ra càng sớm chừng nào, thì nguy cơ của một cuộc chiến (từ nguyên nhân này) càng sớm được chấm dứt hoặc giảm bớt. Nước đang gây áp lực không có ý nói rằng "Hãy rút ra nếu không chúng tôi sẽ cố ý phát động một cuộc chiến tranh." Quyết định không ở về phía họ, và cũng không tùy thuộc vào việc họ thể hiện một giải pháp biểu thị cho một hành động cuối cùng. Phía Nga có thể giả thiết rằng nước nọ sẽ làm mọi điều có thể để ngăn chặn một cuộc chiến toàn diện; nhưng họ cũng phải nhận ra rằng với những vật thể này bay lờn vờn, phát nổ bất cứ lúc nào, và với việc chính họ đáp trả bằng bất cứ cách nào mà họ cảm thấy bắt buộc phải làm, thì cũng không hoàn toàn rõ ràng rằng nước nọ, và người Nga, biết làm cách gì để cho chiến tranh toàn diện không xảy ra.

Phản minh họa này chỉ có dụng ý như một phép so sánh tương tự cho các hành động khác trong đó việc đưa ra nguy cơ của một cuộc chiến toàn diện có thể không dễ dàng được nhận thức là một phần thiết yếu của những gì đang diễn ra. Lấy một tình huống gần gũi hơn, giả sử một đoàn xe bọc thép được đưa đến Berlin trong trường hợp việc tiếp cận bằng bộ binh bị ngăn chặn, hoặc giả sử, một khi việc thắt chặt giao thông đối với Berlin trở nên không thể chịu đựng nổi, các binh lính được gởi đến để chiếm giữ một hành lang; giả sử các hành động được thực hiện, cho dù có chủ định hay không, tạo ra một khả năng Đông Đức nổi dậy xảy ra. Chúng ta phân tích như thế nào về bản chất của sức ép đối với người Nga? Tôi nghĩ câu trả lời phần lớn sẽ là họ phải đổi mặt với nguy cơ của một cuộc chiến mà cả hai bên đều rất mong muốn

không xảy ra, nhưng cả hai bên có thể không có khả năng ngăn chặn. Một suy luận lý tính cho hành động trực tiếp, cho dù ở mức độ mà tự nó có thể thực hiện rất ít, có thể là việc tạo ra có chủ ý một nguy cơ mà chúng ta cùng chia sẻ với người Nga, trao cho họ quyền lựa chọn hoặc chấm dứt nguy cơ bằng cách hành động hoặc rút lui để đáp ứng các yêu sách của chúng ta.

Lẽ dĩ nhiên đây không phải là cách giải thích duy nhất cho hành động như vậy. Mỹ có thể chiến thắng về mặt quân sự nếu cuộc chiến dừng lại ở quy mô nhỏ, và việc người Nga mở rộng cuộc chiến đòi hỏi một bước nhảy không liên tục mà họ có thể bị ngăn chặn không thực hiện được do lo sợ gây ra một sự đáp trả không liên tục. Trong trường hợp đó, cuộc chiến tranh hạn chế ban đầu sẽ chưa đựng một mối đe dọa “ngăn chặn” chống lại việc mở rộng cuộc chiến. Mặc dù vậy, lý do quan trọng của việc tại sao đe dọa một cuộc chiến dù chỉ ở quy mô nhỏ lại có thể có hiệu quả chính là ở chỗ một cuộc chiến như thế hứa hẹn một sự gia tăng nhỏ nhưng đáng kể về khả năng xảy ra một cuộc chiến lớn, khả năng đủ nhỏ để người Nga tin là phương Tây có thể tự mình tạo ra, nhưng đủ lớn khiến cho nó không có lợi cho họ nếu để nó xảy ra.<sup>5</sup>

Cần chú ý cách giải thích này gợi ý rằng đe dọa về một cuộc chiến giới hạn có thể có tiềm năng ngay cả khi có ít kỳ vọng chúng ta sẽ giành chiến thắng. Trong các cách diễn đạt này, một cuộc chiến tranh cục bộ hạn chế không chỉ là hành

<sup>5</sup> Theo quan niệm của tác giả, việc đưa quân đội Hoa Kỳ đến Lebanon năm 1958 không chỉ vừa là nguy cơ vừa là thành công, mà còn là thành công chính nhờ nguy cơ - một nguy cơ mà đối phương có thể làm giảm nhẹ hoặc làm trầm trọng thêm tùy theo cách đáp trả của họ.

động quân sự cục bộ; nó chưa đựng một yếu tố “trả đũa” - không phải là *một ít* sự trả đũa, mà là một *khả năng* nhỏ của một cuộc chiến lớn.

## CHÍNH SÁCH BÊN BỜ VỰC CHIẾN TRANH

Lập luận của phần này dẫn đến một định nghĩa của chính sách bên bờ vực chiến tranh và khái niệm “bờ vực chiến tranh”. Bờ vực không phải là, theo quan điểm này, phần cạnh sắc của một vách đá nơi người ta có thể đứng chắc chắn, nhìn xuống, và quyết định là có nhảy xuống hay không. Bờ vực là một sườn dốc cong nơi người ta có thể đứng với nguy cơ trượt chân nào đó, sườn dốc trở nên dốc hơn, và nguy cơ trượt chân lớn hơn nếu người ta tiến đến phía vực thẳm. Nhưng sườn dốc và nguy cơ trượt chân đều khá bất thường; cả người đang đứng trên đó lẫn những người quan sát không thể hoàn toàn biết chắc là nguy cơ lớn như thế nào, và nó gia tăng như thế nào khi người ta tiến thêm vài bước. Bên miệng hố chiến tranh, người ta không làm khiếp đảm một kẻ thù được cột chặt với mình bằng cách tiến đến quá gần miệng vực đến mức nếu một người *quyết định* nhảy xuống thì anh ta có thể làm thế trước khi bắt cứ ai khác có thể ngăn chặn được anh ta. Tình trạng bên miệng hố chiến tranh bao hàm việc đứng trên sườn dốc nơi người ta có thể té cho dù có nỗ lực hết sức để tự cứu mình, kéo theo kẻ thù của mình.<sup>6</sup>

---

<sup>6</sup> Trẻ em hiểu điều này một cách hoàn hảo.

Chính sách bên bờ vực chiến tranh do vậy là việc thiết lập có chủ ý một nguy cơ chiến tranh có thể nhận ra, một nguy cơ mà một bên không hoàn toàn kiểm soát được. Nó là chiến thuật cố ý để cho tình hình vuột khỏi tầm tay đôi chút, chỉ bởi vì việc nằm ngoài tầm tay có thể là không chịu đựng nổi đối với bên kia và buộc bên kia cùng tham gia. Nó đồng nghĩa với việc quấy rối và hăm dọa kẻ thù bằng cách để anh ta đối mặt với một nguy cơ cùng chia sẻ, hoặc ngăn chặn anh ta bằng cách chỉ ra rằng nếu anh ta thực hiện một bước đi ngược lại, anh ta có thể gây khó khăn cho cả hai đến mức cả hai cùng trượt khỏi bờ vực cho dù có muốn hay không, và kéo anh ta theo.

Cần giải thích về quan niệm rằng chúng ta nên “giữ cho kẻ thù phải phán đoán” về cách đáp trả của chúng ta, đặc biệt là việc liệu chúng ta sẽ đáp trả hay không. Đôi khi người ta tranh luận rằng chúng ta không cần đe dọa kẻ thù bằng khả năng chắc chắn trả đũa hay khả năng chắc chắn kháng cự, mà chỉ cần làm cho kẻ thù sợ hãi với khả năng rằng chúng ta sẽ đánh trả. Quan niệm này có thể sai nếu nó có nghĩa là đương đầu với người Nga bằng một sự đáp trả có thể, tùy thuộc vào quyết định của Mỹ, cách này hay cách khác. Người Nga có thể phỏng đoán rằng sau biến cố Mỹ sẽ chọn lựa cách không đánh trả, đặc biệt là nếu Nga thực hiện sự gây hấn của họ bằng những hành vi lè ở mức độ vừa phải. Và nếu Mỹ không mong muốn sắp xếp để *phải* giáng trả, và thậm chí không mong muốn *nói* rằng Mỹ chắc chắn sẽ giáng trả, thì có thể là đường như Mỹ xác nhận sự hiểu biết của Nga về chọn lựa của Mỹ sẽ là gì nếu Mỹ dành cho mình lối thoát nào khác.

Do đó, nếu chúng ta e ngại rằng một cam kết tuyệt đối về một đe dọa có thể thất bại về mặt mục đích và ràng buộc chúng ta vào một hành động mà chúng ta không muốn bị ràng buộc vào, thì có thể chẳng còn gì nhiều để cứu vãn bằng cách cố gắng thuyết phục kẻ thù rằng chúng ta chỉ có thể quyết định thực hiện điều đó dù gì đi nữa.

Nhưng tình hình thì khác nếu Hoa Kỳ ở vào một tư thế mà hoàn toàn rõ ràng đối với người Nga là Hoa Kỳ có liên quan một cách đầy đủ đến nỗi mặc dù có thể có lối thoát nhưng Hoa Kỳ *có thể* sẽ không thoát ra. Nói rằng chúng ta có thể có hoặc có thể không trả đũa một cuộc xâm lược vào một nước trung lập nào đó, tùy thuộc vào việc nó phù hợp như thế nào lúc đó, và rằng chúng ta sẽ không để cho kẻ thù đưa ra quyết định này thay cho chúng ta, mà cũng không để cho kẻ thù dự đoán điều gì, có thể đương đầu với kẻ thù với điều *sẽ* xảy ra như một sự nghỉ binh. Nhưng tham gia như thế hoặc gần như thế ở một nước trung lập với quân đội hoặc với những cam kết khác mà chính chúng ta không hoàn toàn chắc chắn là liệu chúng ta có thể tránh được một cuộc chiến trong trường hợp xảy ra sự xâm chiếm không, có thể làm cho kẻ thù phải phỏng đoán.

Nói tóm lại, việc cố gắng giữ cho kẻ thù phải phán đoán có thể có ý nghĩa chừng nào mà chúng ta không cố gắng giữ cho kẻ thù phải phán đoán về động cơ của chúng ta. Nếu kết quả một phần được quyết định bởi những sự kiện và những quy trình mà rõ ràng ít nhiều nằm ngoài hiểu biết và kiểm soát của chúng ta, thì chúng ta tạo ra nguy cơ *thật sự* cho kẻ thù.

## **QUY TRÌNH KHÔNG HOÀN HẢO TRONG VIỆC RA QUYẾT ĐỊNH**

Điều quan trọng của mối đe dọa về việc người ta “có thể” trả đũa hay vội vã tiến hành chiến tranh - quyết định đó ít nhiều nằm ngoài sự kiểm soát của anh ta - chính là ý niệm rằng một số các quyết định quan trọng nhất của chính phủ được đưa ra theo một quy trình không hoàn toàn dự đoán được, không hoàn toàn “nằm trong tầm kiểm soát”, không hoàn toàn có chủ đích. Điều đó ám chỉ rằng một quốc gia có thể đi vào một cuộc chiến tranh lớn có phần nào đó không chủ động, theo một quy trình ra quyết định mà có thể gọi là “không hoàn hảo” theo nghĩa rằng việc đáp trả đối với các sự kiện ngẫu nhiên cụ thể nào đó không thể được dự báo một cách chính xác theo các tính toán trước, rằng việc đáp trả một sự kiện xảy ra tình cờ nào đó có thể tùy thuộc vào các quy trình ngẫu nhiên hay vô trật tự, rằng sẽ có những thông tin sai lạc, truyền đạt sai lạc, sự hiểu lầm, sự sử dụng quyền lực sai, nỗi hoảng sợ, hay những thất bại con người hay cơ học.

Ý tưởng này không phản ánh một quan điểm hoài nghi bất thường về quy trình ra quyết định. Trước tiên, các quyết định thật sự được đưa ra trên cơ sở chứng cứ không đầy đủ và sự cảnh báo mơ hồ; và thật vô lý khi phủ nhận về *nguyên tắc* khả năng xảy ra một hành động không thể cứu vãn được thực hiện do một báo động giả. (Hơn nữa, người ta không cần phải bị ám ảnh bởi khả năng xảy ra báo động giả để nhận ra rằng

có thể có những cấp độ mà bên dưới đó mỗi nguy hiểm cụ thể này không thể bị đẩy đến mà không gây ra những mối nguy hiểm khác lớn hơn!)

Thứ hai, chiến tranh có thể xảy ra do đôi bên kiên quyết bám lấy những vị trí không thể hòa giải mà từ đó cả hai đều không chịu xuống thang, đặc biệt là khi việc xuống thang đưa đến suy nghĩ, cho dù là tạm thời, về một điều kiện tổn hại đến mặt quân sự. Và cũng không phải là hoài nghi khi nhìn nhận rằng hai chính phủ có thể đánh giá sai về các cam kết của nhau.

Nhưng trong trường hợp thứ ba, ngay cả một chính phủ ngay ngắn, có những nhà lãnh đạo có trách nhiệm và tương đối bình tĩnh, vẫn nhất thiết phải có một hệ thống ra quyết định không hoàn hảo, đặc biệt là trong lúc khủng hoảng. Sở dĩ như thế là vì nhiều lý do, một trong số đó chính là trong bất cứ chính phủ nào không mang tính chất hoàn toàn độc tài tập trung thì luôn có nhiều người tham gia trong một quyết định, và họ không có cùng hệ thống giá trị, không có cùng đánh giá về các ý định của kẻ thù, không có cùng đánh giá về các khả năng quân sự. Một quyết định được đưa ra nhanh chóng trong lúc khủng hoảng có thể tùy thuộc vào việc ai có mặt ở đó, tùy thuộc vào việc liệu các nghiên cứu cụ thể nào đó đã hoàn tất chưa, vào sáng kiến và tính mạnh mẽ được thể hiện bởi các nhà lãnh đạo cụ thể nào đó cùng các cố vấn là những người đang phản ứng trước một sự khiêu khích hoàn toàn chưa có tiền lệ. Một số phần nào đó của quyết định có thể được đưa ra trên cơ sở phân cấp, và người được phân quyền ra quyết định không nhất thiết đưa ra được một quyết định giống như quyết định lê ra có thể được đưa ra bởi tổng thống hay thủ tướng hay nội các trên cơ sở tham vấn với

những nhà lãnh đạo quốc hội. Thậm chí cũng có thể có một số những mâu thuẫn cần thiết trong quy trình ra quyết định, chẳng hạn như những vấn đề về hiến pháp mà không thể giải quyết được trước gây ra khó khăn trong việc chuẩn bị đầy đủ cho những tình huống bất ngờ nào đó bởi vì sự cần thiết phải vi phạm luật pháp hay một tiền lệ chỉ được ngầm chấp nhận, chứ không được chuẩn bị cho việc đó. Cuối cùng, nhu cầu giữ bí mật đã đặt những hạn chế lên việc chuẩn bị trước cho những tình huống bất ngờ có thể được thực hiện.

Vì lý do này, không có cái gọi là một kế hoạch “cố định”, ý định hay chính sách của một chính phủ bao hàm mọi tình huống bất ngờ – ngay cả đối với tất cả những tình huống bất ngờ quan trọng có thể dự đoán trước. Các tính toán cân nhắc sẽ diễn ra thế nào, những lợi ích gì được mang lại, và trình tự ra quyết định tập thể có hiệu quả như thế nào trong các cuộc khủng hoảng trong tương lai, tất cả đơn giản là không thể xác định đầy đủ trước được.

Nếu trên tất cả những điều này chúng ta nhận ra rằng có những giới hạn con người thông thường về khả năng tri thức và xúc cảm của những người ra quyết định của chính phủ trong quá trình thực hiện những bước đi nguy hiểm bên bờ vực chiến tranh, thì cần phải nhận thức rõ ràng rằng sẽ đi đến tình trạng mà từ đó có vẻ như quốc gia có thể tự giải thoát một cách thành công nhưng có nguy cơ dễ nhận thấy nào đó rằng, dù nỗ lực trong phạm vi các giới hạn có được, quốc gia này có thể không thành công.

Người ta không kỳ vọng vào việc một chính phủ khêu gợi sự chú ý đến những thất bại của chính mình trong khía cạnh

này và thông báo cho kẻ thù rằng việc kiểm soát không hoàn hảo các hành động của chính mình là một phần hệ trọng của chiến lược của chính họ. Cũng có những lý do thuyết phục thuộc phạm vi quan hệ công chúng để không chỉ ra cho kẻ thù thấy rằng người ta thậm chí có thể bị ảnh hưởng nhẹ trước những sai lầm ghê gớm trong đánh giá hay báo động giả, hay người ta không biết chắc làm cách nào để thoát ra khỏi một tình huống rủi ro. Cũng dễ hiểu là một chính phủ tham gia vào một cuộc chiến tranh hạn chế không nói rằng họ đã bị lôi cuốn vào hành động quân sự này do nguy cơ có thể dẫn đến một cuộc chiến tổng lực. Vấn đề là những việc như thế này diễn ra mà không có trao đổi gì.

Nhưng quan điểm căn bản về sự đe dọa có yếu tố may rủi là quan trọng ngay cả nếu chúng ta không sử dụng nó một cách có ý thức, mà là ngầm ngầm. Trước tiên nó có thể được sử dụng chống lại chúng ta. Kế tiếp, chúng ta có thể đánh giá sai một số chiến thuật mà chúng ta có sử dụng nếu chúng ta không nhận ra được sự hiện diện của thành tố nguy-cơ-chiến-tranh-tổng-lực mà có thể là một phần đáng kể trong ảnh hưởng của chúng ta đối với kẻ thù, cho dù chúng ta chưa bao giờ đánh giá cao về nó. Nếu - chẳng hạn như - thành tố đó là một phần quan trọng trong vai trò của các lực lượng chiến tranh tranh hạn chế tại châu Âu, phân tích của chúng ta về vai trò đó có thể sai lầm nghiêm trọng nếu chúng ta không nhận ra thành tố đó. Ý tưởng thông thường rằng một sợi dây kích nổ có thể hoạt động mà cũng có thể không hoạt động, rằng người Nga có thể dự đoán nó hoạt động hoặc có thể không, đều mắc phải hai sai lầm cực đoan đơn giản về một chuỗi các khả năng phức tạp hơn nhiều.



**PHẦN IV**

**TẤN CÔNG BẤT NGỜ:  
MỘT NGHIÊN CỨU VỀ  
SỰ NGỜ VỰC LÃN NHAU**



## LO SỢ LẦN NHAU VỀ TẤN CÔNG BẤT NGỜ

Nếu tôi bước xuống cầu thang để tìm hiểu tiếng động phát ra trong đêm tối, với một khẩu súng trong tay, và chợt thấy mình đối mặt với một tên trộm cũng đang cầm một khẩu súng, thì sẽ có nguy cơ về một kết cuộc mà cả hai đều không mong muốn. Ngay cả nếu tên trộm muốn rút lui trong lặng lẽ, và tôi cũng mong muốn hắn làm như thế, thì vẫn có nguy cơ hắn có thể *nghĩ* rằng tôi muốn bắn, và hắn bắn trước. Tệ hơn nữa, có nguy cơ hắn có thể nghĩ rằng *tôi* nghĩ *hắn* muốn bắn. Hoặc hắn có thể nghĩ rằng *tôi* nghĩ *hắn* nghĩ rằng *tôi* muốn bắn. Và tương tự như thế. "Tự vệ" là điều mơ hồ, khi một bên chỉ cố ngăn chặn để không bị bắn trong khi tự vệ.

Đây là vấn đề tấn công bất ngờ. Nếu yếu tố bất ngờ mang lại lợi thế, thì tấn công trước là việc đáng làm. Nỗi lo sợ rằng phía bên kia có thể chuẩn bị tấn công do tin tưởng sai lầm rằng chúng ta chuẩn bị tấn công tạo ra cho chúng ta một động cơ để tấn công, và cũng bào chữa cho động cơ của bên kia. Nhưng, nếu những lợi ích giành được từ một cuộc tấn

công bất ngờ thành công thậ́m chí còn kém hơn việc không gây chiến, thì không có cơ sở “cơ bản” nào cho bên nào tổ chức tấn công. Tuy nhiên, có vẻ như là một sự kích thích nhỏ nhất nơi mỗi bên để đưa ra đòn đánh đầu tiên – một sự kích thích tự nó là quá nhỏ để có thể khởi động một cuộc tấn công – có thể trở nên phức tạp thông qua một quy trình những kỳ vọng tương hỗ, với động cơ tấn công tăng lên được tạo ra bởi các chuỗi liên hoàn theo kiểu “Họ nghĩ chúng ta nghĩ họ nghĩ chúng ta nghĩ... họ nghĩ chúng ta nghĩ họ sẽ tấn công; do đó họ nghĩ chúng ta sẽ; do đó họ sẽ; vì vậy chúng ta phải.”

Thật thú vị là vấn đề này, mặc dù nó bộc lộ mạnh mẽ nhất trong những tình huống thường được xem như là xung đột, như giữa phía Nga và Hoa Kỳ hay giữa tên trộm và tôi, thì tương đồng về mặt lý luận với vấn đề của hai hay nhiều bên không có sự tin tưởng lẫn nhau. Nếu mỗi bên bị kích thích bỏ trốn đi với những tài sản chung; nếu mỗi bên có một chút nghi ngờ rằng bên kia cũng đang suy nghĩ tương tự; nếu mỗi bên nhận ra rằng bên kia cũng có thể nghi ngờ, và có thể tự hoài nghi rằng mình là đối tượng bị nghi ngờ; thì chúng ta có một ma trận lợi ích tương đồng với ma trận của vấn đề tấn công bất ngờ. Nếu sức nóng nầm ở một số người trong đám đông, phần còn lại của đám đông có thể bị kích thích tách ra để tránh không phải la lối, và những người gặp nguy có thể bị kích thích la lên nhằm tự vệ. Do vậy cấu trúc trò chơi “tự vệ loại trừ” cũng tương tự như cấu trúc của trò chơi “tin tưởng hợp tác.”

Ý tưởng trực giác cho rằng các khả năng xảy ra ban đầu của một cuộc tấn công bất ngờ trở nên lớn hơn – có thể tạo

ra một hiệu ứng “đa tầng” - như một kết quả phức hợp trong nỗi lo sợ của mỗi người về việc mà người khác lo sợ là điều mà tôi muốn phân tích trong chương này. Đặc biệt hơn, tôi muốn phân tích liệu và làm thế nào hiện tượng này có thể phát triển thông qua một chuỗi tính toán *hợp lý* các khả năng xảy ra hoặc thông qua việc lựa chọn chiến lược *hợp lý*, bởi hai bên tham gia là những người biết đánh giá bản chất tình huống của mình. Chính ý tưởng trực giác, ngay cả nếu nó được hình thành một cách sai lạc, cũng thể là một hiện tượng có thật và là hành vi động lực; người ta có thể mơ hồ nhận thức rằng tình huống tự nó mang tính chất bùng nổ, và phản ứng bằng cách cho bùng nổ trước. Nhưng điều mà tôi muốn khám phá là liệu hiện tượng “các kỳ vọng phức hợp” có thể được trình bày như một quy trình quyết định hợp lý không. Liệu chúng ta có thể xây dựng một mô hình rõ ràng cụ thể cho tình huống trong đó hai người chơi duy lý là những nạn nhân của lối tư duy ẩn định các kỳ vọng của họ về nhau không?<sup>1</sup>

---

<sup>1</sup> Các nhà lý thuyết trò chơi thường nhìn nhận vấn đề này như là một trò chơi có tổng lợi ích khác không đối lập với điều mà, đối với các trò chơi có tổng lợi ích bằng không, được gọi là “trò chơi kép.” Phiên bản trò chơi có tổng lợi ích khác không được xem xét ở đây liên quan đến câu hỏi liệu có nên bắn không, chứ không phải khi nào thì nên bắn.

## *CHUỖI VÔ HẠN CÁC KHẢ NĂNG*

Chúng ta có thể bắt đầu bằng việc xây dựng bài toán như sau. Một người chơi chơi theo một loạt những khả năng xảy ra, một chuỗi tiềm năng vô hạn những khả năng đó. Khả năng ước tính đầu tiên,  $P_1$ , mà bên kia “thực sự” mong muốn tấn công, nghĩa là, bên kia sẽ tấn công ngay cả nếu chính bên đó không lo sợ về một cuộc tấn công. Khả năng thứ hai,  $P_2$ , nghĩa là bên kia *nghĩ* rằng tôi “thật sự” muốn tấn công anh ta, nghĩa là, tôi sẽ tấn công anh ta ngay cả nếu tôi không lo sợ một cuộc tấn công vào tôi. Khả năng thứ ba,  $P_3$ , trong đó *anh ta* nghĩ rằng *tôi* nghĩ *anh ta* “thật sự” sẽ; Khả năng thứ tư,  $P_4$ , trong đó *anh ta* nghĩ rằng *tôi* nghĩ *anh ta* nghĩ rằng *tôi* “thật sự” sẽ. Các khả năng thứ năm, thứ sáu và thứ bảy, v.v... được thiết lập bằng cách kéo dài chuỗi “anh ta nghĩ” và “tôi nghĩ” với một khả năng xảy ra riêng biệt gắn liền với mỗi thành viên trong chuỗi. Khả năng cuối cùng xảy ra việc anh ta sẽ tấn công được cho bởi công thức:

$$I - (I - P_1)(I - P_2)(I - P_3)\dots$$

Sự phiền toái đối với công thức này là không có điều gì tạo ra chuỗi đó. Mỗi một khả năng xảy ra là một dự kiến *bắt chốt*, phản ảnh các dữ liệu bổ sung liên quan đến cấu trúc thông tin cụ thể của một tình huống cụ thể. Chúng ta không thể, bắt đầu bằng một vài điều kiện trong chuỗi như dữ liệu, dự đoán phần còn lại cho đến vô hạn, hay xa đến đâu, và rồi xử lý một

cách toán học trên toàn bộ chuỗi. Số điều kiện trong một chuỗi như thế chỉ có thể nhiều bằng mức mà một bên tham gia có thời gian tính toán, hoặc khả năng hiểu biết giới hạn trong đầu, bởi vì anh ta phải tạo ra một điều kiện mới trong chuỗi bằng một quy trình dự đoán độc lập. Quả thật, chúng ta có thể thiết lập các trò chơi cụ thể với những cấu trúc thông tin mà sẽ tạo ra một công thức cho chuỗi - chẳng hạn như, một chuỗi các nắc trong một bánh quay ru-lét có thể xác định liệu người chơi kia có được thông tin về hệ giá trị "thật sự" của tôi hay không, liệu tôi có được thông tin liệu anh ta đã được thông tin hay không, liệu anh ta có được thông tin liệu tôi đã được thông tin về điều anh ta đã được thông tin hay không, và tiếp tục như thế - nhưng đây là những trò chơi đặc biệt, và có thể không làm sáng tỏ nhiều về tình huống chung mà chúng ta đang cố gắng nắm bắt. Điều chúng ta cần là một công thức của vấn đề cho phép chúng ta làm việc với một số lượng giới hạn các thông số phán đoán, có lẽ trình bày các điều kiện ban đầu hay "khách quan" trong một chuỗi, trong một bối cảnh mà tự động làm phát sinh các giá trị của bất cứ khả năng xảy ra bổ sung nào có thể được hình dung thông qua việc lặp đi lặp lại vô định chuỗi "Anh ta nghĩ tôi nghĩ." Chúng ta nên hình thành công thức cho vấn đề theo cách làm cho dự đoán của mỗi người cũng chính là một hàm dự đoán của bên kia.

## TRÒ CHƠI BẤT HỢP TÁC “CÓ THỂ GIẢI QUYẾT CHẶT CHẼ”

Ở lần thử đầu tiên, chúng ta có thể trao cho cho mỗi người chơi một tham số cơ bản đại diện cho khả năng *rằng anh ta sẽ tấn công nếu anh ta không nên tấn công*. Giá trị của những tham số này được biết rõ và được biết là cả hai người chơi cùng biết. Điều tôi muốn nói về “không nên” chứa đựng trong giả thiết hành vi gồm hai phần sau đây.

Phần đầu tiên của giả thiết hành vi của chúng ta là, nếu cả hai người chơi cùng nhận thức rằng một chính sách chung về không tấn công là điều tốt nhất trong tất cả các kết cuộc khả dĩ đối với cả hai bên, họ sẽ nhận ra “đáp số” này và chọn lựa cách tránh xa. Nếu, chẳng hạn như, ma trận lợi ích được trình bày như trong Hình 19, mỗi bên tự tin vào sự tin tưởng lẫn nhau và sẽ chọn một chiến lược mang lại cho đôi bên một kết cuộc khả dĩ tốt nhất. Điều này có vẻ như là một yêu cầu khá khiêm tốn về lý trí của hai người chơi.<sup>2</sup> (Tôi nghĩ rằng đây là một vấn đề đáng phải đặt câu hỏi, đặc biệt nếu tính vượt trội của chính sách cùng không tấn công so với chính sách đơn

---

<sup>2</sup> Trong thuật ngữ của Luce và Raiffa, nếu trò chơi bất hợp tác có một “đáp số đúng nghĩa”, thì “đáp số” đó được xem như chiếm ưu thế ở đây. *Game and Decisions* (Trò chơi và các Quyết định), trang 107. Thực ra điều kiện ở đây ít nhiều mạnh mẽ hơn, bởi vì đáp số được cả hai người chơi ưa chuộng so với tất cả các kết cuộc khác, chứ không chỉ so với tất cả các điểm cân bằng khác.

phương tấn công bất ngờ là nhỏ, hoặc quá nhỏ không đủ tạo cho cả hai bên hoàn toàn tự tin rằng họ hiểu biết về nhau. Và khả năng này - khả năng một ai đó sẽ bị kích thích để phá vỡ nguyên tắc nhằm đạt được sự an toàn, hoặc vì lo sợ rằng bên kia cũng cố tìm cách đạt được sự an toàn - là điều có thể xuất hiện trong phần hai của giả thiết hành vi, sẽ được trình bày ngay sau đây).

	I	II
i	0 0	-.5 .5
ii	-.5 -.5	1 1

H. 19

Phần hai của giả thiết là có một khả năng nào đó, gọi là  $P_r$ , cho bên R và  $P_c$  cho bên C, trong đó người chơi sẽ tấn công khi anh ta chọn (hoặc nên chọn) một chiến lược không tấn công, nghĩa là, quyết định của anh ta sẽ đối lập với phần một trong giả thiết của chúng ta. Điều này đã được đề cập trong ý nghĩa một người chơi có thể tấn công ngay cả khi anh ta "không nên" tấn công. Chúng ta sẽ để ngỏ việc tham số này thể hiện điều gì: có thể nó được xem như khả năng người chơi bất hợp lý, hoặc khả năng ma trận lợi ích được nhận thức sai và rằng anh ta "thật sự" ưa thích đơn phương tấn công bất ngờ, hoặc khả năng rằng ai đó sẽ phạm sai lầm và đưa lực lượng tấn công một cách khinh suất. Tham số này, đối với mỗi người chơi, mang tính chất "ngoại sinh" trong mô hình quyết định của chúng ta: đó là một dữ liệu được cung cấp từ bên

ngoài. Nó không được phát sinh từ sự tương tác giữa hai người chơi.

Hai tham số này,  $P_1$  và  $P_2$ , được giả thiết là khá dễ dàng được cả hai người chơi nhận thấy; không có gì là bí mật hay mang tính phỏng đoán liên quan đến các thông số đó. Giả thiết này có thể dường như gợi lên câu hỏi mà chúng ta đang tìm cách trả lời, nhưng thực ra thì không. Hai khả năng ngoại sinh về tấn công này tự chúng không chỉ ra khả năng gì khiến cho hai người chơi thực ra sê tấn công. Chúng chỉ là một yếu tố. Vấn đề là xem xét liệu với tất cả các nguồn gốc cơ bản của sự bất định thì dự đoán tương hỗ giữa hai người chơi có tạo ra động cơ bổ sung để tấn công hay không. Chúng ta phải đưa ít nhất *một số dữ liệu* vào trong vấn đề để cho kỳ vọng và phỏng đoán có thể hoạt động. Cách duy nhất để giữ các nguồn phán đoán đầu vào ở mức tối thiểu là làm cho hai tham số này hoàn toàn dễ nhận biết; nếu không chúng ta phải trình bày mỗi bên dự đoán như thế nào về các tham số đó, bên này đoán bên kia đoán về chúng như thế nào, bên này đoán bên kia đoán rằng chính bên này đoán về chúng thế nào, và cứ tiếp tục như thế. Một lần nữa chúng ta sẽ gặp phải một chuỗi vô hạn những chi tiết không dự tính trước, với khó khăn lớn thêm trong xử lý việc phân phối các khả năng xảy ra của việc phân phối các khả năng xảy ra. Cách duy nhất để thoát ra, và cung cấp một điểm khởi đầu để tính toán xem mỗi bên lo sợ bên kia lo sợ điều gì, là làm sao cho điều bất định cơ bản này của mỗi bên trở thành một vấn đề về ghi nhận. Điều chúng ta muốn biết là một nguồn bất định cơ bản "khách quan" đã tạo ra một siêu cấu trúc của những mối lo lắng chủ quan về mối lo lắng của nhau như thế nào.

Bây giờ chúng ta có một tình huống trông có vẻ như là nó sẽ tạo ra một tình huống tự vệ phức hợp mà chúng ta đã nói đến. Người chơi đầu tiên phải xem xét liệu khả năng tấn công của người kia có phải là chuyện nghiêm túc hay không; anh ta cũng phải cân nhắc rằng người kia cũng lo lắng theo cách tương tự. Ngay cả một người chơi mà khả năng tấn công “không hợp lý” của chính anh ta là con số không cũng phải cân nhắc rằng người chơi kia có thể tấn công không những theo cách không hợp lý mà còn xuất phát từ nỗi lo sợ rằng người đầu tiên, lo sợ sự tấn công của người thứ hai, có thể cố gắng tấn công trước để ngăn chặn trước. Do vậy có vẻ như là chúng ta có thể gặp phải sự phức hợp của các động cơ.

Nhưng chúng ta không gặp phải. Chúng ta không có được bất cứ loại hiệu ứng “đa tầng” thông thường nào từ tình huống này. Các khả năng xảy ra tấn công từ cả hai phía không tương tác để tạo ra một khả năng cao hơn, trừ phi khi chúng tạo ra sự chắc chắn. Nghĩa là, kết cuộc của trò chơi này, bắt đầu bằng các khả năng có hạn của việc tấn công “không hợp lý” đối với hai phía, không phải là việc mở rộng những khả năng xảy ra đó qua việc lo sợ tấn công bất ngờ; hoặc là cùng tấn công hoặc là không tấn công. Nghĩa là, nó là một cặp *quyết định*, chứ không phải là một cặp *khả năng xảy ra* liên quan đến hành vi.

Chúng ta xây dựng bài toán này bằng việc tính toán lại các phần thường trong ma trận ban đầu, sử dụng hai tham số trình bày khả năng xảy ra tấn công “không hợp lý.” Ô bên trái phía trên trong ma trận vẫn như cũ. Ô bên phải phía dưới có kết quả được tính toán lại, là mức trung bình cân bằng của cả bốn ô. Bởi vì, nếu cả hai người chơi cùng chọn chiến lược

không tấn công, thì có một khả năng tương đương với  $(I - P_i)(I - P_j)$  trong đó tấn công không xảy ra, một khả năng tương đương với  $P_i(I - P_j)$  trong đó R sẽ tấn công và C sẽ không, một khả năng tương đương với  $P_j(I - P_i)$  trong đó C sẽ tấn công và R sẽ không, và một khả năng tương đương với  $P_iP_j$ , trong đó cả hai sẽ cùng tấn công. Cùng một cách đó, các kết quả trong ô phía dưới bên trái là mức trung bình cân bằng của các kết quả trong cột bên trái; vì nếu C chọn tấn công, anh ta chắc chắn sẽ tấn công, trong khi nếu R chọn không tấn công, anh ta thực sự có hoặc không có với các khả năng lần lượt là  $P_j$  và  $(I - P_j)$ .

	I	II
i	0 0	- .4 .4
ii	.4 -.4	.64 .64

H. 20

Như vậy, với các khả năng tấn công không hợp lý tương đương với 0,2 cho mỗi người chơi, ma trận ban đầu của chúng ta sẽ tạo ra một ma trận chỉnh sửa như trong Hình 20.<sup>3</sup> Với các khả năng tấn công không hợp lý tương đương với 0,8 cho C và 0,2 cho R, chúng ta có Hình 21. Và với các khả năng 0,8 từng phần cho tấn công không hợp lý, chúng ta có Hình 22.

<sup>3</sup> Trong thực tế chúng ta xem xét những người chơi trong khi chọn lựa - theo ngôn ngữ của lý thuyết trò chơi - giữa một chiến lược "thuần" và một chiến lược "hỗn hợp" và chiến lược hỗn hợp được xác định bằng một tham số độc lập. (Dĩ nhiên người ta có thể pha trộn thêm nữa các chiến lược thuần với chiến lược hỗn hợp, nhưng trong ví dụ hiện tại không cần thiết phải làm như thế).

	I	II
i	0	- .1
ii	.4	.46
	- .4	- .14

H. 21

	I	II
i	0	- .1
ii	.1	.04
	- .1	.04

H. 22

Các khả năng tấn công không hợp lý như trong ma trận sửa đổi đầu tiên của chúng ta, nghĩa là các khả năng 0,2 cho mỗi người chơi, tỏ ra là vô hại. Nghĩa là, chúng vô hại *liên quan đến việc lựa chọn các chiến lược*. Chúng tạo ra một ma trận lợi ích mới mà vẫn có một "đáp số chặt chẽ" ở góc phải bên dưới. Giá trị của trò chơi bị giảm đối với mỗi người chơi, bởi vì không có cách nào thoát ra khỏi hai khả năng cơ bản đó; nhưng việc suy nghĩ về các khả năng đó vẫn chưa dẫn đến chỗ bực bội đối với chúng. Mỗi người chơi đã xem xét chúng một cách kỹ lưỡng, đã nhìn thấy rằng vẫn có một giải pháp chung được ưa thích là không tấn công, và qua giả thiết ban đầu đã chọn lựa chiến lược đó.

Ma trận cuối cùng trong số các ma trận sửa đổi của chúng ta, với một khả năng xảy ra 0,8 cho mỗi người chơi, thì đối

xứng và không ổn định; mỗi người chơi giờ đây mong muốn tấn công hơn là hy vọng vào việc cùng không tấn công, và mỗi bên đều biết rằng bên kia cũng muốn như thế. Đây là một tình huống éo le, tương ứng với "mâu thuẫn của người tù" khá quen thuộc trong lý thuyết trò chơi; giải pháp có hiệu quả duy nhất sẽ là một thỏa ước có tính chất ràng buộc để chọn không tấn công (mà vẫn khiến cho các bên chịu đựng một giá trị giảm bớt là 0,04), nếu các thỏa ước mang tính chất ràng buộc là khả dĩ về mặt thể chế và nếu cuộc chơi bị buộc phải hoãn lại để cho các bên chơi một cơ hội đạt được một thỏa ước như thế.<sup>4</sup>

Ma trận thứ hai trong số các ma trận sửa đổi này cũng không ổn định, mặc dù không theo cách đối xứng. Tính không hợp lý có thể có nơi người chơi C đòi hỏi người chơi R phải dự đoán trước điều đó bằng cách tấn công trước trong tự vệ; người chơi C, biết rõ điều này, cũng tấn công.<sup>5</sup>

<sup>4</sup> "Mâu thuẫn của người tù", trong lý thuyết trò chơi, nói đến một tập hợp các lợi ích cho cả hai người chơi những động cơ lớn - ngược với khi không có một thỏa ước có hiệu lực cưỡng chế - để chọn lựa các chiến lược mà khi kết hợp lại sẽ mang đến cho cả hai người chơi một kết quả kém hài lòng hơn so với trường hợp nếu cả hai có những sự chọn lựa đối nghịch. Tên gọi này xuất phát từ vấn đề của hai người tù, khi bị thẩm vấn tách riêng, có thể thú nhận một tội nhẹ hơn cùng phạm phải hoặc tố cáo lẫn nhau về một tội nặng, một người tố cáo được tự do trừ phi chính anh ta bị tố cáo, người hoặc những người bị tố cáo nhận những bản án nặng. Xem Luce và Raiffa, trang 94<sup>6</sup>.

<sup>5</sup> Một trường hợp hơi khác, và khá thú vị, xảy ra nếu chúng ta đặt  $P_r$  tương đương với 0,2 và  $P_c$  tương đương với 0,6. Ma trận điều chỉnh (cho riêng R) sẽ là:

0	.2
-.4	.12

R vẫn có một "chiến lược chi phối" về tấn công; anh ta thực hiện tốt hơn qua việc tấn công, bất chấp C thực hiện điều gì đi nữa. Nhưng trong trường hợp này, khác biệt với trường hợp minh họa trong Hình 19, anh ta đạt được ít hơn so với nếu cả hai bên cùng chọn lựa tấn công. Chính sự hiểu biết của C về chiến lược chi phối của R khiến cho cả hai chỉ nhận được con số không. Tính "không hợp lý" của C, thể hiện trong  $P_c$ , tạo cho R một động lực tấn công nhằm "tự vệ"; nhưng một yếu tố trong động lực đó - một sự "không thuần khiết" nho nhỏ trong động cơ tự vệ - là khả năng của R trong việc giành được yếu tố bất ngờ và như thế có khả năng thực hiện tốt hơn là mức đối phó với một cuộc tấn công sắp diễn ra. Nếu R không có khả năng gây bất ngờ đối với C, ngay cả khi anh ta nỗ lực, kết quả của anh ta trong ô bên phải phía trên của ma trận ban đầu sẽ là số không, chứ không phải 0,5, và ma trận điều chỉnh cho R sẽ là:

0	0
-4	.08

Điều này làm "tệ hại thêm" cả hai lợi ích cho R trong cột bên phải, nhưng ô phía trên tệ hơn ô phía dưới. Do đó, nó triệt tiêu động cơ tấn công của R, và C biết điều đó, do đó kết quả là cùng không tấn công. Như thế, không những nó giúp cho cả hai người chơi nếu người "không hợp lý" hơn không có khả năng tấn công; nó còn có thể giúp cho cả hai người chơi nếu "nạn nhân" không có khả năng giành được sự bất ngờ ngay cả trong việc "tự vệ". Điều kiện cho trường hợp đặc biệt này, xét đến các thông số được sử dụng trong đoạn kế tiếp của phần nội dung, là

$$I - h < P_c < I/(I + h).$$

Điểm này có thể làm cho trở nên tổng quát hơn. Giả sử giá trị của việc chiến thắng một cuộc chiến, biểu thị bằng  $h$ , có thể lớn hơn  $I$ ; nếu nó lớn hơn, và nếu một chiến lược tấn công luôn giành thắng lợi trong khi bên kia không tấn công, cả hai người chơi sẽ có chiến lược chi phối là "tấn công." Cả hai đều sẽ nhận được kết quả bằng không, trong khi họ có thể nhận được nhiều hơn nếu họ đã chọn cách ngược lại. Bây giờ giả sử rằng khả năng giành được yếu tố bất ngờ, và do đó chiến thắng, chỉ là  $Q$ , do đó giá trị kỳ vọng đạt được thông qua cuộc tấn công đơn phương chỉ là  $Qh$ . Nếu  $Qh$  nhỏ hơn  $I$ , chúng ta sẽ trả lại một ma trận với một đáp số chắc chắn được ưa thích hơn là cùng không tấn công; và, với khả

Các giới hạn đối với các giá trị của hai tham số của chúng ta,  $P_c$  và  $P_r$ , mà ngoài phạm vi đó chúng khiến cho tình huống trở nên không ổn định và kích thích việc cùng tấn công, là - xem  $h$  như giá trị đạt được qua tấn công bất ngờ đơn phương,  $-h$  là giá trị đạt được qua việc bị tấn công trong khi không tấn công,  $0$  là giá trị đạt được qua việc tấn công đồng thời, và  $I_0$  là giá trị của việc cùng không tấn công, đối với mỗi người chơi.

$$P_c < I - h_r$$

$$P_r < I - h_c^6$$

năng tấn công "không hợp lý", trò chơi sẽ ổn định nếu  $P_c < I - Qh$  và  $P_r < I - Qh$ . Giả sử rằng  $P_c$  và  $Q$  đáp ứng điều kiện đầu tiên trong số những điều kiện này: thế thì cả  $R$  lẫn  $C$  sẽ có được lợi thế nếu điều kiện thứ hai cũng được đáp ứng. Nếu  $P_c$  nằm ngoài khả năng điều khiển,  $R$  có lẽ sẽ mong muốn rằng  $Q$ , là khả năng riêng của anh ta trong việc gây bất ngờ cho kẻ thù, sẽ nhỏ hơn  $(I - P)/h$ . Chỉ khi nào như thế thì anh ta và  $C$  có thể nhận được nhiều hơn con số không. Nếu  $R$  có thể, bằng phí tổn của riêng mình, nâng cao hệ thống cảnh báo của "kẻ thù", hoặc nếu anh ta có thể làm giảm bớt khả năng gây bất ngờ bằng một cách dễ nhận biết, để giữ  $Q$  ở dưới mức giới hạn, anh ta có lẽ sẽ làm như thế. Nguyên tắc này cũng tương tự như nguyên tắc đối với hai người chơi, ít nhiều nghi ngờ lẫn nhau, giữ hai móc khóa riêng trên bàn hợp tác. Nếu một bên không thể có được chiếc móc khoá, bên kia sẽ cung cấp cho anh ta bằng phí tổn riêng của mình; chỉ bằng cách đó họ mới có thể cùng làm ăn với nhau.

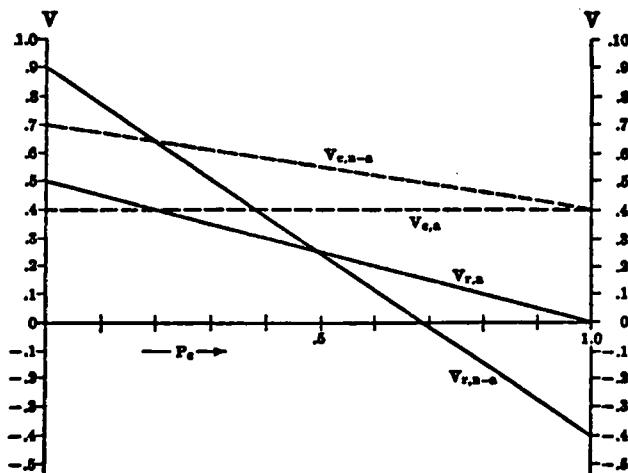
<sup>6</sup> Một công thức tổng quát hơn, bao hàm trường hợp bất đối xứng, và sử dụng  $R_{11}$ ,  $R_{12}$ ,  $R_{21}$ , và  $R_{22}$  để biểu thị các kết quả đối với  $R$  trong hàng I cột I, hàng I cột 2, và cứ như thế,

$$\frac{P_c R_{23} - R_{12}}{1 - P_c} < \frac{R_{12}}{R_{11} - R_{21}}$$

Tử số là "chi phí" của sai lầm trong việc tấn công; mẫu số là "chi phí" của sai lầm trong việc không tấn công. Có thể lưu ý rằng tiêu chuẩn vẫn như thế, như thế  $P$  và  $(I - P)$  là những khả năng xảy ra chắc chắn hơn là những khả năng xuất phát từ, và tuân thủ theo, một mô hình hành vi "duy lý".

Hình 23 minh họa điều xảy ra đối với “giá trị của trò chơi” cho mỗi người chơi, và cho mỗi chiến lược, vì ứng với mỗi một khả năng  $P$  sẽ dao động từ 0 đến 1.0. Đặt  $P_c$  bằng 0,2, và sắp xếp những giá trị của trò chơi ngược lại với  $P_c$  (dựa vào ma trận trong Hình 19), sẽ tạo ra các giá trị cho C và R như biểu đồ minh họa. Khi  $P_c = 0,5$ , trò chơi trở nên không ổn định, và giá trị của trò chơi bằng 0 đối với cả hai người chơi.

Việc trò chơi này không hoàn toàn tuân thủ theo ý nghĩa khởi đầu của “các khả năng phức hợp” được minh họa bằng việc chúng ta có thể phớt lờ tham số nhỏ hơn trong hai tham số nếu chúng không bằng nhau. Nếu cả hai đều nhỏ hơn mức giới hạn quan trọng, thì chúng có là gì đi nữa cũng không quan trọng; nếu một tham số nằm bên trên giới hạn chỉ một chút xíu nhỏ nhoi, thì cũng không có khác biệt gì nếu tham số kia bằng 0 hay 1.0. Chúng do vậy có thể có tiềm năng lớn hơn mức chúng đóng góp vào giá trị của việc cùng không tấn công, bởi vì chúng có thể khiến cho những người chơi chuyển từ một chiến lược không tấn công sang chiến lược tấn công. Nhưng họ sẽ làm như thế theo cách được ăn cắp ngã về không. Tính có lẽ tấn công hoặc bị giới hạn ở các khả năng ngoại sinh, hoặc trở thành điều chắc chắn.



H. 23

Giá trị của trò chơi đối với R và C, như một chức năng của  $P_c$ ;  $P_c$  bằng 0.2.  $V_{c,a} = 0.9 - 1.3P_c$ : giá trị của trò chơi đối với R, chiến lược cung không tấn công;  $V_{r,a} = 0.5 - 0.5P_c$ : giá trị của trò chơi đối với R, là bên tấn công trong khi C chọn cách không tấn công;  $V_{c,a-e} = 0.7 - 0.3P_c$ : giá trị của trò chơi đối với C, chiến lược cung không tấn công;  $V_{r,a-e} = 0.4$ : giá trị của trò chơi với C, là bên tấn công trong khi R chọn cách không tấn công.

## TRÒ CHƠI NHƯ MỘT CHUỖI CÁC NƯỚC ĐI NỐI TIẾP

Chúng ta có được cùng một kết quả nếu chúng ta thử một trò chơi với các nước đi nối tiếp cho ma trận mà chúng ta đã đang sử dụng. Giả sử R được cho phép tự do lựa chọn, tấn công hay không tấn công, trong khi C bị giới hạn phải chờ đợi;

và C có thể tấn công chỉ *sau khi* R đã có một cơ hội chọn lựa và thực hiện sự chọn lựa đó, và chỉ khi *nếu* R đã không tấn công. Nay giờ chúng ta tiếp tục xây dựng trò chơi này, bằng cách cho phép C có được một sự lựa chọn sớm hơn, trước cả sự chọn lựa của R, sao cho C được đi trước, sau đó đến R, sau đó cuối cùng là C. Sau đó chúng ta cho R một lượt đi sớm hơn, sao cho R chọn lựa, rồi C chọn lựa, sau đó R chọn, sau đó C chọn (chứng nào mà không bên nào chọn tấn công).

Trò chơi này đưa đến điều gì? Với nước đi cuối cùng của mình, C sẽ chọn không tấn công nếu ma trận giống như trong Hình 19; anh ta thực sự sẽ tấn công sau đó với khả năng  $P_c$ . Với nước đi cuối cùng của mình, R biết được điều mà C sẽ chọn, và thực hiện một sự chọn lựa đón đầu tùy thuộc vào  $P_c$ . Ở nước đi trước đó, C biết được điều mà R sẽ chọn, xem xét  $P_r$ , và thực hiện một sự chọn lựa đón đầu. Ở nước đi trước đó, R biết được điều mà C sẽ chọn để thực hiện trong cả hai lượt tiếp theo, và xem xét khả năng  $I - (I - P_c)^2$ , rằng C có thể tấn công tại một trong hai nước đi kế tiếp, và thực hiện sự chọn lựa của mình theo cách đón đầu. Và tiếp tục như thế. Nếu mỗi người chơi có n số lần di chuyển, với khả năng  $P_r$  hoặc  $P_c$  về việc tấn công không thích hợp tại *mỗi nước đi*, kết quả tùy thuộc vào liệu  $P_c = I - (I - P_c)^n$  và  $P_r = I - (I - P_r)^{n-1}$  có đáp ứng các điều kiện rút ra trước đó không. Nếu như thế, mỗi người chơi biết rằng người kia sẽ không chọn tấn công ở bước kế tiếp, và chính anh ta cũng sẽ không chọn khi đến lượt mình. Nhưng nếu  $P$  lớn hơn giới hạn đối với một trong các điều kiện anh ta sẽ mong muốn tấn công và bên kia biết được điều đó, do đó bắt cứ ai có lượt đi đầu tiên cũng sẽ tấn công ngay lập tức.

Nói cách khác, chúng ta đang làm phức tạp các khả năng xảy ra, nhưng vẫn có một hiệu quả được ăn cả ngã về không, mà không có việc người chơi nào phải *kết hợp* các tham số về tính không hợp lý nơi cả hai người chơi trong quy trình phức hợp. Hoặc là khả năng đối với ít nhất một trong số người chơi là đủ lớn và trò chơi đủ dài để khiến cho người chơi đầu tiên tấn công, hoặc là không có bên nào tấn công. Và, nếu chúng ta làm cho các khả năng tổng thể về việc tấn công không hợp lý độc lập với số lượt đi, bằng cách để cho khả năng ở mỗi lượt đi tương đương với  $I - (I - P)^{1/n}$ , sao cho tổng số phức hợp chỉ là  $P_c$  hay  $P_r$ , kết quả của trò chơi này thì *độc lập với số lượt đi*. Nếu chúng ta suy nghĩ về trò chơi này, thế thì, *tương tự* như tình huống anh-ta-nghĩ-tôi-nghĩ, với mỗi lượt đi tương ứng cho một vòng xoáy nghi ngờ, chúng ta có một mô hình trong đó các nỗi lo sợ lẫn nhau liên tiếp về việc mỗi bên lo sợ điều gì chẳng tạo ra khác biệt nào: hoặc có một cơ sở “khách quan” cho một trong các người chơi tấn công, hoặc tất cả cùng rút lui không tấn công.

### XEM XÉT LẠI VẤN ĐỀ

Điều tương tự như thế có vẻ là đúng sự thật nếu chúng ta trở lại với tên trộm dưới cầu thang. Nếu hắn cư xử “duy lý” như đã được xác định trong giả thiết về hành vi mà chúng ta nêu trên, hắn phải cân nhắc khả năng tôi sẽ bắn chỉ vì mong muốn như thế; và hắn phải cân nhắc rằng tôi có thể bắn anh ta nếu tôi nghĩ có khả năng cao là hắn sẽ bắn tôi chỉ vì mong

muốn như thế. Nhưng, nếu cả hai chúng tôi biết hai “khả năng” (ngoại sinh) cơ bản này là gì, chúng tôi không cần phải đi xa hơn. Hoặc là các khả năng cơ bản này là đủ để khiến cho ít nhất một trong chúng tôi phải bắn để ngăn chặn sự bất ngờ, và do đó khiến cả hai cùng bắn, như thế các nỗi lo sợ ở cấp độ thứ hai và cao hơn là thừa, hoặc là các khả năng đó tự chúng không đủ để khiến cho một trong hai chúng tôi phải bắn nhằm tự vệ, và chúng tôi biết điều đó và không có gì phải lo sợ bên ngoài phạm vi chính các khả năng ngoại sinh. Nếu cả hai chúng tôi có thể đơn giản thấy rằng không có ai hoàn toàn bị kích thích phải bắn chỉ vì lo sợ về khả năng ngoại sinh rằng bên kia “thật sự” muốn bắn, thế thì chúng tôi phải có thể nhìn thấy rằng không ai cần phải lo sợ một hành động ngăn chặn, rằng không có ai sau đó cần lo sợ rằng người kia lo sợ điều đó, và tiếp tục như thế.<sup>7</sup>

Nhưng một tình huống khác diễn ra nếu tôi bắn không phải do tính toán mà do hồi hộp. Giả sử rằng sự hồi hộp nơi tôi tùy thuộc vào việc tôi sợ hãi thế nào, và sự sợ hãi nơi tôi tùy thuộc vào mức độ tôi suy nghĩ rằng anh ta có thể bắn tôi; và giả sử anh ta cũng hành động theo một cách tương tự. Thế thì khi tôi cân nhắc khả năng ngoại sinh rằng anh ta có thể bắn tôi chỉ vì mong muốn thế, điều đó khiến tôi hồi hộp; sự hồi hộp này nâng cao khả năng rằng tôi có thể bắn anh ta cho dù tôi không thích làm thế. Anh ta thấy được sự hồi hộp nơi tôi và chính anh ta cũng trở nên hồi hộp; điều đó khiến tôi

---

<sup>7</sup> Ví dụ như, nếu cả hai có thể thông tin cho nhau và kiểm tra được sự hiểu biết của người kia, họ có thể đạt được một thỏa thuận không chính thức không chọn việc bắn mà không có lý do gì để đánh lừa - giả sử rằng hai tham số cơ bản này là rất rõ ràng đối với cả hai.

càng hoảng sợ, và tôi càng có xu hướng muốn bắn. Anh ta thấy được tính chất gia tăng này trong sự hồi hộp của tôi, và cũng gia tăng sự hồi hộp tương ứng, và càng làm cho tôi thêm hoảng sợ; và khả năng rằng tôi sẽ bắn lại càng gia tăng. Bây giờ chúng ta có thể biểu thị sự hồi hộp của mỗi người như là một chức năng trong sự hồi hộp của người kia, và khả năng xảy ra việc bắn như một chức năng của sự hồi hộp, và chúng ta có được một cặp biểu thức sai số đồng thời mà có vẻ sẽ đưa đến một cách chính xác loại hiện tượng mà chúng ta đã khởi đầu nghiên cứu.<sup>8</sup>

Và lý do họ làm là vì mô hình này không liên quan đến *tiêu chí ra quyết định*; nghĩa là nó không bao hàm một giả thiết hành vi mà có thể cho chúng ta biết được một người sẽ chọn cái nào trong hai chiến lược. Thay vào đó, "mô hình hồi hộp" của chúng ta là mô hình trong đó người ta phản ứng trước nỗi lo sợ tấn công bằng cách *thay đổi khả năng xảy ra* là chính họ sẽ tấn công. Chỉ duy nhất qua cách này, bằng việc đối phó với *khả năng* trong quyết định của một người, chứ không phải đối phó với *quy tắc* của quyết định - nghĩa là, không phải bằng mô hình trong đó người chơi tính toán chiến lược tốt

---

<sup>8</sup> Có sự bất đối xứng quan trọng trong vấn đề hình thành ở đây. Chúng ta đã cho phép khả năng xảy ra rằng một người có thể bắn khi mà lẽ ra anh ta không nên và người kia biết được điều đó - trường hợp "hồi hộp" - nhưng không có khả năng một người có thể không bắn khi anh ta cần phải bắn, và người kia biết được điều đó. (Có thể có rủi ro nào đó rằng tên trộm có viên đạn thối trong súng hay hắn quên nạp đạn cho khẩu súng, và tôi có thể biết rằng có một trường hợp như thế, hắn có thể biết rằng tôi biết điều đó, và tiếp tục như thế). Khả năng này có vẻ hẳn nhiên mang tính chất giúp ổn định, có xu hướng làm giảm bớt khả năng xảy ra một quyết định tấn công cũng như khả năng ngoại sinh của việc tấn công do vô ý hay không hợp lý.

nhất của mình và đi theo nó - chúng ta mới có thể có được hiện tượng “cùng làm trầm trọng lẫn nhau” mà tôi đã mô tả ở phần đầu của chương này.

Lúc này, liệu điều đó có nghĩa rằng hiện tượng của chúng ta không phải là hiện tượng có thể được trình bày bởi những người chơi duy lý, quyết đoán hay không? Làm sao chúng ta có thể hình dung ra một người chơi phản ứng trước một thay đổi trong môi trường của mình, hay trước một thông tin nhỏ, bằng cách *quyết định* rằng anh ta sẽ làm điều gì đó “ít nhiều có khả năng xảy ra hơn” so với trước kia? Một người chơi duy lý có thể hồi hộp, trong trường hợp đó lý thuyết của chúng ta mang tính chất sinh lý hơn là tri thức; nhưng liệu chúng ta có thể hình dung được việc một người chơi duy lý nhìn một cách khác về tên trộm và thay đổi sự điều chỉnh trên bánh quay roulette của anh ta được không?<sup>9</sup>

Tuy nhiên, các quyết định cá nhân và tập thể có thể khác nhau trong khía cạnh này. Chúng ta có thể nghĩ đến một quyết định tập thể thông qua biểu quyết, với những thành viên khác nhau có những hệ thống giá trị khác nhau và do đó có những ngưỡng khác nhau trong việc phản ứng trước khả năng bị tấn công, sao cho quy mô biểu quyết tấn công sẽ là một chức năng của khả năng dự đoán được về việc bị tấn công. Nếu việc biểu quyết cũng tùy thuộc nhiều vào các yếu tố may rủi, chẳng hạn như số người vắng mặt vào ngày biểu quyết, khả năng của việc giành đa số phiếu biểu quyết tấn

<sup>9</sup> Chú ý rằng tính hợp lý thông thường đối với một chiến lược hỗn hợp - nghĩa là, đối với việc điều chỉnh lại một cách thích hợp bánh quay roulette của một người để có quyết định - không can hệ gì đến trường hợp hiện nay.

công trở thành một chức năng gia tăng của khả năng xảy ra trong quyết định riêng của kẻ thù, mà đến lượt nó lại là một chức năng của khả năng của người chơi tập thể đầu tiên. Do vậy chúng ta có thể có được hiện tượng mà chúng ta muốn cho những người chơi “duy lý” nếu chúng ta coi là hợp lý một người chơi tập thể có những giá trị khác biệt và một hệ thống biểu quyết.

Tuy nhiên, có một cách để áp dụng mô hình của chúng ta ngay cả cho một người chơi duy nhất, quyết đoán, và duy lý. Nó có thể mang tính phổ quát rộng rãi trong các vấn đề hợp tác và tấn công bất ngờ. Và nó trực tiếp bao hàm một phần đáng kể của vấn đề thực sự trong việc gây bất ngờ quân sự, nghĩa là việc phụ thuộc quyết định vào một hệ thống báo động không hoàn hảo, và khả năng xảy ra các sai lầm “loại-1” và “loại-2” trong quy trình ra quyết định.

## *HÀNH VI KHẢ NĂNG TẠO RA BỞI MỘT HỆ THỐNG BÁO ĐỘNG KHÔNG HOÀN THIỆN*

Giả thiết rằng nguy cơ bị tấn công bất ngờ có thể được giảm bớt bằng việc sử dụng một hệ thống báo động. Nhưng hệ thống báo động này không phải là không thất bại. Một hệ thống báo động có thể mắc sai lầm theo hai cách: hoặc là nó có thể khiến chúng ta làm một chiếc máy bay tấn công với một con hải âu, để rồi không làm gì cả, hoặc là nó có thể khiến chúng ta làm một con hải âu với một chiếc máy bay tấn công,

để rồi khiêu khích kẻ thù bằng một cuộc tấn công do khinh suất. Cả hai khả năng làm lẩn này có thể được giảm bớt về mặt giả thiết bằng cách chi thêm tiền và tăng cường công nghệ cho hệ thống. Nhưng, với một khoản chi phí xác định nào đó, thông thường thì đúng đồi với các tiêu chí ra quyết định rằng việc thắt chặt các tiêu chí liên quan đến một loại sai lầm lại làm nới lỏng ra các tiêu chí liên quan đến loại sai lầm khác. Việc đòi hỏi chứng cứ về một cuộc tấn công sắp xảy ra trước khi "trả đũa" đồng nghĩa với việc đòi hỏi thêm chứng cứ rằng chúng thực sự là những chú hải âu để khỏi phải cho máy bay của chúng ta xuất kích.

Nhưng lúc này chúng ta có thể có một mô hình về một người ra quyết định hợp lý là người phản ứng lại một dự báo về khả năng bị tấn công *không phải* bằng một *quyết định hành động* hay *không hành động rõ ràng*, mà *bằng việc điều chỉnh khả năng anh ta có thể tấn công một cách sai lầm*. Nếu phản ứng của từng người chơi trước một nguy cơ gia tăng về tấn công bất ngờ là thiên về tấn công khinh suất, thì khả năng tấn công của từng người chơi lúc này chính là một hàm gia tăng của khả năng tấn công của người kia.<sup>10</sup> Một hệ thống báo động như thế chính là cơ chế đối ứng, hợp lý của sự hồi hộp nơi chúng ta trong việc đối mặt với tên trộm.

Để xây dựng một mô hình như thế (theo tính đối xứng cho đơn giản) chúng ta lại có thể dùng *h* biểu thị cho giá trị của

<sup>10</sup> Như thể hiện dưới đây, điều này không *nhất thiết* là như thế; nếu nguy cơ gia tăng về việc bị tấn công gắn liền với tính dễ bị tấn công giảm bớt của kẻ thù trước một cuộc tấn công bất ngờ, thì phản ứng của một bên hoàn toàn có thể theo hướng ngược lại với mô tả trong phần này.

việc “chiến thắng” một cuộc chiến,  $-h$  biểu thị cho việc “thua” trong một cuộc chiến,  $o$  là giá trị kỳ vọng của việc tấn công đồng thời (cơ hội 50-50 thắng hoặc thua), và  $I.o$  là giá trị của việc không tiến hành chiến tranh. (Lần này chúng ta có thể để cho  $h$  lớn hơn  $I$ , chừng nào mà  $(I - R)h$  trong ma trận bên dưới vẫn còn nhỏ hơn  $I$ . Nhưng nếu “việc chiến thắng” phải trả giá đắt,  $h$  sẽ là một số thập phân nhỏ). Chúng ta giả thiết rằng việc tạo bất ngờ thành công mang lại chiến thắng trong cuộc chiến; “sự bất ngờ thành công” có nghĩa là một bên tấn công trong khi bên kia không và hệ thống báo động của bên kia đánh lừa anh ta. Hãy xem  $R$  biểu thị cho tính đáng tin cậy của hệ thống báo động của một bên, nghĩa là, khả năng một cuộc tấn công nếu nó xảy ra sẽ được xác định và yếu tố bất ngờ được ngăn chặn. Khi đó ma trận lợi ích sẽ như trong Hình 24.

Khả năng một người chơi sẽ tấn công khi mà anh ta không nêu, nghĩa là anh ta sẽ tiến hành khi mà sự lựa chọn “hợp lý” của anh ta phải là ngược lại với việc tấn công (trong nghĩa đã được sử dụng trước đó), sẽ bao gồm hai phần. Phần trước tiên, biểu thị bằng  $A$ , là khả năng ngoại sinh của việc tấn công không hợp lý; nó loại trừ khả năng xảy ra một cuộc tấn công bị châm ngòi bằng báo động giả. Khả năng xảy ra một cuộc tấn công thông qua báo động giả được biểu thị bằng  $B$ . Như thế, hai loại sai lầm trong hệ thống báo động được thể hiện bằng  $B$  và  $(I - R)$ ; và đặc điểm chính của mô hình là  $B = f(R)$ ,  $f'(R) > 0$ . Nghĩa là nếu chúng ta càng giảm bớt  $(I - R)$  như là một nguồn gốc của sai lầm, chúng ta càng gia tăng  $B$ , và ngược lại.

		I	II
		0	-(1-R_e)h
i		0	(1-R_e)h
ii	(1-R_r)h		1
	-(1-R_r)h		1

H. 24

Sự lựa chọn chiến lược của mỗi người chơi liên quan đến cặp giá trị của  $B$  và  $R$  mà sẽ làm giảm tối thiểu mất mát dự kiến của anh ta, nghĩa là, làm gia tăng tối đa giá trị kỳ vọng của trò chơi đối với anh ta. Để  $V$ , biểu thị cho giá trị kỳ vọng của của trò chơi đối với  $R$ , vấn đề của hệ thống báo động đối với  $R$  là chọn cặp giá trị đối với  $R$  và  $B$ , tương hợp với  $B = f(R)$ , mà sẽ làm tối đa hóa<sup>11</sup> biểu thức

$$\begin{aligned}
 V_r &= (I - P_c)(I - P_r) + P_r(I - P_c)h(I - R_c) \\
 &\quad - P_c(I - P_r)h(I - R_r) \\
 &= (I - A_c)(I - B_c)(I - A_r)(I - B_r) \\
 &\quad + (A_c + B_c - A_r B_c)(I - A_c)(I - B_r)h(I - R_c) \\
 &\quad - (A_c + B_c - A_c B_r)(I - A_r)(I - B_c)h(I - R_r).
 \end{aligned}$$

Thêm vào đó, theo bảng phân tích ma trận trước đó,  $R$  cần phải xem xét ma trận lợi ích "điều chỉnh" mà có kết quả từ

<sup>11</sup> Để thuận tiện cho minh họa, chúng ta giả thiết rằng một cuộc tấn công bất cẩn xuất phát từ báo động sai cũng chính là một loại tấn công được trù liệu trước, với cùng khả năng giành được yếu tố bất ngờ. Trong khi đó, chúng ta đang bỏ qua chiều kích thời gian của  $B$ , mà có lẽ cần phải được nghĩ đến như khả năng xảy ra báo động giả trên mỗi đơn vị thời gian, trong khi  $(I - R)$  là khả năng xảy ra sai lầm đối với mỗi cuộc tấn công sắp đến, và  $A$  rất có thể có ít nhiều cả hai yếu tố. Như thế, lĩnh vực thời gian được xem như là cố định trong mô hình này.

việc sử dụng những giá trị “tối ưu” của  $R$ , và  $B$ , cùng với các giá trị quan sát được (hay giá trị tối ưu kỳ vọng) của  $R_c$  và  $B_c$ , để xem liệu chính sách cùng không tấn công có còn là kết quả cùng mong muốn hay không. Các điều kiện để có được cùng mong muốn không tấn công, với các hệ thống cảnh báo được điều chỉnh tối ưu, sẽ là:

$$P_c = (A_c + B_c - A_c B_c) < \frac{I_c h(I_c R_c)}{I_c h(R_r - R_c)}$$

$$P_r = (A_r + B_r - A_r B_r) < \frac{I_r h(I_r R_r)}{I_r h(R_c - R_r)}$$

Với tính chất đối xứng, các mẫu số của các số hạng bên phải chỉ là  $I$ .

Thực tế, như chúng ta sẽ gặp dưới đây, việc khảo sát lần hai này có thể là không cần thiết; đối với một số giả thuyết về hành vi, việc điều chỉnh “tối ưu”  $R$  và  $B$  (đối với bất cứ giá trị nào ít hơn  $R = I$ ) đòi hỏi rằng các điều kiện để ổn định ma trận chỉnh sửa phải được đáp ứng.

Tuy nhiên, sẽ vẫn cần phải xác định những người chơi sẽ hành xử như thế nào. Nói chung, chúng ta có thể làm cho một trong ba giả thuyết, tùy thuộc nhiều hay ít vào sự khác biệt giữa “hành vi tham số”, thành một “trò chơi ngầm”, và một “trò chơi mặc cả.”

## SỰ ĐIỀU CHỈNH ĐỘNG (HÀNH VI THAM SỐ)

Trước tiên chúng ta có thể thử giả thiết rằng mỗi người chơi nhìn nhận khả năng bị tấn công là một tham số chứ không phải biến số trong hàm tổn thất của anh ta, và thực hiện tương tự với tính đáng tin của hệ thống báo động của đối thủ anh ta. Nghĩa là, anh ta trực tiếp *quan sát* các giá trị  $B$  và  $R$  của đối thủ, và lựa chọn cặp giá trị cho  $B$  và  $R$  của riêng mình nhằm giảm tối thiểu các tổn thất kỳ vọng của anh ta. Giả thiết này có xu hướng làm cho lựa chọn  $B$  của từng người thành một hàm gia tăng của khả năng xảy ra rằng bên kia sẽ tấn công. (Nó chỉ “có xu hướng”, bởi vì có khả năng sự thay đổi tương ứng trong  $R$  của bên kia mang lại một sự khuyến khích bù trừ, như sẽ được trình bày dưới đây). Nếu chúng ta nghĩ hai người chơi liên tục điều chỉnh các giá trị  $B$  và  $R$  của họ, mỗi người đều theo dõi các giá trị  $B$  và  $R$  của người kia, nhưng luôn luôn phản ứng theo tham số đối với khả năng hiện tại về việc bị tấn công và không dự đoán hành vi của bên kia như một hàm số của họ, chúng ta có được một hệ thống “bội số” động đơn giản - ổn định hay bùng nổ tùy thuộc vào các giá trị tham số và hình thái của  $f$ . Chúng ta có thể thể hiện giá trị tối ưu  $B$  của mỗi người chơi như là một hàm số của người kia, giải quyết hai biểu thức, và diễn dịch các điều kiện ổn định để đạt cân bằng. Chúng ta cũng có thể tính toán “các bội số” liên quan đến các thay đổi  $B$  và  $R$  của mỗi người chơi sang các chuyển đổi hàm số  $f$  hay sang các thay đổi các tham số  $A$ .

Một cách hiển nhiên, để tìm hàm số “hành vi-tham số” cho người chơi R chúng ta tối đa  $V_r$ , liên quan đến  $R_r$ , tùy thuộc vào  $B_c = f(R_r)$ , nhưng xem  $B_c$  và  $R_r$  là cố định. Sử dụng công thức cho trước đối với  $V_r$ , chúng ta có

$$f' = \frac{P_c h(I - R_r)}{(I - P_c)(I - h(I - R_c)) - P_c h(I - R_r)}$$

và, sao cho  $h(I - R_c) < I > h(I - R_r)$ ,  $f'' > 0$ .

Bởi vì  $f'$  được giả thiết là số dương, mẫu số cũng phải dương nếu  $V_r$  được tối đa hóa với  $R < I$ ; nhưng điều kiện đòi hỏi mẫu số phải dương cũng chính là điều kiện mà  $P_c$  phải đáp ứng để cho người chơi R vẫn mong muốn cả hai cùng không tấn công. Như thế, nếu cả hai người chơi có những sự điều chỉnh tối ưu với  $R < I$ , các giá trị tối ưu đó của  $R$  và  $B$  cũng là tất yếu tương thích với mong muốn chung cùng không tấn công.

Mối quan hệ của  $B_r$  đối với  $B_c$ , theo giả thiết hành vi này, nghĩa là, việc giảm bớt hàm số phát sinh mà tạo giá trị  $B$  tối ưu của  $R$  đối với các giá trị cho trước của  $B_c$ , thì có được bằng cách lấy vi phân hai về của phương trình trên:

$$\begin{aligned} \frac{\frac{dB_r}{dB_c}}{f'' dB_c} \quad (\text{theo hàm số hành vi của người chơi } R) &= \frac{dB_r}{dB_c} \frac{dR_r}{df'} \frac{df'}{dB_c} \\ &= \frac{f'}{f'' dB_c} = \frac{f'}{f''} \left( \frac{\partial f'}{\partial B_c} + \frac{\partial f'}{\partial R_c} + \frac{dR_c}{dB_c} \right) \\ &= \frac{f'}{f''} - \frac{\partial f'}{\partial B_c} + \frac{\partial f'/\partial R_c}{\varphi} \end{aligned}$$

trong đó  $B_c = \phi(R)$  biểu thị cho hàm tương ứng của người chơi C.

Bởi vì  $\partial f'/\partial R_c$  là số âm, các giá trị nhỏ của  $\phi'$  có thể khiến cho trị số  $dB_c/dB_r$  của người chơi R mang giá trị âm; điều đó là có thể qua việc nâng cao “chi phí” của việc cuộc tấn công khinh suất ở mức đủ lớn hơn việc gia tăng nguy cơ bị tấn công. Nói cách khác,  $B_r$  là một hàm số không chỉ của  $B_c$  mà còn là của  $\phi'(B_r)$  nữa;  $B_r$  có xu hướng được gia tăng khi  $B_c$  gia tăng nhưng giảm xuống khi  $R_c$  gia tăng, trong khi  $B_c$  và  $R_c$  cùng gia tăng khi chúng ta xem xét rút trực  $B_c$  ra.

Tình trạng cân bằng ổn định đòi hỏi rằng trị số  $dB_c/dB_r$  của người chơi R và trị số  $dB_r/dB_c$  của người chơi C có một kết quả thấp hơn 1, nghĩa là, với việc đánh giá  $B_r$ , theo trục tung và  $B_c$  theo trục hoành, đường cong của C sẽ phân cắt đường cong của R từ bên dưới. Biểu thức “bội số” tổng quát liên quan đến những thay đổi của  $B$  và  $R$  sang việc thay đổi các hàm số (hay sang việc thay đổi các giá trị của A) có chứa 1 trừ đi kết quả này trong mẫu số.

Như đã nhận định trước đó, mẫu số trong biểu thức cho  $f'$  biến mất, và  $R_r$ ,  $B_r$ , và  $f'$  tăng mạnh, khi  $h$  tiến đến gần điều kiện cần thiết cho một ma trận bất ổn định. (Thực sự, sự ổn định của trò chơi ma trận này, khác với sự ổn định của sự cân bằng hành vi-tham số, không phải là một khái niệm thích hợp cho giả thiết hành vi-tham số; suy nghĩ về ma trận và dự đoán trước hành động của người khác nghĩa là dự đoán hành động của anh ta, chứ không phải quan sát hành vi đó và điều chỉnh theo).

Cũng có thể lưu ý rằng người chơi R có thể bỏ qua A, trong

các tính toán của anh ta. Nó xuất phát từ công thức để đạt  $B$ , và  $R$ , tối ưu. Một cách trực giác, có điều này là vì tính chất ngẫu nhiên duy nhất trong đó hoặc là giá trị của  $R$ , hoặc là giá trị của  $B$ , có thể tạo ra sự khác biệt cũng chính là tính chất ngẫu nhiên khiến cho  $R$  không phát động cuộc tấn công "vô lý"; nếu anh ta tấn công, thì  $B$  và  $R$  không liên quan đến anh ta. (Tuy nhiên,  $A$ , có ảnh hưởng đến điều kiện đối với một ma trận ổn định, bởi vì nó tham gia vào điều kiện mà  $P$ , phải đáp ứng. Do đó trong việc dự đoán sự điều chỉnh của  $C$ ,  $R$  sẽ phải đưa vào xem xét  $A$ , Nhưng "dự đoán" hành vi của  $C$ , hơn là chỉ liên tục quan sát  $R_c$  và  $K_c$ , sẽ làm cho hành vi của  $R$  trở nên phi tham số, mâu thuẫn với giả thuyết lúc này. Nếu người chơi  $R$  đang xem xét giá trị của việc sử dụng tiền bạc để nâng cao hệ thống bảo động của mình,  $A$ , sẽ tác động tới việc tính toán bởi vì nó tác động tới khả năng hệ thống này tạo ra khác biệt; việc xem xét này không nằm trong mô hình hiện nay).

## TRÒ CHƠI NGẦM

Chúng ta có thể thực hiện một giả thiết hành vi khác, mà có thể dẫn đến cùng một kết quả. Thay vì giả sử rằng đương nhiên mỗi người chơi thấy được  $R$  và  $B$  của người kia được điều chỉnh như thế nào, và phản ứng lại; chúng ta có thể giả sử rằng mỗi người chơi biết được các cơ hội công nghệ của người kia - mối quan hệ hàm số giữa  $R$  và  $B$  đối với người chơi kia - nhưng không thể quan sát một cách đáng tin cậy xem

người kia đã điều chỉnh các giá trị  $R$  và  $B$  như thế nào. Nghĩa là, mỗi người hiểu được cơ chế trong hệ thống báo động của người kia, nhưng không bao giờ có thể chắc chắn được rằng người kia đã đưa ra những chỉ dẫn về diễn giải chứng cứ như thế nào khi chúng xuất hiện trên hệ thống – nghĩa là quy tắc ra quyết định của người kia. Giả định này đưa ra cho chúng ta một trò chơi bất hợp tác, trong đó mỗi người chơi phải chọn một giá trị cho  $B$  (nghĩa là, cho  $R$ ), mà không biết được người kia đã chọn giá trị gì nhưng biết được ma trận lợi ích của người kia.

Trong trường hợp này, chúng ta có một ma trận lợi ích với một “điểm cân bằng” chính xác tại điểm đó, nếu có, thì giả thiết hành vi-tham số đã tạo ra sự cân bằng ổn định<sup>12</sup>. Nói cách khác, cái đã là “đáp số” theo giả thiết hành vi-tham số thì vẫn là một ứng viên cho việc được xem là một “đáp số” trong hình thái bất hợp tác của trò chơi. (Trong cả hai trường hợp không nhất thiết điểm cân bằng phải là duy nhất. Nếu nó không phải, thì giả thiết đầu tiên làm cho kết quả tùy thuộc vào các điều kiện ban đầu và “các cú sốc”; giả thiết thứ hai có xu hướng làm phức tạp vấn đề trí tuệ trong việc nhận dạng các chiến lược “lời giải”.)

Lời giải này, dĩ nhiên, là không đủ cho cả hai người chơi. Nó là một ví dụ về “mâu thuẫn của người tù” đã được đề cập ở trên (trang 214); việc gia tăng hổ tương trong các giá trị của

---

<sup>12</sup> Một *điểm cân bằng*, trong lý thuyết trò chơi, là một cặp chiến lược cho cả hai người chơi sao cho mỗi chiến lược là tối ưu tương ứng với chiến lược kia. (Có thể có nhiều điểm cân bằng như thế).

B đã đơn thuần gia tăng khả năng tấn công từ mỗi phía.<sup>13</sup> Có những giá trị nhỏ hơn cho các trị số  $B$  mà sẽ làm cho cả hai bên khá hơn; và nếu các khả năng xảy ra tấn công lén lút có chủ đích bởi hai người chơi là bằng nhau (các trị số  $A$  bằng nhau), một thỏa thuận về việc không có hệ thống báo động nào, nghĩa là không có khả năng xảy ra báo động giả, sẽ là sự mặc cả được ưa thích cho cả hai bên nếu họ bị giới hạn ở những cuộc mặc cả tạo ra cho họ các hệ thống báo động giống nhau.<sup>14</sup>

---

<sup>13</sup> Các nhà kinh tế học có thể nhận ra rằng tình huống này gợi đến việc hai nhà sản xuất cùng phân bổ các nguồn tài nguyên sản xuất giới hạn của mình cho hai loại hàng hóa. Một loại hàng hóa, "an ninh chống lại báo động giả", bao hàm các yếu tố kinh tế bên ngoài; loại kia, "an ninh chống lại sự bất ngờ" bao hàm các yếu tố phi kinh tế bên ngoài.

<sup>14</sup> Nếu các trị số của  $A$ ,  $B$  và  $R$  là bằng nhau,  $V$ , và  $V$  bằng  $(I - P)^2$ , mà có giá trị cực đại khi  $B = 0$ . (Nếu  $B$  có một giá trị cực tiểu nào đó lớn hơn 0, chúng ta có thể xác định nó từ  $A$ ). Nếu các trị số của  $B$  và  $R$  là bằng nhau, còn các trị số  $A$  không bằng nhau,

$$dV/dB = -2(I - B)(I - A_c)(I - A_f) + (A_c - A_f)(h/f'),$$

mà có thể có giá trị dương với  $A_c$  lớn hơn  $A_f$  và  $f'$  nhỏ. Trong trường hợp này một trong hai người chơi - người với  $A$  nhỏ hơn - có mong muốn có được hệ thống báo động nào đó ngay cả nếu nó phải có chung cho cả hai bên, so với việc không có gì cả; nhưng nó bao hàm các giá trị nhỏ hơn cho  $B$  và  $R$  so với cái mà hành vi tham số (hay một trò chơi bất hợp tác) sẽ đưa đến, như có thể thấy qua việc đặt phần diễn giải trên bằng 0 và so sánh công thức phát sinh cho  $f'$  với công thức tương ứng với hành vi tham số.

## *TRÒ CHƠI MẶC CẢ*

Nếu chúng ta xem xét khả năng hai người chơi thương lượng để giảm bớt tính nhạy cảm của các hệ thống báo động của họ, vì lợi ích của việc cùng giảm bớt  $B$  với cái giá là các trị số  $R$  nhỏ hơn, và giả sử rằng một thỏa thuận như thế là có thể thực hiện, thì không có cách nào thật sự thuyết phục để rút ra một giải pháp độc nhất mà không đòi hỏi việc đề cập sâu hơn về một khung mặc cả. Nếu *đáp số* phải là đối xứng và *trò chơi* cũng đối xứng, nghĩa là, nếu họ thương lượng về *một cặp giá trị chung cho  $R$  và  $B$* , kết quả sẽ như vừa được đề cập - các giá trị là 0 cho  $B$ , ngay cả nếu điều này nghĩa là 0 cho  $R$ , không hề có hệ thống báo động nào cả. Nếu các hệ thống báo động cần phải đồng nhất, thì sẽ có một sự khác biệt quan trọng nào đó giữa các khả năng cơ bản về việc xảy ra tấn công đột kích có chủ đích đối với cả hai người (giữa  $A$ , và  $A_j$ ) mà bên ngoài phạm vi đó cần phải có một khoản chi phí bổ sung để đạt được một thỏa thuận về việc hủy bỏ các hệ thống báo động.

Nhưng, nói chung, điều này trở thành một vấn đề mặc cả mở rộng. Nó thậm chí còn mở rộng hơn so với công thức đề cập lúc này đề cập đến, bởi vì những người chơi có thể không chỉ xử lý các giá trị của  $R$  và  $B$  mà dĩ nhiên còn có thể đe dọa tấn công trực tiếp lúc này, hoặc sử dụng những dàn xếp về mặt định chế để xác định các giá trị của  $A$ .

Có một khó khăn trong việc tuân thủ bất cứ thỏa thuận nào

về việc giảm bớt các giá trị của  $R$  và  $B$  vì lợi ích chung; đó là các giá trị của mỗi bên đối với  $R$  và  $B$  có thể là không quan sát được. Chúng tùy thuộc - ít nhất ở một tầm mức quan trọng - vào các chuẩn mực mà sẽ chi phối các quyết định trong tương lai, chứ không chỉ duy nhất vào cơ chế vật chất, có thể quan sát được của một hệ thống báo động. Chúng tùy thuộc vào việc người ta sẽ chờ đợi bao lâu để có thể "chắc chắn", và vào các rủi ro nào mà người ta sẽ chấp nhận trong một tình huống khẩn cấp.Thêm vào đó, việc phá vỡ thỏa thuận, nếu nó có dẫn đến điều gì, thì chính là dẫn đến chiến tranh; do đó những lời buộc tội qua lại và những kiện cáo về tổn thất là nằm ngoài vấn đề nếu mô hình của chúng ta thể hiện chiến tranh tổng lực hơn là một vụ việc xâm phạm biên giới hay một vụ khiêu khích nhỏ của một bên đối với bên kia.

Rất có thể rằng  $R = B = 0$  là có thể quan sát được về mặt định tính - việc "thiếu vắng" về mặt vật chất của một hệ thống nào đó. Ngay cả khả năng xảy ra này cũng không có, như một hệ thống có thể thi hành, nếu ma trận là không ổn định với  $R = 0$ , nghĩa là, với  $h > I$ . Trong trường hợp đó, "nguy cơ" nào đó trong hình thái của  $B$  là cần thiết để đặt giá trị  $R$  an toàn trong phạm vi  $h(I - R)$  là nhỏ hơn  $I$ .

Cũng có thể khó khăn để có được một thỏa thuận mà công khai thừa nhận các trị số  $A$ , bởi vì nó có thể khó khăn về mặt chính trị để thừa nhận rằng trị số  $A$  của một bên là lớn hơn 0.

Những người chơi do vậy có thể bị thúc đẩy dựa vào các dàn xếp mà hoặc là làm giảm bớt một cách có thể quan sát được khả năng gây bất ngờ của họ, hoặc nâng cao một cách có thể quan sát được các đường cong biến đổi của họ và của

lẫn nhau liên quan  $R$  với  $(I - B)$ . Cả hai bên có thể, chẳng hạn như, đồng ý chi nhiều hơn cho hệ thống báo động, để làm cho nó có hiệu quả hơn; và bên có tiềm lực tài chính lớn hơn có thể mong muốn tài trợ cho việc nâng cao hệ thống báo động của đối phương, thay vì để mặc nó trong một tình trạng mà hoặc là làm trầm trọng thêm cảm giác mất an toàn của bên kia, hoặc là làm cho chính bên này dễ bị tổn thương trước báo động giả. Một thỏa thuận về việc sắp xếp các lực lượng không có tiềm năng gây ra tấn công bất ngờ, nhưng thay vào đó là tính dễ bị tổn thương cao hơn trước tấn công bất ngờ, có vẻ như là điều được đề nghị. Nghĩa là, thay vì làm cho  $R$  và  $B$  trở thành các điều khoản trong một thỏa thuận, họ có thể bị tính không quan sát được của  $R$  và  $B$  buộc phải sử dụng các hàm số  $f$  và  $\phi$ , xem như mỗi một trong số các hàm số này bao hàm cả hệ thống báo động riêng của một bên lẫn lực lượng tấn công của kẻ thù (hay của đối tác). (Tuy nhiên, cũng cần lưu ý ở đây rằng “các cải tiến” trong các hệ thống báo động – các biến đổi của các hàm số  $f$  và  $\phi$  theo hướng  $B$  nhỏ hơn đối với một mức độ cho trước của  $R$  và ngược lại – không phải là có tác dụng ổn định trong mọi trường hợp. Những gì làm tăng chi phí biến của  $R$  có thể dẫn đến các giá trị cao hơn cho  $B$ ; đó sẽ là các cải tiến quá mức xét từ quan điểm của cả hai người chơi hợp lại, tương tự như một “cải tiến” trong ma trận mâu thuẫn của người tù mà tăng lợi ích của mỗi người chơi cho các chiến lược bất hợp tác).

Công thức trò chơi mặc cả cũng thích hợp với việc phân tích chiến thuật mặc cả. Ví dụ như, nếu một bên hành động theo tham số và bên kia biết được và tính toán đến điều đó,

bên đầu tiên đi theo một "hàm phản ứng"<sup>15</sup> mà sẽ được đưa vào trong công thức cho V của bên kia, là bên sẽ cố tối đa hóa yếu tố V này. Nói chung, việc phân tích "nước đi chiến lược" thuộc loại đã được thảo luận trong các Chương 2, 5 và 7, là có liên quan đến phiên bản trò chơi tấn công bất ngờ này, trò chơi theo nguyên tắc đối tác.

### *NHIỀU HƠN HAI NGƯỜI CHƠI*

Một biến thể thú vị của bài toán này sẽ xảy ra nếu số bên tham gia chơi được加倍, hoặc nếu một bên thứ ba được đưa vào như là một đại diện tự chủ. Trong phạm vi mà việc tấn công từ các khu vực khác phải được lường trước, thì động cơ cùng giảm bớt các hệ thống báo động cũng giảm theo. Tuy nhiên, vẫn đúng là bất cứ hai bên tham gia nào trong một trò chơi lớn hơn có thể tìm thấy lợi thế nào đó trong việc cùng gia tăng các hệ thống báo động của mình, theo hướng làm giảm bớt nguy cơ báo động giả, bằng cách đưa vào xem xét "các yếu tố phi kinh tế ngoại vi" lẫn nhau mà họ bỏ qua không cẩn nhắc khi hành động theo tham số. Hai người canh gác có vũ trang tuần tra cùng một tòa nhà, mỗi người tùy thuộc vào một sự xui khiến nào đó phải bắn khi bắt gặp ai đó, sẽ thấy dễ chịu hơn nếu họ có thể tìm ra cách thức nào đó nhằm đạt được một thỏa thuận có thể thực hiện được để bớt sẵn sàng bắn khi bắt gặp, để giảm bớt khả năng bắn lẫn nhau. (Thực

<sup>15</sup> So sánh với chú thích 24, Chương 5, liên quan đến khái niệm "hàm phản ứng."

ra, vấn đề hai người canh gác là một cách thể hiện mô hình ban đầu của chúng ta, nếu chúng ta để cho các tham số ban đầu là  $P_c$  và  $P$ , thể hiện cho các khả năng tương quan mà một người bị bắt gặp trong đêm tối là tên trộm thay vì là người canh gác kia. Chúng ta phải đưa ra một tình trạng không chắc chắn nào đó nơi hành vi của tên trộm - nghĩa là, cho hắn ta tham gia trò chơi như là một người chơi thứ ba duy lý đang cố gắng dự đoán các quyết định của những người kia - để có thể tăng thêm các yếu tố phức tạp vào những gì chúng ta đã có).<sup>16</sup>

---

<sup>16</sup> Arthur Lee Burns, ở trường Đại học Quốc gia Úc, đã bàn luận một số bài toán thú vị về một thế giới ba người hay nhiều hơn. Sự khiêu khích có chủ đích để gây ra chiến tranh giữa hai người do một bên thứ ba ác ý là một khả năng xảy ra khi một hành động công khai xuất phát từ một nguyên nhân mơ hồ có thể được đưa vào trong mô hình nghi ngờ lẫn nhau; và việc phân tích có thêm sự phong phú bổ sung khi một bên xem xét các hệ thống báo động mà, vì các lý do kỹ thuật hoặc bởi lý do ràng buộc lẫn nhau, cho phép một hoặc hai bên tham gia chính có thể chứng kiến điều gì đang xuất hiện trên màn hình ra-đa của đối phương. Xem thêm "Cơ sở của một cuộc chiến do kích thích" (Rational of Catalytic War) (Trung tâm Nghiên cứu Quốc tế, Bản ghi nhớ nghiên cứu số 3; Đại học Princeton, 1959).

## TẤN CÔNG BẤT NGỜ VÀ GIẢI TRỪ QUÂN BỊ

"Giải trừ quân bị" bao trùm một sự đa dạng của sự phối hợp, một số thuộc phạm trù lý trí, một số thuộc lĩnh vực tình cảm, hợp tác giữa các kẻ thù tiềm năng để giảm bớt khả năng xảy ra chiến tranh hay giảm bớt quy mô cũng như tính khốc liệt của nó. Hầu hết các đề xuất đều dựa trên một tiền đề là việc giảm bớt số lượng và tiềm năng của vũ khí, đặc biệt là với các loại vũ khí "tấn công" hay các loại vũ khí mà có thể có chủ đích hay vô ý gây ra thương vong cho dân thường hay tạo ra sức tàn phá lớn, nhằm hỗ trợ cho mục đích này. Có sự phối hợp toàn diện; có sự phối hợp nhằm xác định các lĩnh vực cụ thể trong đó lợi ích chung là rất rõ ràng, khi mà nhu cầu về sự tin tưởng lẫn nhau là tối thiểu, và khi mà một bước khởi đầu đáng kể có thể được thực hiện, mà nếu thành công, sẽ là một bước đầu tiên tiến đến giải trừ quân bị toàn diện hơn. Trong số các phối hợp ít toàn diện hơn này, các biện pháp an toàn chống lại sự tấn công bất ngờ đã trở nên ngày càng nổi bật hơn kể từ đề xuất "bầu trời mở" lần đầu của Tổng thống Dwight D. Eisenhower vào năm 1955.

Việc tập trung vào sự tấn công bất ngờ đã không phản ánh việc từ bỏ mối quan tâm nhiều tham vọng hơn vào việc giải giới. Thay vào đó nó thể hiện triết lý về việc lựa chọn một lĩnh vực có cơ may thành công nhất, nhằm thiết lập một truyền thống nào đó trong việc hợp tác thành công. Việc tìm kiếm các biện pháp an toàn chống lại tấn công bất ngờ đã được xem xét rộng rãi, trong chính phủ Mỹ cũng như tại các nước khác, không phải như một *biện pháp thay thế* cho giải trừ quân bị, mà như một *loại* giải trừ quân bị và là một bước đi khả dĩ hướng tới giải trừ quân bị rộng hơn.

Tuy nhiên, mặc dù các kế hoạch hóa giải việc tấn công bất ngờ có thể là giải trừ quân bị kiểu truyền thống, chúng thể hiện một điều gì đó mang tính cải tiến. Đè xuất ban đầu về bầu trời mở là một điều không chính thống trong quan niệm cơ bản của nó cho rằng vũ khí tự chúng không mang tính khiêu khích chừng nào chúng rõ ràng vẫn còn nằm trong kho - chừng nào vai trò của chúng vẫn chỉ là phòng ngừa thay vì tấn công. Đè xuất này cũng không chính thống trong việc nhắc nhở mạnh mẽ rằng, cũng quan trọng như việc giữ bí mật không cho kẻ thù biết và trong một số vấn đề để khiến cho kẻ thù phải suy đoán về các kế hoạch của chúng ta, nó còn có thể thậm chí quan trọng hơn khi thấy rằng kẻ thù *không* còn được tùy nghi phỏng đoán về các ý định của chúng ta đối với việc tấn công bất ngờ chống lại họ *nếu trong thực tế chúng ta không có kế hoạch thực hiện tấn công như thế*. Chúng ta quan tâm đến việc không chỉ bảo đảm cho chính mình qua trực quan rằng kẻ thù không chuẩn bị một cuộc tấn công chống lại chúng ta; chúng ta còn quan tâm đến việc tránh an

kẻ thù thông qua trực quan của họ rằng chúng ta không chuẩn bị một cuộc tấn công có chủ đích chống lại họ.

Như Tướng Leslie R. Groves nhận xét trong một bài diễn văn, "Nếu Nga biết rằng chúng ta không tấn công trước, thì điện Kremlin sẽ rất ít có xu hướng tấn công chúng ta... Sự miễn cưỡng của chúng ta trong việc tấn công trước là một bất lợi về mặt quân sự đối với chúng ta; nhưng, như một nghịch lý, nó cũng là một tác nhân ngăn chặn cuộc xung đột thế giới hôm nay."<sup>1</sup> Chúng ta sống trong một kỷ nguyên trong đó động cơ mạnh mẽ ở cả hai bên - có lẽ là động cơ chủ lực - để khởi đầu một cuộc chiến tổng lực bằng một cuộc tấn công bất ngờ chính là nỗi lo sợ trở thành người đi sau tội nghiệp vì không thể tấn công trước. "Tự vệ" trở nên phức tạp một cách đặc biệt nếu chúng ta phải lo lắng về việc kẻ thù tấn công chúng ta nhằm ngăn chặn chúng ta tấn công họ nhằm ngăn chặn họ tấn công chúng ta... Vấn đề tấn công bất ngờ, khi được xem xét như là một vấn đề nghiêm túc lẫn nhau và làm trầm trọng thêm tình trạng "tự vệ", gợi ý rằng không chỉ có các bí mật mà chúng ta không mong muốn giữ kín, mà còn cả những khả năng quân sự mà chúng ta không mong muốn có được.

Đĩ nhiên, thậm chí sẽ là tốt hơn nếu phía bên kia cũng không có những thứ đó. Do vậy có thể có những lợi thế khi suy nghĩ về vấn đề tấn công bất ngờ như một vấn đề thích hợp để đàm phán.

Sự đổi mới trong phương pháp tiếp cận vấn đề tấn công bất ngờ còn đi xa hơn nữa. Nó có liên quan đến việc kế hoạch

---

<sup>1</sup> The New York Times, ngày 29-12-1957, trang 20.

được thiết kế để bảo vệ cái gì và các loại vũ khí nào được nhìn nhận. Một kế hoạch chống tấn công bất ngờ có mục đích không chỉ làm cho việc *tấn công* trở nên khó khăn mà còn làm giảm bớt hay triệt tiêu lợi thế trong việc tấn công *trước*. Nó phải giả thiết rằng nếu lợi thế của việc tấn công trước có thể bị triệt tiêu hoặc bị giảm bớt đáng kể, thì sự kích thích để tấn công *nếu có* sẽ được giảm bớt.

Người ta thường thừa nhận rộng rãi rằng Mỹ có *khả năng* quân sự hùng như đù để xóa sổ Liên Xô, và ngược lại. Và người ta cũng thừa nhận rộng rãi rằng, nếu một trong hai bên tấn công bên kia bằng một đòn tấn công nguyên tử lớn, thì đất nước bị tấn công sẽ có một *sự kích thích* mạnh mẽ để đánh trả lại bằng một sức mạnh tương ứng hoặc lớn hơn. Nhưng, nếu một bên có thể xóa sổ bên kia, thì chuyện bên nào tấn công trước có ý nghĩa gì? Câu trả lời, dĩ nhiên, là người Mỹ không đặc biệt quan tâm đến việc sống lâu hơn Liên Xô một ngày; Mỹ lo lắng về việc liệu một cuộc tấn công bất ngờ có thể mang lại những triển vọng *phá hủy năng lực trả đũa* mà không bị ngăn chặn bởi một mối đe dọa trả đũa hay không. Đây không phải là việc *năng lực* hiện có của Mỹ có thể hủy diệt Nga để ngăn chặn một cuộc phản công của Nga nhắm vào Mỹ, mà là *năng lực* của Mỹ để trả đũa sau khi chính Mỹ bị tấn công. Mỹ phải giả thiết rằng một cuộc tấn công đầu tiên từ phía Nga, nếu nó diễn ra, sẽ nhắm vào chính *năng lực* mà Mỹ dựa vào để trả đũa.

Có sự khác biệt giữa một sự cân bằng sức hủy diệt trong đó *bất cứ* bên nào có thể xóa sổ bên kia với một tình trạng cân bằng trong đó *cả hai* bên có thể thực hiện điều đó bất chấp việc bên nào tấn công trước. Nó không phải là loại “cân bằng”

- sự ngang bằng đơn thuần hoặc sự đối xứng của tình huống - mà tạo ra sự ngăn chặn lẫn nhau; nó là *tình trạng ổn định* của sự cân bằng. Sự cân bằng chỉ ổn định khi mà không bên nào, trong việc tấn công trước, có thể hủy diệt năng lực trả đũa của bên kia.

Sự khác biệt giữa sự cân bằng ổn định và cân bằng không ổn định được minh họa bằng một loại vũ khí tấn công khác mà không có biện pháp tự vệ nào từng được thiết kế để chống lại nó.<sup>2</sup> “Luật cân bằng” của vùng Viễn Tây cổ cho phép *một* trong hai người có thể giết người kia; nó không bảo đảm rằng cả hai đều bị giết. Những hậu quả căng thẳng của hệ thống vũ khí này có thể thấy hầu như mỗi tối trên truyền hình. Lợi thế của việc bắn trước làm gia tăng động cơ bắn. Như cách mà người sống sót diễn đạt, “Hắn ta chuẩn bị bắn tôi để tự vệ, do đó tôi phải bắn hắn ta để tự vệ.” Hoặc, “Hắn ta, nghĩ rằng tôi chuẩn bị bắn hắn ta để tự vệ, nên đã chuẩn bị bắn tôi để tự vệ, nên tôi phải bắn hắn ta để tự vệ.” Nhưng nếu cả hai cùng được đảm bảo là sẽ sống đủ lâu để bắn trả với mục tiêu không bị suy suyển, thì sẽ chẳng có lợi thế gì trong việc vội vàng bắn và chẳng có lý do để sợ rằng người kia sẽ cố làm điều đó.

Ý nghĩa đặc biệt của tấn công bất ngờ do vậy nằm ở khả năng dễ tấn công của các lực lượng trả đũa. Nếu các lực lượng này tự chúng không dễ bị tấn công – nếu mỗi bên tự tin rằng

---

<sup>2</sup> Một sử gia quân sự, bình luận về cái gọi là “sự thật lịch sử” theo đó chưa bao giờ có một loại vũ khí mà con người không thể thiết kế một loại vũ khí đối trọng hay một biện pháp tự vệ chống lại nó, đã nhắc nhở chúng ta rằng “sau năm thế kỷ sử dụng các loại vũ khí cầm tay với tên lửa có thuốc nổ... vẫn chưa tìm được câu trả lời thích hợp cho viên đạn đó” (Bernard Brodie, *The Absolute Weapon* (Vũ khí tuyệt đối) [New York, 1946], trang 30-31).

lực lượng của mình có thể vượt qua một cuộc tấn công, nhưng đồng thời biết rằng mình không thể hủy diệt năng lực đánh trả của bên kia - sẽ chẳng có động cơ nào đủ mạnh để một bên tấn công trước. Và lại càng có ít nhu cầu phải phản ứng nhanh trước một động thái có thể là một báo động giả.

*Do vậy các kế hoạch ngăn chặn tấn công bất ngờ có mục tiêu tức thời là đảm bảo sự an toàn của vũ khí hơn là sự an toàn của con người.* Các kế hoạch tấn công bất ngờ, tương phản với các loại đòn xuất về giải giáp, được dựa trên *sự ngăn chặn* như là sự bảo vệ cơ bản chống lại tấn công. Chúng tìm kiếm việc hoàn thiện và ổn định sự ngăn chặn lẫn nhau - nâng cao tính thống nhất của các hệ thống vũ khí riêng biệt nào đó. Và một cách chính xác chính các loại vũ khí có sức hủy diệt lớn nhất đối với con người là đối tượng mà một kế hoạch chống tấn công bất ngờ tìm cách bảo vệ - các loại vũ khí trả đũa, các loại vũ khí có vai trò trừng phạt hơn là để chiến đấu, để làm tổn thương kẻ thù sau đó, chứ không phải để tước vũ khí của kẻ thù trước đó. Một loại vũ khí mà chỉ có thể làm tổn thương *con người*, và không thể có khả năng làm tổn hại sức tấn công của bên kia, là loại vũ khí hoàn toàn mang tính phòng vệ: nó không cung cấp cho người sở hữu nó động lực nào để tấn công trước. Chính loại vũ khí được thiết kế hay được sử dụng để phá hủy các mục tiêu “quân sự” - để tìm kiếm tên lửa hay các máy bay ném bom của kẻ thù - mới có thể khai thác lợi thế tấn công trước và do đó xui khiến người ta hành động như thế.

Với việc xác định vấn đề tấn công bất ngờ là khả năng dễ bị tấn công các lực lượng trả đũa của mỗi bên trước yếu tố bất ngờ, chúng ta đang ở một điểm mà tại đó các biện pháp chống

lại tấn công bất ngờ khác biệt một cách sâu sắc với các ý nghĩa thông thường hơn của việc giải trừ quân bị. Chúng ta cũng đang ở tại nguồn của một số bất thường và nghịch lý mà sẽ phải đổi mới nếu chúng ta cần phải nhận ra những ưu điểm và khiếm khuyết của các kế hoạch cụ thể nào đó và cần hiểu biết các nước đi đằng sau chúng. Cũng ngay tại điểm này, chúng ta bắt đầu nêu câu hỏi liệu các kế hoạch chống lại tấn công bất ngờ có thể được xem như "các bước đầu tiên" hướng tới giải trừ quân bị toàn diện hơn theo nghĩa truyền thống không, hay thay vào đó là không tương xứng với các hình thức giải trừ quân bị khác. Liệu các biện pháp để bảo vệ kế hoạch tấn công bất ngờ (SAC) có thể được xem như các bước đầu tiên hướng tới việc hủy bỏ chúng không? Liệu chúng ta có thể trước hết áp dụng các biện pháp hợp tác để hoàn thiện và bảo vệ năng lực của mỗi bên trong việc trả đũa với quy mô lớn, vì lợi ích của sự ngăn chặn lẫn nhau, và thực hiện điều đó như một bước hướng tới việc từ bỏ đe dọa trả đũa quy mô lớn từ một thế giới căng thẳng và phức tạp không?

Hoặc thay vào đó chúng ta có nên nhìn nhận các biện pháp bảo vệ chống lại tấn công bất ngờ như một sự nhượng bộ - một thừa nhận ngầm về "sự ngăn chặn lẫn nhau" là nguồn gốc tốt nhất của sự ổn định quân sự mà chúng ta có thể tìm thấy - và một sự nhìn nhận rằng mặc dù có thể là chúng ta không thể thay thế sự cân bằng sức hủy diệt bằng bất cứ điều gì tốt hơn, vẫn có nhiều việc chúng ta có thể thực hiện để làm cho sự cân bằng đó ổn định hơn thay vì bất ổn.<sup>3</sup>

---

<sup>3</sup> Trong trường hợp độc giả cảm thấy rằng cách lập luận trình bày ở đây là đúng về nguyên tắc nhưng thật sự không thú vị bởi vì tính không dễ bị tấn công liên tục của các lực lượng trả đũa của chúng ta được đảm

Một khi chúng ta đã xác định vấn đề tấn công bất ngờ là khả năng dễ bị tấn công của lực lượng trả đũa của mỗi bên trước đòn tấn công đầu tiên của bên kia, thì điều trở nên cần thiết là đánh giá sức mạnh quân sự, các biện pháp tự vệ, và các đề xuất về việc kiểm tra hay hạn chế vũ khí, với chính loại tổn thương chiến lược này trong đầu. Chẳng hạn như chúng ta không đánh giá các lực lượng chiến lược của Mỹ hay Liên Xô bằng cách đếm số máy bay ném bom, số tên lửa, số tàu ngầm và số hàng không mẫu hạm của hai bên, như kiểu chúng ta muốn xem bên nào có thể đưa ra lực lượng diễu binh ấn tượng nhất trong thời bình. “Bên nào vượt trội” trong cuộc chạy đua vũ trang thường thường sẽ là: *bất cứ bên nào tấn công trước*. Và nếu chúng ta phải hoạch định trên cơ sở giả thiết bảo thủ rằng bên kia sẽ tấn công trước, thì 200 máy bay ném bom được giữ an toàn khỏi cuộc tấn công có thể có giá trị như 2.000 máy bay mà chỉ có 105 có cơ hội an toàn.

Một cách thẩm định các biện pháp tự vệ cũng xuất hiện theo một cách khác nếu chúng ta chủ yếu dựa vào sự ngăn chặn. Chicago không thể được che giấu, hay chôn trong một hang động chống công phá, hoặc đặt sâu 10 dặm dưới lòng đất; nhưng sự che giấu, phân tán, hầm trú ẩn kiên cố, hay việc báo động từ trên không là những biện pháp phòng vệ có ý nghĩa trong việc duy trì lực lượng rắn đe. Một hệ thống phòng không hữu hiệu của Chicago mà chỉ có cơ may 50-50 trong việc cứu thành phố khỏi một quả bom nhiều triệu tấn

---

bảo không có gì phải lo lắng, tôi xin gợi ý độc giả tìm đọc phần thảo luận khá thuyết phục của Albert Wohlstetter trong “Cân bằng sức hủy diệt mong manh” (The Delicate Balance of Terror), tạp chí *Foreign Affairs*, 37: 211-234 (tháng 1-1959).

có thể sẽ là một viễn cảnh làm nản lòng, và chúng ta có ít hứa hẹn đảm bảo rằng chúng ta có thể làm tốt tới mức đó. Nhưng một hệ thống phòng không hữu hiệu mà có thể đảm bảo việc tồn tại của phần lớn lực lượng tấn công chiến lược có thể là quá đủ để đảm bảo với người Nga một cái giá đắt trong việc trả đũa. Tương tự như thế, một hệ thống phòng thủ Chicago mà đòi hỏi kẻ thù phải tăng gấp ba lần quy mô tấn công của họ có thể là một viễn cảnh tồi; có thể nó chỉ có ý nghĩa rằng kẻ thù sẽ đầu từ vào một cuộc tấn công ban đầu lớn hơn. Nhưng một hệ thống phòng thủ của lực lượng trả đũa của chúng ta mà đòi hỏi kẻ thù phải tăng gấp ba lần quy mô cuộc tấn công của họ có thể làm gia tăng một cách đáng kể sự khó khăn của kẻ thù trong việc vượt qua được hệ thống bảo động của chúng ta, và làm thay đổi đáng kể khả năng có thể của kẻ thù trong việc ngăn chặn thành công sự trả đũa.

Loại tính toán này cũng thích hợp đối với việc đánh giá các hạn chế vũ khí. Nếu chúng ta chỉ xem xét vấn đề của một cuộc tấn công của Nga vào các thành phố Mỹ, thì có thể dường như là không quan trọng mấy với đối phương cho dù họ bắn một tên lửa đạn đạo xuyên lục địa tầm gần hay tầm xa; sự chính xác cũng chẳng tạo ra khác biệt đáng kể nào với một quả bom hạt nhân triệu tấn nhắm vào các vùng đại đô thị. Nhưng nếu kẻ thù đang cố phá hủy một quả tên lửa hay một máy bay ném bom đã được che giấu sâu dưới lòng đất với những bức tường bê tông gia cố thì sự chính xác không còn được xem là vô dụng nữa. Một sai số mục tiêu với khoảng cách hai hay ba dặm có thể không là gì khi tấn công một đô thị lớn; một nỗ lực phá hủy một loại vũ khí trả đũa được che giấu kỹ lưỡng có thể đòi hỏi sử dụng nhiều tên lửa để có được một đòn đánh

đủ chính xác. Do vậy *các giới hạn khu vực* đối với việc đặt các tên lửa đạn đạo xuyên lục địa có thể có vẻ như là một hình thức giải trừ quân bị không hiệu quả theo nghĩa thông thường; nhưng trong việc làm ổn định sự ngăn chặn - trong việc làm giảm bớt tính dễ bị tấn công của lực lượng trả đũa của mỗi bên trước lực lượng của bên kia - thì việc tách rời các địa điểm đặt tên lửa của mỗi bên khỏi các địa điểm của bên kia, bằng cách làm giảm sự chính xác, có thể tạo ra sự khác biệt thật sự. (Đối với các máy bay hay tên lửa không được cất giấu, thì dĩ nhiên sự tương đồng như mục tiêu thành phố rủi thay lại là đúng).

Đối với một số câu hỏi, sự nhấn mạnh vào vấn đề tấn công bất ngờ có thể dẫn đến sự đảo ngược trực diện câu trả lời mà người ta có được từ các cách xem xét "giải giáp" truyền thống hơn. Hãy xem xét trường hợp về việc giới hạn số tên lửa mà hai bên được phép có (nếu chúng ta có thể đạt đến điểm mấu chốt trong các cuộc thương lượng với người Nga mà từ đó một hiệp ước hạn chế số tên lửa là thích hợp và việc giám sát có vẻ là khả thi). Giả sử chúng ta đã quyết định, từ việc xem xét các mục tiêu dân thường và các động cơ của kẻ thù, rằng chúng ta sẽ cần phải kỳ vọng tối thiểu có 100 tên lửa còn lại sau đòn tấn công chống tên lửa của kẻ thù để có thể thực hiện một đòn trả đũa trùng phạt tương xứng - nghĩa là, để ngăn chặn kẻ thù khỏi tấn công trước. Để minh họa hãy giả sử rằng sự chính xác và tính đáng tin cậy nơi hệ thống vũ khí của kẻ thù là ở mức mỗi tên lửa của họ có cơ may 50-50 trong việc phá hủy một tên lửa của chúng ta. Do vậy, nếu chúng ta có 200 tên lửa, kẻ thù cần phá hủy trên một nửa số đó; và với tính đáng tin cậy ở mức 50%, thì kẻ thù cần bắn hơn 200 tên

lửa để cắt giảm nguồn cung còn lại của chúng ta ở mức dưới 100 tên lửa. Nếu chúng ta có 400, kẻ thù cần phải phá hủy ba phần tư số tên lửa của chúng ta; và, với tỷ lệ khâu trừ 50% cho các lần bắn sai và bắn hỏng, kẻ thù sẽ cần phải bắn hơn hai lần con số 400, nghĩa là hơn 800 tên lửa. Nếu chúng ta có 800, kẻ thù cần phải phá hủy bảy phần tám số tên lửa của chúng ta, và với tính đáng tin ở mức 50%, kẻ thù sẽ cần phải bắn hơn ba lần con số đó, nghĩa là hơn 2.400 tên lửa. Và cứ tiếp tục như thế. Con số ban đầu của bên "phòng thủ" càng lớn, thì bên tấn công cần phải có *bội số* càng cao để giảm nguồn cung còn lại của bên nạn nhân xuống dưới con số "an toàn."<sup>4</sup>

Từ quan điểm này, một sự giới hạn trên số tên lửa sẽ có vẻ mang tính ổn định hơn, thì *con số tên lửa được phép càng lớn*. Sở dĩ như thế là vì hai lý do. Trước hết, con số tên lửa ở cả hai bên càng lớn, thì con số tuyệt đối kỳ vọng của số lượng tên lửa còn lại càng lớn để dùng trả đũa khi xảy ra sự cố một trong hai bên có thể tấn công trước, và do vậy khả năng ngăn chặn càng lớn trước một nỗ lực tấn công đầu tiên. Thứ hai, số tên lửa ở cả hai bên càng lớn, thì sự gia tăng con số tuyệt đối cũng như tỷ lệ phần trăm số tên lửa hẵn cũng phải lớn hơn mà bất kỳ bên nào cũng phải giành được để có thể bảo đảm được, với bất cứ khả năng xảy ra cụ thể nào, rằng số tên lửa còn lại của bên kia sẽ ít hơn một con số cụ thể nào đó sau khi bị tấn công. Như thế, khó khăn của việc một bên đánh lửa, bằng cách

---

<sup>4</sup> Điều này giả thiết rằng kẻ thù bắn các tên lửa của họ cùng một lúc hoặc rằng, nếu họ bắn các tên lửa theo từng đợt liên tiếp, họ không có phương tiện trinh sát để biết, trên từng đợt bắn liên tiếp, thì các tên lửa cụ thể nào đã hủy diệt các mục tiêu.

ngụy trang hay che giấu số tên lửa thặng dư, hay phá vỡ thỏa thuận và chạy đua để giành được một con số ưu thế, sẽ gia tăng nhiều hơn so với mức gia tăng các con số ban đầu ở cả hai bên. Thật vậy, nếu các con số khởi đầu đủ lớn để gây ra khó khăn cho những năng lực tài chính của hai kẻ thù, và trong phạm vi những năng lực tài chính này con số tên lửa là cao thì sự ổn định có thể được áp đặt bởi sự hạn chế về kinh tế lên việc mà một trong hai bên có thể thực hiện có liên quan đến việc mà bên đó phải thực hiện để giành được ưu thế.

Vậy thì, sau đây là một trường hợp, trong đó một cuộc "chạy đua vũ trang" không nhất thiết dẫn đến một tình huống càng lúc càng bất ổn. Với những con số cân bằng ở cả hai bên, thì khả năng xóa sổ thành công số tên lửa của bên kia càng trở nên khó xảy ra hơn khi mà số lượng tên lửa của cả hai bên gia tăng. Và *sức chịu đựng* của hệ thống cũng gia tăng theo. Nếu cả hai bên đều có những con số nhỏ, một tỷ lệ hai hay ba phá hủy được một có thể tạo ra ưu thế cho bên có nhiều hơn, với cơ hội tấn công trước và để lại cho bên kia một con số tuyệt đối khá nhỏ để đánh trả. Nhưng nếu con số khởi đầu của cả hai bên là lớn hơn, nó có thể đòi hỏi một tỷ lệ 10 trên một thay vì hai hay ba trên một để có một cơ hội tốt để tấn công mà không bị trừng phạt. Không bên nào phải hoảng sợ nếu có thất thế một chút, và không bên nào có bất cứ hy vọng lớn nào để có thể vượt xa lên trên nhằm tạo ra được ưu thế mà bên đó cần.

Cách nhìn đơn giản hóa ở mức cao này về "cuộc so găng tên lửa" là quá cụ thể để dùng làm lập luận cho các cuộc chạy đua vũ trang thay vì cho việc giải giáp. Nhưng nó cũng minh họa rằng, trong phạm vi lý luận của sự ngăn chặn ổn định,

và liên quan đến các kế hoạch ngăn ngừa tấn công bất ngờ, vấn đề nhiều vũ khí chống lại ít vũ khí phải được phân tích theo các giá trị của nó trong các trường hợp riêng biệt. Nó *không* phải là một kết luận tất yếu rằng việc giải giáp, theo nghĩa đơn thuần của nó, dẫn đến sự ổn định.

Thái độ của chúng ta đối với các tàu ngầm tên lửa, và đối với vấn đề thiết kế các kỹ thuật phát hiện tàu ngầm, có lẽ bị ảnh hưởng nhiều bởi việc liệu chúng ta lo lắng về sự tấn công của kẻ thù hay lo lắng về sự tấn công *bất ngờ* của kẻ thù. Nếu tàu ngầm trong nhiều năm đã chứng minh là khó bị tổn thương trước các tên lửa sát thương dân chúng, thì có lẽ chúng ta nên xem nó không phải như một công trình phát triển gây kinh hãi đặc biệt mà như một công trình trấn an. Nếu quả thực điều tốt nhất chúng ta có thể hy vọng là sự ngăn chặn lẫn nhau và chúng ta chỉ muốn sự cân bằng được ổn định, thế thì tên lửa loại phân cực mang trên một tàu ngầm có độ cơ động cao và sức chịu đựng lớn có thể là loại hệ thống vũ khí mà chúng ta muốn thấy với những con số cân bằng ở cả hai bên. Nếu nó có thể được chứng minh là vừa không thể bị phát hiện và vừa có độ tin cậy cao, thì hệ thống vũ khí đó sẽ có lợi thế là để tấn công thì không cần phải tấn công trước, không lo sợ rằng bên gây hấn có thể hy vọng phá hủy được chính các lực lượng được dùng để ngăn chặn họ. Thật vậy, có thể có vẻ như là bảo đảm hơn nếu chúng ta có năng lực phá hủy các tàu ngầm tên lửa của kẻ thù trong khi kẻ thù không có năng lực phá hủy tàu ngầm của chúng ta. Nhưng nếu năng lực đó đã tồn tại ở cả hai bên rồi và chúng ta không thể cầu mong nó đừng có, thì khi đó mức tối đa mà chúng ta có thể hy vọng là khả năng phá hủy lẫn nhau này tự nó rất bền vững để mỗi

bên thật sự bị răn đe. Từ quan điểm đó, chúng ta có lẽ thậm chí không nên mong ước rằng riêng chúng ta có thể có tàu ngầm hạt nhân “không dễ bị tấn công”; nếu thực tế chúng ta không có ý định mà cũng không có khả năng chính trị để tấn công trước, thì thông thường sẽ là có lợi nếu kẻ thù được bảo đảm một cách tự tin về điều này. Khả năng không dễ bị tấn công rõ ràng của kẻ thù trước đòn tấn công đầu tiên của chúng ta có thể là lợi thế cho chúng ta nếu nó giải tỏa cho kẻ thù khỏi mối quan ngại chính yếu mà có thể kích thích kẻ thù cố gắng tấn công trước. Nếu kẻ thù phải lo lắng về việc để lộ lực lượng chiến lược *của họ* trước cuộc tấn công của *chúng ta*, thì *chúng ta* cũng phải lo lắng về điều đó.

Những suy nghĩ này cũng ảnh hưởng đến thái độ của chúng ta đối với việc tìm kiếm kỹ thuật phát hiện tàu ngầm. Hải quân hiện đang tìm kiếm khẩn cấp một hệ thống phòng thủ chống lại tàu ngầm tốt hơn, và không có vấn đề gì ngoại trừ việc chúng ta phải tăng cường tập trung cao độ vào vấn đề này. Tuy nhiên có lẽ chúng ta phải đồng thời *hy vọng* rằng vấn đề là không thể giải quyết được. Nếu nó là không thể giải quyết (theo nghĩa tương đối trong đó một vấn đề kỹ thuật có thể là không thể giải quyết được) và những chiếc tàu ngầm được dự liệu là những phương tiện tương đối an toàn trong khoảng một thập kỷ, thì sự ngăn chặn ổn định có thể là khả thi về mặt công nghệ. Nếu các tàu ngầm có thể bị tổn thương, thế thì công nghệ vũ trang ít ổn định hơn so với mức chúng ta hy vọng. Mỹ phải *cố gắng* phát hiện tàu ngầm, vì Mỹ không thể để cho người Nga tìm ra một kỹ thuật mà Mỹ không biết, và bởi vì Mỹ phải biết tất cả những gì có thể biết về kỹ thuật phát hiện để làm cho tàu ngầm của Mỹ khó bị phát hiện hơn.

Nhưng giống như một người thiết lập được một thỏa thuận với một đối tác mà anh ta không tin tưởng, chúng ta phải tìm kiếm như điên cuồng một kẽ hở, vì biết rằng đối tác của chúng ta cũng đang tích cực tìm kiếm y như thế, đồng thời hy vọng rằng chẳng có kẽ hở nào được tìm thấy.<sup>5</sup>

Bởi vì chúng ta đã đẩy cuộc tranh luận đến mức này, nên sẽ phải tiếp tục cho đến khi hoàn tất. Nếu vấn đề của chúng ta là đảm bảo với một kẻ thù rằng chúng ta có khả năng đánh một đòn trả thù sau khi chính chúng ta bị tấn công – và trấn an kẻ thù rằng chúng ta biết rằng họ biết điều đó do vậy chẳng có gì xui khiến chúng ta nghi ngờ tiềm năng ngăn chặn của mình và tấn công trước – chúng ta nên nhìn nhận lợi ích của những phát minh công nghệ mà có thể nâng cao hiệu lực chống dân thường của các loại vũ khí trả đũa của mình. Nếu là hợp lý để sử dụng các biện pháp để đảm bảo rằng một phần lớn hơn các lực lượng trả đũa của chúng ta có thể vượt qua được một đòn tấn công đầu tiên nhắm vào họ, thì lý luận tương tự có thể làm cho chúng ta vui mừng về một sự gia tăng hiệu quả của các lực lượng trả đũa đã tồn tại được sau đòn tấn công. Như Bernard Brodie đã nói, "Khi chúng ta xem xét các yêu cầu đặc biệt về sự ngăn chặn, với sự nhấn mạnh vào khía cạnh trả đũa, chúng ta có thể tìm thấy một nhu cầu đối với thậm chí các loại bom siêu phóng xạ. Bởi vì phải tập trung làm sao cho chắc chắn rằng kẻ thù sẽ lo sợ ngay cả số lượng bom nhỏ nhất có thể được sử dụng để trả đũa,

<sup>5</sup> Tài liệu này nói về các nguyên tắc, chứ không phải về tàu ngầm, có lẽ tôi xin được thứ lỗi về việc nêu giả thiết ở đây rằng việc không thể bị phát hiện cấp thời ngoài khơi là tương đương với tính không thể bị tấn công được.

người ta muốn những quả bom này phải có, để xuất hiện trước sự kiện đó, chúng càng kinh khủng càng tốt.”<sup>6</sup>

Yếu tố mới mẻ của cách lập luận này biến mất ngay khi chúng ta nhìn nhận rằng sự “cân bằng sức hủy diệt”, nếu nó ổn định, chỉ đơn thuần là một dạng to lớn và hiện đại của một mô hình cổ xưa: việc trao đổi con tin. Trong những thời cổ xưa, một người cam kết giữ lời hứa bằng cách chuyển giao con tin bằng xương bằng thịt của mình vào tay của một “đối tác” không đáng tin cậy; công nghệ quân sự ngày nay khiến người ta có thể nắm giữ số phận của phụ nữ và trẻ em của một kẻ thù tiềm năng trong phạm vi tầm tay cho dù kẻ thù cắt giữ những phụ nữ và trẻ em đó cách xa hàng ngàn dặm. Chứng nào mà mỗi bên có một sức mạnh hiển nhiên có thể hủy diệt một đất nước và dân số của nó để đáp trả lại một cuộc tấn công của bên kia, thì sự “cân bằng sức hủy diệt” phát triển thành một sự hiểu ngầm được hỗ trợ bởi một sự trao đổi toàn diện tất cả các con tin mà chúng ta có thể nghĩ ra được. Dĩ nhiên, có thể chúng ta không mong muốn trao đổi nhiều con tin đến *nhu thế* chỉ để hỗ trợ cho một sự cảm thông cụ thể đối với một kẻ thù cụ thể. Nhưng trong một thế giới vô luật lệ không có cơ sở cho những vụ kiện về tổn thất do việc vi phạm hợp đồng bất thành văn này, các con tin có thể là công cụ duy nhất mà các đối tác bất tín và đối kháng có thể sử dụng như là phương tiện mặc cả.<sup>7</sup>

<sup>6</sup> Bernard Brodie, *Strategy in the Missile Age* (Chiến lược trong Thời đại Tên lửa) (Princeton, 1959) trang 295.

<sup>7</sup> Cần phải nhấn mạnh rằng ở đây tôi đang chỉ thảo luận vấn đề tấn công bất ngờ lớn. Các ngụ ý về khái niệm “con tin” đối với chính sách phòng thủ dân sự chẳng hạn phụ thuộc vào mối quan hệ của nó với các sự kiện xảy ra khác nữa – ví dụ như chiến tranh hạn chế, ác ý của một

Lối lập luận hép này thì không chỉ là một sự hợp lý hóa lớn lao cho một cuộc chạy đua vũ trang. Nó còn thật sự gợi ý rằng "sự giải trừ quân bị" theo nghĩa đen của nó, nhắm đến tất cả các loại vũ khí không có sự phân biệt - hoặc thậm chí nhắm đến một cách có chọn lọc các loại vũ khí hủy diệt hàng loạt khủng khiếp nhất - có thể tạo ra sự bất ổn định hơn là sự ổn định, và có thể cần phải *hoàn toàn* thành công để không trở nên thảm khốc. Tuy nhiên, có một lĩnh vực quan trọng của việc hạn chế vũ trang không chỉ thích hợp với việc phân tích đã đề cập trước mà còn được mở từ sự phân tích đó.

Nó gợi mở việc phân biệt rõ ràng giữa các loại vũ khí đặc biệt thích hợp với việc khai thác đòn tấn công đầu tiên và các loại vũ khí đặc biệt thích hợp cho vai trò trả đũa. Ở một thái cực là loại vũ khí đơn thuần để đánh trả: loại phương tiện tương đối kém chính xác với một trái bom siêu phong xạ mà có thể giết chết hầu hết mọi thú trong đất nước kẻ thù *ngoại trừ* một lực lượng trả đũa được bảo vệ kỹ hay che giấu kỹ, mà chính nó cũng được bảo vệ kỹ hay che giấu kỹ để không bị tổn hại trước bất kỳ loại vũ khí nào mà bên kia có thể sở hữu. Một cách lý tưởng, loại vũ khí này không chịu bất lợi nào trong khi chờ đợi để tấn công sau nhưng cũng không có lợi thế nào khi dùng để tấn công trước. Ở thái cực đối ngược là một loại vũ khí mà tự nó dễ bị tổn hại đến mức không thể tồn tại được để tấn công sau, hoặc một loại vũ khí để tìm và diệt các lực lượng trả đũa của kẻ thù trước khi chúng được bắn di quá đặc biệt

---

bên thứ ba, sự trả đũa chưa đến mức quy mô lớn v.v... Một trong những mối liên hệ qua lại này giữa tấn công bất ngờ và các sự kiện quân sự khác được đề cập đến trong các trang cuối cùng của chương này.

đến mức nó sẽ mất hầu hết tác dụng nếu nó được giữ lại cho đến khi phía bên kia đã khởi sự tấn công. Các loại vũ khí “tấn công trước” này không chỉ tạo cho người sở hữu chúng động cơ mạnh mẽ để tấn công trước, và động cơ vội vã sử dụng vũ khí khi có sự kiện báo động mơ hồ thay vì chờ đợi và hoàn toàn chắc chắn; chúng là một lời công bố ngầm với kẻ thù rằng một bên sẵn sàng tấn công trước. Hậu quả là chúng mời gọi kẻ thù tấn công trước đó một chút và hành động vội vàng trong trường hợp kẻ thù nghĩ rằng chúng ta nghĩ rằng đã đến lúc hành động nhanh chóng.

Giữa những thái cực của loại vũ khí “thuần” để tấn công trước và loại vũ khí “thuần” để đánh trả, có những loại vũ khí có thể dùng để tấn công trước nhưng không cần thiết phải làm thế, vũ khí mà có thể tồn tại và phục vụ mục đích đánh trả nhưng nó cũng có thể có một ảnh hưởng quan trọng lên các lực lượng trả đũa của bên kia nếu nó được sử dụng tấn công trước. Có lẽ hầu hết các loại vũ khí rơi vào chủng loại này nếu các biện pháp phòng ngừa được áp dụng để bảo vệ chúng. Do vậy chúng ta không thể phân biệt rõ ràng giữa các loại vũ khí tấn công trước và vũ khí tấn công sau, ca ngợi loại vũ khí này và chê bai loại vũ khí kia trong cách tiếp cận vấn đề tấn công bất ngờ. Nếu chúng ta cần phải xem xét việc hủy bỏ tất cả các loại vũ khí mà có thể có ảnh hưởng đối với các lực lượng trả đũa của bên kia, hoặc loại có lợi thế để sử dụng tấn công trước, thì có thể số còn lại không đủ để dùng cho mục đích trả đũa.<sup>8</sup> Nhưng các cuộc thương lượng về tấn công bất ngờ có thể tập trung một cách hữu ích vào thái cực đối nghịch kia.

<sup>8</sup> Hơn nữa, ở đây chúng ta không xem xét vấn đề gì khác ngoại trừ vấn đề tấn công bất ngờ.

Các ứng viên hiển nhiên nhất sẽ là các loại vũ khí được phơi bày lộ liễu và dễ bị tổn hại. Có vẻ như là bất thường khi đề nghị với người Nga là họ che giấu đi sự trân trọng của các lực lượng chiến lược bên phía họ, hay ngược lại phía họ gợi ý chúng ta cần phải bảo vệ tốt hơn các lực lượng của chúng ta. Điều thường thấy hơn là các gợi ý từ bỏ các loại vũ khí mang tính chất khiêu khích lộ liễu đối với bên kia. Hãy lưu ý sự khác biệt về mặt tinh thần giữa điều này với định hướng "cấm bom." Cho dù các ngụ ý tuyên truyền của đế tài này là gì đi nữa, ít nhất nó cũng có giá trị xem xét sự ngăn chặn như là một việc cần phải nâng cao, chứ không phải là một việc cần hủy bỏ.

Thứ hai, các biện pháp hạn chế việc triển khai lực lượng mà làm ảnh hưởng đến tiềm năng chống trả của họ thay vì tiềm năng tấn công dân thường có thể được tìm kiếm. Tuy nhiên, các biện pháp này sẽ không được tìm kiếm cho đến khi có nhận thức rõ ràng rằng các kế hoạch tấn công bất ngờ là có chủ đích nhắm đến việc bảo vệ, chứ không phải làm suy yếu, năng lực đánh trả của mỗi bên. Phần thảo luận trên về hiệu quả chuỗi đối với các yêu cầu về tên lửa, cho dù các giá trị cụ thể của nó là gì, cũng gợi ý rằng loại biện pháp hạn chế này không phải là vô ích.

Thứ ba, có thể có một cách thăm dò hữu ích nào đó về các biện pháp hợp tác, hay các mô hình hành vi điều chỉnh hỗ tương, mà giúp giảm bớt nguy cơ chiến tranh do hiểu lầm. Ngay cả việc trao đổi thông tin tự nguyện cũng có thể giúp ích, nếu Mỹ và người Nga có thể đơn phương lựa chọn những ~~mô hình hành vi~~ mà, khi sự thật được rõ ràng, mang tính chất trấn an. Đây có lẽ là ý tưởng đằng sau các đề xuất thanh tra hoạt động không lưu ở khu vực Bắc cực, và có thể có một số

loại hoạt động khác trong đó có thể có lợi ích chung từ một số quy tắc lưu thông nào đó. Điều hấp dẫn liên quan đến các biện pháp này - cũng như liên quan đến việc thảo luận thẳng thắn về các hậu quả tệ hại của các hệ thống vũ khí tấn công trước - là ở chỗ chúng có thể tạo ra một số cách hiểu biết mà không nhất thiết phải được thể hiện trong các thỏa thuận chính thức, và có thể hỗ trợ cho các hoạt động điều chỉnh đơn phương ở cả hai bên.

Thứ tư, có thể có những hoạt động sắp xếp để đối phó với các cuộc khủng hoảng và các tình huống khẩn cấp có thể đe dọa làm nổ ra một cuộc chiến không có chủ định. Phần sau của chương này sẽ thảo luận điểm này ở một mức độ nào đó.

Thứ năm, có thể có những biện pháp mà, bằng việc làm cho yếu tố bất ngờ trở nên ít có cơ may xảy ra, khiến cho hoạt động tấn công trước trở nên ít hấp dẫn. Điểm này đưa chúng ta trở lại loại đề xuất bầu trời mở.

Hầu hết việc thảo luận công khai về vấn đề tấn công bất ngờ trong một vài năm qua đều có liên quan đến các biện pháp mà có thể làm giảm bớt khả năng xảy ra yếu tố bất ngờ, hơn là các biện pháp để hạn chế tầm mức mà vũ khí có thể gây ra nếu yếu tố bất ngờ đạt được. Đề xuất bầu trời mở dựa trên ý tưởng rằng với sự quan sát đầy đủ về lực lượng quân sự của mỗi bên thì không bên nào có thể giành được yếu tố bất ngờ và, do thiếu lợi thế của sự bất ngờ, sẽ bị ngăn chặn.

Vấn đề kỹ thuật của việc thiết kế một kế hoạch thanh tra thực tế mà có thể đưa ra cho mỗi bên sự báo động tương ứng về một cuộc tấn công của bên kia đã trở nên khó khăn hơn

nhiều kể từ khi đề xuất bầu trời mở được đưa ra lần đầu tiên. Với việc các loại vũ khí khinh khí có thể giảm bớt số lượng máy bay mà có thể cần đến trong một cuộc tấn công bất ngờ, với việc các loại tên lửa hứa hẹn giảm bớt tổng lượng thời gian giữa các hành động ban đầu trong việc chuẩn bị sẵn sàng một đòn tấn công và việc kích nổ các loại vũ khí đúng vào mục tiêu, và với các hệ thống cơ động như các tàu ngầm tên lửa trong tầm kiểm tra, thì có vẻ như là việc thanh tra đơn thuần không đi kèm với bất cứ giới hạn nào về hành vi của những sự việc được thanh tra sẽ là cực kỳ khó khăn hoặc cực kỳ không hiệu quả. Ý tưởng kiểm tra các hình ảnh chụp để có những chỉ dẫn chiến lược về các hoạt động di chuyển và tập trung lực lượng đơn giản là quá lỗi thời. Vấn đề hiện nay có vẻ như là một vấn đề về giám sát chuyên sâu các lực lượng chiến lược bởi một tổ chức rộng lớn mà có thể truyền đi các thông điệp xác thực, báo cáo hoạt động đáng ngờ trong phạm vi tối đa một vài giờ, và cuối cùng trong vòng một vài phút, theo một cách thức không bị ảnh hưởng một cách quá sức chịu đựng bởi các báo động giả. Không có gì đảm bảo thực tế rằng điều này có thể thực hiện được.

Điều này không có nghĩa là các kế hoạch thanh tra chống lại tấn công bất ngờ không có triển vọng thành công. Nó có nghĩa là một kế hoạch không nhằm đến *điều gì khác ngoài việc thanh tra* có thể có những triển vọng rất hạn chế. Nhưng nếu người ta không thể gởi các quan sát viên đi theo dõi mọi máy bay, tên lửa và tàu ngầm tới mọi nơi mà chúng di chuyển, thì người ta vẫn có thể kêu gọi các máy bay, tên lửa và tàu ngầm tập hợp tại nơi mà chúng dễ được quan sát hơn. Nếu các biện pháp hạn chế đối với việc triển khai lực lượng được

sử dụng để làm cho công tác thanh tra dễ quản lý hơn, thì có thể đạt được một điều gì đó. Nhưng mặc dù có hứa hẹn trong ý tưởng kết hợp thanh tra và các biện pháp hạn chế vũ khí, thì cũng vẫn có những vấn đề nghiêm trọng.

Một trong những vấn đề đó là sự bất tương đồng có thể có giữa nhu cầu thanh tra và nhu cầu che giấu. Khi các tên lửa đạt tới độ chính xác đủ, thì hầu như không thể bảo vệ được các lực lượng trả đũa của một bên chỉ bằng cách sử dụng bê tông cốt thép để che chắn, hoặc nếu không thể như thế, thì sẽ phải chịu tổn thất rất cao. Tính cơ động và sự che giấu có thể là nguồn an toàn cho các lực lượng trả đũa; nếu kẻ thù có thể đánh trúng bất cứ thứ gì mà họ phát hiện, và có thể phá hủy bất cứ thứ gì mà họ đánh trúng, thì phải làm cho kẻ thù không thể phát hiện ra nó. Trong phạm vi mà kẻ thù có thể theo dõi, giám sát liên tục các loại vũ khí trả đũa của chúng ta, thì kẻ thù sẽ có thông tin liên tục về vị trí của chúng.

Trong những cách thức khác, một kế hoạch thanh tra ở quy mô cần có để đảm bảo chống lại tấn công bất ngờ có thể đưa ra thông tin thái quá về việc dàn binh của các lực lượng phía bên kia và khiến cho các lực lượng đó dễ tổn thương. Chẳng hạn như nhiều người biết việc đã từng có một thời điểm xảy ra các cơn lốc lớn làm tê liệt một phần rất lớn các máy bay ném bom B-36 vốn chiếm giữ một vai trò chủ chốt trong lực lượng trả đũa của Mỹ. Những ngụ ý về tấn công bất ngờ của một sự kiện như vậy rõ ràng là rất khác biệt, tùy thuộc vào việc liệu kẻ thù chỉ biết chung chung rằng loại sự việc này có thể xảy ra cho chúng ta, hoặc thay vào đó kẻ thù có thông tin xác định khi nào sự việc xảy đến và biết chính xác liệu nó có

thể hành động không bị trở ngại trong một vài ngày hay không. Hãy tưởng tượng tình trạng căng thẳng mà có thể xảy ra nếu nhân lực chiến lược của *một trong hai* bên bắt đầu bị ảnh hưởng bởi một đại dịch nghiêm trọng đe dọa đến khả năng huy động tạm thời của họ trước con mắt quan sát của phía bên kia. Sẽ là tốt hơn nhiều - nếu chúng ta và kẻ thù thỉnh thoảng lâm vào một tình huống rất mất cảnh giác vì những lý do không thể ngăn ngừa được - khi một trong hai bên không ở trong tình trạng biết quá rõ về những tình huống mất năng lực đôi khi xảy ra của bên kia.

Cuối cùng, trong khi có thể có những việc sắp đặt có khả năng cao trong việc đưa ra báo động về việc chuẩn bị tấn công của kẻ thù, thì giá trị của hệ thống tùy thuộc vào việc chúng ta có thể làm gì nếu chúng ta nhận được báo động. Chúng ta có thể đưa ra đòn tấn công phủ đầu của riêng mình, với hy vọng tấn công trước; nhưng đây là một xu hướng không hấp dẫn nếu báo động là mơ hồ. Một báo động sai như thế dẫn đến chiến tranh. Và một báo động đúng lại là khúc mở đầu cho việc ngăn chặn vào giờ chót.-

Ở thái cực khác chúng ta có thể chỉ chờ đợi và "chuẩn bị sẵn sàng." Và nếu những việc mà chúng ta có thể làm để chuẩn bị sẵn sàng có thể làm giảm đáng kể khả năng kẻ thù tấn công thành công - nếu họ đưa ra khả năng cho rằng chúng ta sẽ trả đũa một cách mạnh mẽ - chúng ta có thể mong muốn mau chóng biểu lộ cho kẻ thù thấy rằng chúng ta sẵn sàng, với hy vọng rằng sự biểu lộ tăng cường của chúng ta sẽ ngăn chặn quyết định cuối cùng của họ.

Câu hỏi quan trọng là chúng ta làm gì để tạo ra tình trạng sẵn sàng. Nếu câu trả lời chỉ đơn giản là "Hãy cảnh giác hơn",

thế thì tại sao chúng ta không cảnh giác hơn ngay từ đầu? Hầu hết những việc hiển nhiên mà một bên sẽ thực hiện nếu bên đó nhận được báo động về một cuộc tấn công chính là những việc mà một bên có lẽ mong muốn thực hiện liên tục qua việc nhìn nhận khả năng thường trực về một cuộc tấn công. Và nếu Bộ Tư lệnh Không quân Chiến lược của chúng ta liên tục nỗ lực hết sức để rút ngắn thời gian cần thiết trong việc chuẩn bị và cho xuất phát các máy bay khi có báo động, hoặc trong việc bảo vệ an toàn các máy bay được che giấu, hoặc giữ các máy bay an toàn trên không trung trong tình trạng sẵn sàng chiến đấu, thì có thể là kẻ thù chẳng thể làm được gì nhiều hơn trong tình huống cấp bão.

Tuy nhiên, có những điều mà một quốc gia có thể làm khi đối mặt với một cuộc tấn công sắp xảy ra nhưng không thể làm liên tục và vô hạn định. Người ta có thể di tản hay xuống hầm, nhưng không thể mãi mãi như thế. Người ta có thể cho các lực lượng trả đũa của mình xuất phát khỏi mặt đất để không còn là các mục tiêu ném bom của kẻ thù; nhưng chúng không thể ở mãi trên không trung. Người ta có thể huy động binh lính của mình trực chiến liên tục hai mươi bốn giờ, nhưng không thể trong nhiều ngày liên tục. Người ta có thể cho hạ cánh tất cả các máy bay thương mại để nâng cao tính tin cậy của hệ thống báo động, nhưng thiệt hại về kinh tế có thể sẽ khủng khiếp nếu phải cấm tất cả các chuyến bay thương mại và tư nhân lâu dài chỉ để có lợi trong việc dễ phát hiện ra máy bay của kẻ thù. Nói cách khác, có những việc mà người ta có thể thực hiện để “chuẩn bị sẵn sàng” khi đối mặt với một cuộc tấn công có thể xảy ra mà người ta không thể kỳ vọng thực hiện được liên tục.

Nhưng có một câu hỏi khác. Chúng ta có thể duy trì tình trạng đó được bao lâu? Giả sử rằng chúng ta không thể lúc nào cũng duy trì tất cả máy bay trên không trung, mà sự thật là thế, và rằng có thể là quá tốn kém về mọi khía cạnh (tai nạn cũng như chi phí nhiên liệu và phi hành đoàn) để duy trì trung bình khoảng một nửa số máy bay trên bầu trời, và rằng sẽ có một sự gia tăng đáng kể số máy bay trên không trung trong tình huống báo động gấp nếu nhận được một lời cảnh báo nghiêm trọng. Điều này có thể có nghĩa rằng kẻ thù sẽ không bị ngăn chặn bởi thể hiện thông thường của chúng ta, nhưng sẽ bị ngăn chặn bởi thể hiện mà chúng ta có thể áp dụng khi chúng ta nhận được sự báo động. Liệu điều đó có nghĩa là kẻ thù sẽ từ bỏ ý định ngay khi nhận thấy là chúng ta sẵn sàng hay không? Hay kẻ thù có thể sẽ chờ đợi cho đến khi nhiên liệu đã cạn, các phi công đã mệt mỏi, và các máy bay phải hạ cánh xuống? Và nếu như thế, chúng ta có phải tấn công trước hay không?

Vấn đề “mệt mỏi” này có xu hướng làm tê liệt bất cứ cách thức siêu báo động nào mà người ta có thể áp dụng. Giải pháp gồm hai phần. Trước tiên, người ta phải cố gắng thiết kế một hệ thống đáp trả siêu báo động có sức chịu đựng cao và bớt mệt mỏi, nhìn nhận rằng điều này có nghĩa là phải chấp nhận giảm bớt tính hiệu quả cao của nó. Thứ hai, và thích hợp nhất với chủ đề lúc này, là người ta có thể buộc phải tham gia vào một loại thương lượng về giải giáp mạnh với kẻ thù trong thời điểm mà thực tế người ta áp dụng các biện pháp để bảo vệ cho tính dễ bị tổn thương của mình trước việc trả đũa. Nếu Mỹ có thể duy trì tình trạng siêu báo động trong một vài ngày, thì Mỹ sẽ có một vài ngày để cố gắng đàm phán hay thương lượng

một mức độ nào đó về việc “giải trừ quân bị” của phía Nga mà phù hợp với họ đồng thời đủ để trấn an cho Mỹ để cho phép Mỹ trở lại tình trạng “bình thường” thay vì xúc tiến một cuộc chiến tranh tổng lực. Điều này có thể có nghĩa là phải thiết kế và tạo lập một kế hoạch nhiều tham vọng hơn với các biện pháp chống tấn công bất ngờ so với các biện pháp khả thi về chính trị của giai đoạn trước đó. Điều đó có nghĩa là thương lượng không chỉ dưới áp lực thông thường về việc biết rằng tấn công lén là một nguy cơ dài hạn, mà là thương lượng với sự lưu ý rõ ràng rằng nếu các biện pháp để làm cho việc tấn công bất ngờ thành công trở thành không thể đã không được thiết kế, được thống nhất với nhau, và áp dụng theo một thời hạn chót mau chóng, thì chiến tranh mà cả hai bên đều đồng thuận đã trở thành điều không thể tránh khỏi.

Những phản ánh này không ngu ý rằng việc tăng cường báo động sẽ là hoặc vô dụng hoặc làm rối rắm thêm. Điều mà chúng thể hiện là việc báo động tự nó có thể là không đủ. Việc tăng cường báo động cung cấp một *cơ hội*, nhưng cơ hội phải được khai thác với kỹ năng. Và các công việc chuẩn bị cho điều mà người ta sẽ thực hiện để đối phó có thể phải được thực hiện từ rất sớm trước đó. Khó mà có thời gian để đưa ra một tối hậu thư cho phía Nga vào lúc Mỹ bắt gặp Nga đang chuẩn bị tấn công. Quyết định loại tối hậu thư nào mà vừa đáp ứng các nhu cầu của Mỹ và vừa chấp nhận được đối với người Nga thì không chỉ khó khăn về mặt trí tuệ mà còn khó khăn về mặt kỹ thuật, tùy thuộc vào việc xem xét quy trình để xác nhận sự tuân thủ là gì. Có lẽ Mỹ có thể đưa ra một tối hậu thư có hiệu quả nếu và chỉ nếu Mỹ đã hoạch định cẩn thận từ trước tối hậu thư đó có thể chứa đựng những nội dung gì.

Có hai tiêu chí rất rạch ròi để đánh giá tính hiệu quả của một hệ thống thanh tra kiểm soát, hay để thiết kế chính hệ thống đó. Một tiêu chí là hệ thống có khả năng tiếp cận sự thật ở mức độ như thế nào bắt chấp các nỗ lực che giấu sự thật đó; tiêu chí kia là hệ thống đó có khả năng thế nào trong việc giúp người ta bộc lộ sự thật một cách thuyết phục khi họ có lợi ích để làm như thế. Sự khác biệt thì cũng giống như sự khác biệt giữa một kế hoạch để phát hiện người có tội và một kế hoạch cho phép người vô tội khẳng định được sự vô tội của mình. Nói một cách nôm na, một kế hoạch tiếp cận giả thiết về sự vô tội theo một cách thức tiêu cực, bởi thiếu chứng cứ tích cực ngược lại; kế hoạch kia dựa vào chứng cứ tích cực, và tùy thuộc vào các tình huống cụ thể trong đó lợi ích của người ta nằm ở chỗ để cho sự thật được bộc lộ.

Sự khác biệt giữa hai tình huống này tùy thuộc vào sự phân biệt rạch ròi giữa một kế hoạch giảm thiểu nỗi lo sợ về tấn công bất ngờ *có chủ đích* và một kế hoạch giảm thiểu nỗi lo sợ về cuộc chiến tranh do khinh suất, do “sự cố” hay không có chủ đích - một cuộc chiến tranh xuất phát từ một báo động giả, hay từ một đánh giá sai lầm về sự đáp trả của đối phương trước một báo động giả, hay trước một sự giải thích sai về một sự cố không cố ý, hay trước một ý đồ xấu xa có tính chất kích thích của một bên thứ ba quan tâm đến việc thúc đẩy chiến tranh, hay trước một tình huống trong đó việc hiểu biết của mỗi bên cho rằng bên kia có thể sắp sửa tấn công phủ đầu bùng nổ ra bởi việc thúc đẩy một cuộc chiến tranh xuất phát từ nỗi sợ hãi lẫn nhau. Trong trường hợp một cuộc tấn công bất ngờ có hoạch định và có chủ đích, bên gây hấn có mọi lý do để che giấu sự thật. Nhưng trong trường hợp “chiến tranh

khinh suất", cả hai bên đều có lợi ích mạnh mẽ trong việc chuyển tải sự thật nếu sự thật có thể thực sự được chuyển tải trong một cách thức đáng tin cậy kịp thời để ngăn chặn quyết định sai lầm của bên kia.

## *HIẾU SAI VỀ TẤN CÔNG*

Hãy xem xét câu hỏi này: làm cách nào Mỹ có thể chứng minh với Liên Xô là Mỹ không chuẩn bị cho một cuộc tấn công bất ngờ, khi mà trong thực tế Mỹ không có nhưng Nga lại tưởng rằng Mỹ có thể có? Làm cách nào Nga có thể chứng minh với Mỹ rằng họ không khởi xướng một cuộc tấn công bất ngờ, nếu trong thực tế họ không có nhưng họ biết rằng Mỹ lo sợ họ có thể làm điều đó.

Rõ ràng việc chỉ nói sự thật thì sẽ không đủ. Quả thực có thể có một số tình huống trong đó sự tiếp xúc thuần bằng lời nói cũng đủ làm giảm mối nghi ngờ của mỗi bên. Nếu người Nga - chỉ đơn cử một ví dụ đại khái - tình cờ chịu tổn thất vì một vụ nổ hạt nhân tại một trong những căn cứ của chính họ, thì có thể đó cũng là một điều hữu ích cho cả hai bên nếu họ có thể đơn giản trấn an Mỹ một cách nhanh chóng rằng họ biết đó là một tai nạn, và rằng họ không xem đó là một báo hiệu về một cuộc tấn công từ phía Mỹ, và vân vân. Nhưng, trong hầu hết các trường hợp mà người ta có thể tưởng tượng ra, sẽ là không đủ nếu như chỉ khẳng định rằng một bên đang không tiến hành một cuộc tấn công chiến lược hoặc một bên không ở trong tư thế đe dọa. Phải có cách nào đó để xác minh

một số sự kiện, những sự kiện được cho là có liên quan tới việc bày binh bố trận. Mỹ sẽ không chỉ phải chứng minh rằng Mỹ không *có ý định* lợi dụng vị trí của mình, mà còn phải chứng minh rằng vị trí thực sự của Mỹ là một vị trí *không thể* được lợi dụng để lừa dối kẻ thù nếu như họ tin lời Mỹ và kiềm chế sức mạnh quân sự của chính họ.

## *SỰ HIẾU SAI TRONG CHIẾN TRANH HẠN CHẾ*

Đặc biệt là trong quá trình của một cuộc chiến tranh hạn chế bên này hay bên kia đều có thể thực hiện một hành động mà có thể bị hiểu sai là một cuộc đột kích chiến lược. Chẳng hạn, giả như Mỹ đã sử dụng những loại máy bay mà sẽ được chọn sử dụng trong một cuộc tấn công chống lại các căn cứ của Nga, và bay theo hướng có thể bị hiểu là nhắm vào chính Liên Xô - có thể cũng như trong trường hợp nếu chúng bay từ các căn cứ Bắc Phi hoặc phi đội ở Địa Trung Hải tới các quốc gia giáp ranh giới phía nam của Liên Xô. Hoặc cách khác, giả sử như máy bay của Liên Xô đang xung kích trong một cuộc chiến tranh hạn chế có thể bị hiểu, dựa trên chứng cứ tạm thời Mỹ có thể có được, là tấn công vào tất cả các căn cứ và các tàu sân bay ở nước ngoài của Mỹ, nhưng thực sự nó là một cuộc tấn công giới hạn chứ không phải là một phần của một nỗ lực tổng thể nhằm tiêu diệt khả năng trả đũa của Mỹ.

Vấn đề được đặt ra là liệu có cách nào để giảm khả năng hiểu sai trong trường hợp này không, trường hợp mà sự hiểu sai có thể dẫn tới việc một bên hoặc sẽ bắt đầu trả đũa trước,

tấn công phủ đầu càng nhanh càng tốt, hoặc áp dụng tình trạng siêu cảnh giác mà có xu hướng cao có thể dẫn tới báo động giả. Người ta có thể mong muốn đem hết sức mình ra để chứng minh rằng các hành động bổ sung - các hành động liên quan tới những lực lượng khác tại các vùng khác trên thế giới, những hành động hầu như chắc chắn xảy ra nếu đây là một cuộc phản công tổng lực - không xảy ra trong thực tế.

### SỰ HIẾU SAI LẦN NHAU

Hãy xét tới một trường hợp khác được Gromyko mô tả tại một cuộc họp báo.

Rốt cuộc, những vì sao băng và sự nhiễu điện tử cũng được phản chiếu trên các màn hình radar của Liên Xô. Nếu trong những trường hợp như thế các máy bay Liên Xô, có chở bom nguyên tử và bom khinh khí, đang hướng về nước Mỹ và các căn cứ của Mỹ trong các quốc gia khác, các phi đội của cả hai bên, đã nhận ra nhau ở một nơi đâu đó trong vùng Bắc cực, trong những tình huống như thế sẽ đương nhiên đi đến kết luận rằng một cuộc tấn công đích thực của địch quân đang xảy ra, và nhân loại sẽ thấy mình bị vướng vào vòng xoáy của chiến tranh nguyên tử.

Tạm thời giả sử rằng có thể tin được là một tình huống tương tự như tình huống được mô tả có thể phát sinh, vậy bằng cách nào người ta có thể làm chậm và đảo ngược sự hiểu sai lầm nhau của cả đôi bên? Nếu có một cách nào đó

để đảo lộn chiêu hướng của cả hai bên, theo một cách thức phân kỳ hợp lý và xác thực, thì đôi bên có thể đi đến đồng thuận về một loại rút quân cân bằng.

Môi trường mặc cả không phải là một môi trường thuận lợi. Trong trường hợp tốt nhất thì thời gian tiến hành các cuộc thương lượng cũng chỉ được tính bằng giờ, còn trong trường hợp tệ hại thì chẳng có một chút thời gian nào cả. Những yêu cầu để đạt kết quả mỹ mãn theo lô-gích phân tích thì có thể được chia làm hai phần. Thứ nhất phải tìm cho ra một "giải pháp" nào đó - một mô hình hành động làm đảo ngược xu hướng tấn công lẫn nhau, và tạo ra sự rút lui ổn định một cách năng động về một tình trạng báo động ít mang tính đe dọa hơn, một tình trạng mà không tạo ra cho bên nào một lợi thế đầy nguy hiểm trong quá trình đó, và nằm trong các khả năng vật chất của các lực lượng có liên quan. Yêu cầu thứ hai chính là sự tuân thủ có thể quan sát được, biện minh được, và chứng minh được bằng một cách nào đó. Chúng ta không thể tiến hành phần mặc cả của chúng ta trừ phi chúng ta có được phương tiện đáng tin cậy để theo dõi sự tuân thủ của bên kia, và đối với họ thì cũng đúng như vậy. Có thể tin được rằng chúng ta có thể có lợi ích trong chuyện lừa dối; nhưng rất có thể xảy ra rằng trong những tình huống này chúng ta lại mong ước có được một hệ thống giám sát chống lừa dối mà chúng ta có thể tin cậy, để nếu chúng ta tuân thủ phần mặc cả của mình thì bên kia sẽ không chút hề nghi gì về điều đó. Vấn đề về bản chất là việc thực thi hợp đồng. Và động cơ trong trường hợp này, đối với mỗi bên, là phải hết sức chuyển tải sự thật liệu trong thực tế mình có tuân theo đúng kế hoạch không.

Ví dụ này không chỉ cho thấy rõ nhu cầu cần có *sự sắp đặt trước* nào đó cho việc quan sát và xác minh, trên cơ sở có rất ít thời gian để đưa các thanh sát viên đến hiện trường; nó còn chứng minh cho tầm quan trọng của việc suy tính trước về loại đe dọa nào phải đưa ra, và việc thiết kế kế hoạch rút lui của chính mình theo cách có thể tận dụng tối đa bất cứ phương tiện nào chúng ta có thể có để cung cấp một cách có chủ đích thông tin đúng cho kẻ thù trong trường hợp nhất thiết phải làm thế.

Trường hợp này cũng có thể minh họa cho sự khác biệt giữa hai tiêu chí về tính đáng tin cậy của một hệ thống kiểm tra. Có thể rất khó thiết kế hệ thống ra-đa mà *luôn luôn* nắm bắt được kẻ thù - và qua đó kẻ thù cũng có thể nắm bắt được chúng ta - đang cố lén lút tấn công. Một vấn đề đáng kể nữa là làm sao thiết kế ra-đa để nếu cả hai muốn khuyến khích việc giám sát tự nguyện, chúng ta đều có thể dựa vào đó một cách có sức thuyết phục. Trong thực tế, một mặt chúng ta đang tránh tối đa sự theo dõi bằng ra-đa của địch. Một mặt chúng ta cố ý "diễn hành" trước ra-đa của địch, hoặc dựa vào phương tiện nhận biết từ xa khác, miễn là địch cũng làm y như thế cho chúng ta.

## GIÁM SÁT DÀI HẠN

Sự khác biệt giữa những tình huống khủng hoảng và tình huống khẩn cấp này và những vấn đề dài hạn hơn của việc đưa ra các chính sách về hạn chế vũ khí nằm ở loại chứng cứ

cần phải có và ở sức mạnh của động cơ cung cấp chứng cứ đó. Quá trình kiểm tra “thư thả” hơn thường được xem là chủ yếu dựa vào *chứng cứ tiêu cực*, nghĩa là, *thiếu chứng cứ*. Người ta làm giảm bớt khả năng bỏ lỡ chứng cứ như thế bằng cách mở rộng và tăng cường hệ thống; và người ta cho rằng sự né tránh sẽ trở nên khó khăn do nhu cầu phải giữ kín các hoạt động trong một thời gian dài. Nhưng trong lúc khủng hoảng người ta cần phải có chứng cứ chắc chắn hơn; người ta không có thời gian để tìm kiếm và lần theo các manh mối; chẳng có thời gian để thử nghiệm hệ thống và mở rộng nó hoặc tăng cường nó nếu nó không hiệu quả. Do đó, một hiệp định giải quyết khủng hoảng phải dựa trên *chứng cứ tích cực*. Thay vì tìm kiếm chứng cứ về những gì phe kia đang *không* làm, ta cần phải có chứng cứ cho thấy họ *đang* làm gì. Và lý do tại sao chứng cứ như thế lại có thể xuất hiện trong một cuộc khủng hoảng chính là động cơ cung cấp nó - tính cấp bách phải đạt được một sự hiểu biết hoặc thỏa thuận tùy thuộc vào chứng cứ đó lớn hơn - có thể được tăng thêm trong một tình trạng khẩn cấp như thế.

## XÂY DỰNG HỆ THỐNG QUÁ MỨC

Với mục đích ít nhất phải chuẩn bị một mức độ nào đó cho các cuộc khủng hoảng và những tình huống ngoài dự kiến, có một lý lẽ vững chắc để tiến hành một số dàn xếp dự phòng linh động để liên lạc với các kẻ thù tiềm năng và kiểm tra lẫn nhau. Đặc biệt có một lý lẽ vững chắc để xây dựng một hệ

thống kiểm tra quá mức liên quan đến việc sử dụng như đã được thỏa thuận. Việc có khả năng sẵn sàng mở rộng và tăng cường hệ thống, hoặc tăng thêm các phương tiện và thanh sát viên cho hệ thống đó, có thể có liên quan rất nhiều tới tính hữu dụng của hệ thống trong thời điểm khủng hoảng. Nói khác đi, chúng ta không nên đánh giá tính đáng tin cậy và tính hữu dụng của một hệ thống chỉ về mặt động cơ của các bên tham gia trong hoạt động "chuẩn tắc". Chúng ta nên nhìn nhận rằng các cơ hội có thể nảy sinh khi có một động lực mạnh mẽ để mở các cuộc thương lượng cấp thời về những biện pháp hạn chế vũ khí, ít ra là những hạn chế tạm thời, mà chẳng có thời gian để thiết lập các hệ thống quan sát và thông tin liên lạc *bất thường*.

Nói cụ thể hơn: trong trường hợp cần phải thiết lập một hệ thống kiểm tra để theo dõi một hiệp định nhằm định chỉ các cuộc thử nghiệm hạt nhân, chúng ta nên thận trọng cân nhắc xem làm cách nào cả hai bên có thể tận dụng các thanh sát viên và những phương tiện của họ trong trường hợp xảy ra một cuộc khủng hoảng quân sự gay gắt. Tính lưu động của các thanh sát viên, vị trí của họ, các phương tiện thông tin liên lạc của họ, thiết bị huấn luyện kỹ thuật và giám sát của họ, tính đáng tin cậy của họ, và số người trong nhóm của họ, cần phải được đánh giá và trù tính không chỉ với ý định dò tìm các vụ thử hạt nhân, mà còn với một quan điểm về việc họ đáp ứng nhu cầu hết sức quan trọng là cần có phương tiện thanh tra, xác minh, và liên lạc, trong một cuộc khủng hoảng đe dọa cả hai bên về một cuộc chiến ngoài ý muốn.

Từ những điểm vừa đề cập, thì vẫn chưa rõ ràng rằng tính ổn định của sự cân bằng sức hủy diệt - tức là việc thiếu sự kích thích để chủ định tấn công bất ngờ, và tính miễn nhiễm của tình huống này với báo động giả - sẽ bị tác động mạnh bởi những hoạt động bố trí quân sự mà Mỹ cố tiến hành với người Nga. Khi thiên nhiên tiết lộ những bí mật về khoa học và kỹ thuật của nó trong những năm tới, có thể thấy rằng mỗi bên (nếu họ làm những gì phải làm và làm đủ nhanh chóng) có thể bảo đảm một cách đáng kể tính không dễ bị tấn công của các lực lượng trả đũa của mình bất luận bên kia có làm gì đi nữa, và bảo đảm điều đó một cách thuyết phục, để đưa đến một khả năng ngăn chặn lẫn nhau ở mức ổn định cao. Cách khác, thiên nhiên có thể đã đặt những bí mật xấu xa ở phía trước, để người Mỹ và người Nga liên tục tìm ra những phương thức mới để tiêu diệt các lực lượng trả đũa với một tốc độ nhanh hơn tốc độ mỗi bên tìm những phương thức mới để bảo vệ chúng. Chỉ có một hy vọng - không có suy đoán - đó là ngay cả với tài năng và các hành động tốt nhất của hoạt động ngoại giao người Mỹ và người Nga sẽ tìm ra những biện pháp hợp tác nhằm ngăn chặn xu hướng dẫn tới tình trạng bất ổn. Do vậy người ta có thể tìm được sự ổn định mà không cần hợp tác, hoặc không thể tìm được nó dù có hợp tác. Hơn nữa, một loại hợp tác nào đó với người Nga, hoặc sự hạn chế lẫn nhau, chính thức hay không chính thức, ngầm hay công khai, có thể tỏ ra là tạo ra được một sự khác biệt đáng kể về tính ổn định của sự cân bằng sức hủy diệt. Và tất nhiên khả năng rủi ro thì rất cao. Như vậy dù Mỹ không thể chắc chắn được rằng một chính sách thận trọng về hợp tác để làm cho các lực lượng trả đũa của mỗi bên không dễ bị tấn công sẽ tạo

ra sự khác biệt, Mỹ phải cân nhắc rằng điều đó có thể xảy ra và Mỹ phải tự hỏi liệu trong thực tế Mỹ có muốn một sự cân bằng ổn định một cách hoàn hảo về việc ngăn chặn nếu có cơ hội lựa chọn hay không. Liệu Mỹ có thực sự quan tâm tới một kế hoạch chống tấn công bất ngờ tổng quát và hiệu quả hay không nếu Mỹ biết về một kế hoạch như thế, và nếu Mỹ nghĩ rằng người Nga sẽ chấp nhận nó?

Cho dù thật yên tâm khi biết rằng người Nga sẽ không bị kích thích để lên kế hoạch tấn công lén có chủ đích, thật yên tâm khi biết được rằng họ rất chắc chắn rằng Mỹ sẽ không có làm điều đó đến độ họ sẽ không bao giờ phải vội vã sử dụng đến vũ khí trong hoảng loạn, tuy nhiên có thể lập luận rằng khả năng Mỹ ngăn chặn mọi sự trừ một cuộc đột kích lớn nhằm vào mình thì ít ra cũng tùy thuộc phần nào vào niềm tin của người Nga rằng Mỹ có thể bị xúi giục vào việc tấn công có chủ đích. Người Nga có thể không tin điều này nếu các lực lượng trả đũa của họ không dễ bị tấn công một cách đáng kể trước đòn tấn công đầu tiên của Mỹ. Có thể tranh luận rằng ngoại trừ khi bị khiêu khích tột độ Mỹ sẽ tránh né bắt cứ cuộc tấn công trả đũa nào mà không có nhiều cơ may xóa bỏ hoặc làm suy yếu sự phản công của Nga. Theo lập luận này, một cặp Chỉ huy Lực lượng Không quân Chiến lược (VAC) *không dễ bị tấn công* là một cặp Chỉ huy Lực lượng Không quân Chiến lược *bị vô hiệu hóa*; và cho dù đó có thể là loại tốt nhất trong một thế giới lưỡng cực triệt để, nó lại là một món hàng xa xỉ mà Mỹ không thể mua sắm trong thế giới hiện nay - một thế giới trong đó có một “vùng thứ ba” rộng lớn là nơi Mỹ mong ước ngăn chặn sự gây hấn của Nga bằng một sự đe dọa đáng tin hơn so với sự đe dọa của việc cung tự sát.

Liệu Mỹ có thể đe dọa trả đũa, không chỉ để kháng cự một cách cục bộ, nếu quân Nga chắc chắn có khả năng quân sự để giáng lại Mỹ một đòn với quy mô tùy ý họ muốn không? Khi mà mỗi bên không dễ bị tấn công trước bên kia, thì các lực lượng chiến lược có vai trò gì ngoài vai trò vô hiệu hóa lẫn nhau và bảo đảm, nhờ sự đồng tồn tại, rằng cả hai bên đều không sử dụng lực lượng của mình?

Có một vai trò. Các lực lượng chiến lược sẽ vẫn có khả năng tiến hành “trả đũa” theo nghĩa trừng phạt. Nếu sự đe dọa đánh bại các thành phố của Nga và Trung Quốc từ ban đầu được cho là có tiềm năng bởi vì có thể kéo theo nỗi đau đớn, thiệt hại về kinh tế, sự bất ổn về cơ cấu tổ chức, và mất thế diện, chứ không phải chủ yếu bởi vì tư thế quân sự của địch quân trong chính khu vực họ gây hấn sẽ bị ảnh hưởng nặng nề, thì yếu tố chính của sự đe dọa vẫn còn tồn tại mặc dù Bộ Chỉ huy Lực lượng Không quân Chiến lược của bên kia không dễ bị tấn công.

Mỗi đe dọa trả đũa ô ạt, nếu “ô ạt” được hiểu nghĩa là sự trả đũa không giới hạn, thực sự sẽ mất tính đáng tin cậy cùng với việc Mỹ cũng bị mất niềm hy vọng rằng một cuộc tấn công tổng lực được tiến hành một cách chu đáo có thể thành công trong việc ngăn chặn sự phản công trả đũa. Nhưng nếu như có bao giờ Mỹ xem trả đũa giới hạn hoặc theo cấp độ như phương tiện để gây áp lực lên người Nga để họ ngưng những hành động mà Mỹ không thể chịu đựng nổi, hoặc Mỹ xét tới việc mở rộng một cuộc chiến cục bộ giới hạn bên trong biên giới của Nga sao cho duy trì được cái cớ để có hành động quân sự cục bộ nhưng thực sự lại nhắm tới việc thực hiện trừng phạt gây đau thương cho dân thường và đe dọa thêm

nữa, thì loại hành động trả đũa này, và sự đe dọa trả đũa có thể được tăng thêm tính đáng tin cậy kèm theo sự giảm tính có thể bị tấn công của các lực lượng chiến lược của cả hai bên. Thật nghịch lý, điều đó có hiệu quả vì cùng một lý do rằng *mọi loại chiến tranh giới hạn có thể trở nên ít bị giới hạn hơn khi khả năng của một cuộc tấn công tổng lực trở nên không còn thực hiện được*. Nguy cơ liên quan tới sự trả đũa kém toàn diện hơn một chút sẽ giảm so với lúc này bởi vì sự lo sợ một cuộc tấn công trả đũa toàn lực sẽ giảm nhẹ rất nhiều. Nỗi lo sợ rằng sự trả đũa có giới hạn của Mỹ sẽ bị hiểu là làm là bước đầu trong việc khởi xướng một cuộc chiến tổng lực sẽ giảm bớt; người Nga sẽ buộc phải tin rằng thực sự Mỹ đang chuẩn bị tự sát khi hiểu là rằng trả đũa giới hạn của Mỹ là bước khởi đầu trong việc hủy diệt lẫn nhau.

Nói thế không có nghĩa là chúng ta lập luận rằng sự trả đũa hạn chế, gây ra nguy cơ, nếu không phải chắc chắn sẽ gây ra, một sự trả đũa hạn chế đối ứng, không thể dẫn tới sự phá hủy hoàn toàn, hoặc một cách chậm rãi hoặc bằng cách bùng nổ thành những cuộc tấn công trả đũa ngày càng lớn hơn. Vấn đề giới hạn một cuộc chiến tranh trả đũa có thể không dễ hơn vấn đề giới hạn chiến tranh cục bộ, và có thể nó còn gay go hơn. Tuy nhiên, lý luận ở đây không tùy thuộc vào việc làm cho việc trao đổi các đòn tấn công trùng phạt giới hạn có vẻ an toàn và hấp dẫn so với chiến tranh cục bộ giới hạn, nhưng đủ an toàn và đủ hấp dẫn so với chiến tranh tổng lực để trở thành một đe dọa đáng tin (và không phải một sự hù dọa) trong bất cứ trường hợp nào trong đó chúng ta có thể phải dựa vào sự đe dọa trả đũa.

Do đó các lực lượng chiến lược sẽ bị "vô hiệu hóa" chỉ trong khía cạnh các cuộc tấn công tiềm năng lên mỗi bên; chúng sẽ vẫn còn vai trò trừng phạt tạo ra một cơ sở nào đó cho một mối đe dọa ngăn chặn. Trong khi đe dọa về trừng phạt *tổng lực* có thể mất tính đáng tin cậy với việc các lực lượng trả đũa của cả hai phía đều đạt được tính không dễ bị tấn công, thì sự đe dọa trả đũa giới hạn rất có thể đạt được nó. Dù tác động thực có ra sao đi nữa, chúng ta không thể phản kháng lại một thế giới của những Bộ Chỉ huy Chiến lược Không quân không dễ bị tấn công chỉ bằng cách liên hệ tới nhu cầu ngăn chặn tấn công ở vùng thứ ba; cần phải thể hiện rằng một đe dọa tấn công ngăn chặn cụ thể nào đó (tấn công toàn diện) thì có tiềm năng hơn một sự đe dọa (giới hạn) khác.

Chỉ một người cực kỳ lạc quan mới có thể nghĩ rằng chúng ta có thể luôn có một sự lựa chọn rõ ràng là chấp nhận hoặc bác bỏ một kế hoạch mà sẽ bảo đảm cho các lực lượng trả đũa của cả hai bên trở nên hoàn toàn và liên tục không dễ bị tấn công. Nhưng vấn đề về chuyện gì sẽ xảy ra cho sự ngăn chặn tấn công ở vùng thứ ba, và khả năng trả đũa giới hạn mà nó gợi ra, thì thích hợp với câu hỏi chúng ta có thể cho phép chính mình hy vọng về điều gì.<sup>9</sup>

---

<sup>9</sup> Để có thêm thảo luận xin xem T. C. Schelling và Morton H. Halperin, *Strategy and Arms Control*, The Twentieth Century Fund (New York, 1961).

# PHẦN PHỤ LỤC



# PHỤ LỤC A

## VŨ KHÍ NGUYÊN TỬ VÀ CHIẾN TRANH HẠN CHẾ

Với sự phát triển của các loại vũ khí nguyên tử cỡ nhỏ, có tầm tác động nhỏ thích hợp sử dụng trong phạm vi cục bộ cho các lực lượng bộ binh với thiết bị khiêm tốn, và với việc phát triển các loại vũ khí hạt nhân xuyên sâu lòng đất cùng những tên lửa hạt nhân không đối không, các đặc điểm kỹ thuật của vũ khí hạt nhân không còn tạo ra cơ sở đáng kể, nếu có, cho việc xem xét vũ khí hạt nhân như một chủng loại khác biệt so với các loại vũ khí khác trong việc tiến hành một cuộc chiến tranh cục bộ. Dĩ nhiên là đã có tranh luận rằng có những bất lợi chính trị trong việc Mỹ sử dụng vũ khí hạt nhân trong chiến tranh hạn chế, đặc biệt là khi Mỹ sử dụng chúng trước. Ngay cả những người xem một vụ nổ hạt nhân cũng hợp lẽ như là một vụ nổ bom napalm thiêu cháy một con người cho đến chết cũng phải nhìn nhận một yếu tố chính trị về nỗi sợ hãi toàn cầu chống lại vũ khí hạt nhân.

Phần Phụ lục này bàn về một cơ sở khác nữa để phân biệt vũ khí hạt nhân và các loại vũ khí khác. Nó bao hàm các mối

quan hệ của chúng ta với kẻ thù trong quá trình của việc giới hạn chiến tranh. Vì ích lợi của việc hạn chế chiến tranh hay vì ích lợi của việc hiểu biết chiến tranh giới hạn, có thể là cần thiết phải nhìn nhận rằng có thể tồn tại sự khác biệt giữa vũ khí hạt nhân và các loại vũ khí khác cho dù sự phân biệt không phải là yếu tố vật chất mà là yếu tố tâm lý, yếu tố nhận thức, mang tính pháp lý hay tính biểu tượng. Việc cho rằng các vũ khí hạt nhân có tầm tác động nhỏ được phóng đi với tầm chính xác "cao độ" chỉ là một loại đạn pháo, và do đó không có phân biệt vấn đề các giới hạn trong chiến tranh, chỉ là một lập luận hoàn toàn dựa vào một phân tích về tính hiệu quả của vũ khí, chứ không phải dựa vào một phân tích về quy trình giới hạn - về nơi mà các giới hạn xuất phát trong chiến tranh giới hạn, điều gì làm cho chúng ổn định hay bất ổn định, điều gì tạo ra thẩm quyền cho chúng, và những tình huống hay mô hình hành vi nào mang tính chất dẫn dắt đưa đến việc tìm kiếm và nhìn nhận của cả hai bên về các giới hạn. Cơ sở của lập luận cho rằng "chỉ-là-một-loại-vũ-kí-khác" là, nếu không có cơ sở ràng buộc về tính hiệu quả của vũ khí để tạo lập một ranh giới khác biệt giữa vũ khí hạt nhân và các loại vũ khí khác, thì không có cơ sở nào thích hợp với quy trình hạn chế.

Liệu có phải một quan điểm như thế không áp dụng được cho việc phân biệt những người sử dụng vũ khí hay không? Chẳng có sự khác biệt thêm nào giữa người Nga và người Trung Quốc so với sự khác biệt giữa vũ khí hạt nhân và các loại vũ khí khác; tương tự như thế là sự khác biệt giữa người Trung Quốc và người Bắc Triều Tiên, hoặc giữa người Mỹ và Đài Loan, người Anh và người Jordan, người Ai Cập và người

Angieri. Tuy nhiên tính dân tộc đã là một ranh giới quan trọng trong quy trình hạn chế chiến tranh hay phá hủy những giới hạn. Tương tự như thế, có ít sự khác biệt giữa vùng đất nằm ở 100 dặm về phía bắc của đường biên giới Xô Viết - Iran và vùng đất nằm ở 100 dặm về phía nam, hay giữa vùng đất nằm ở phía trên Yalu và ở phía dưới đó, hay ở hai phía của đường biên giới Hy Lạp - Nam Tư. Tuy nhiên những đường biên giới như thế này đóng một vai trò quan trọng trong quy trình giới hạn, nằm tách biệt hoàn toàn với bất cứ khó khăn vật chất nào trong việc băng ngang qua các dòng sông hay vượt qua những dãy núi mà vô tình trùng hợp với các đường biên giới đó.

Người ta có thể trả lời rằng chúng là những sự phân biệt "pháp lý" và rằng những sự phân biệt pháp lý là những thực thể trong khi sự khác biệt giữa vũ khí hạt nhân và các loại vũ khí khác là sự khác biệt hú cẩu. Nhưng chúng không thật sự mang tính chất pháp lý; chúng mang "tính chất chủ nghĩa pháp lý." Không có thẩm quyền pháp lý nào buộc các bên tham gia trong một cuộc chiến tranh hạn chế phải nhìn nhận các ranh giới chính trị hay các yếu tố dân tộc. Người Nga không bị bắt buộc về mặt pháp lý phải xem việc xâm phạm ở mức thật nhỏ biên giới của họ là một sự thay đổi định tính trong chiến tranh - một hành động nghiêm trọng tách khỏi hành động đụng chạm đến biên giới của họ. Người Trung Quốc không bị bắt buộc về mặt pháp lý phải trả đũa (thay vì chỉ kháng cự) nếu Mỹ cố tình băng ngang dòng sông Áp Lục; họ không mất bất cứ quyền pháp lý nào để từ chối không cho băng ngang bằng cách thỉnh thoảng cho phép lưu thông qua lại. Mỹ không bị ràng buộc về pháp lý phải thừa nhận các phi

công Nga nếu họ tham gia vào cuộc chiến giới hạn, hay những "người tình nguyệt" Nga trong đội bộ binh Cận Đông chiến đấu chống lại Mỹ. Giới hạn đối với việc xâm nhập biên giới, hay việc đưa một yếu tố dân tộc vào một cuộc xung đột, thì cũng tương tự như việc đưa vào sử dụng vũ khí hạt nhân; nó là nguy cơ của việc kẻ thù đáp trả. Và một yếu tố quyết định quan trọng trong sự đáp trả của đối thủ chính là đánh giá của họ về điều mà họ ngầm đồng ý nếu họ không có khả năng đáp trả, hay chỉ thực hiện một sự đáp trả bỗng sung trước hành động không liên tục một cách tượng trưng của Mỹ.

Điều khiến cho biên giới Liên Xô hay Trung Quốc trở thành nơi thích hợp hay bắt buộc để vẽ một đường ranh giới khi xảy ra chiến tranh ở vùng đó thì chủ yếu là do thông thường không có ranh giới nào thích hợp hơn để vẽ. Việc binh lính phương Tây băng qua biên giới Nga đồng nghĩa với việc thách thức - không phải trong thực chất mà mang tính biểu tượng - sự toàn vẹn lãnh thổ của Liên Xô, và biểu lộ hay ít nhất ngũ ý một ý định tiếp tục tiến tới. Trừ phi người ta có thể tìm thấy một giới hạn "hiển nhiên" nào đó bên trong biên giới đó, một giới hạn mà sẽ là rõ ràng đối với người Nga tại nơi mà người ta có ý định ngừng lại sau khi băng qua biên giới, và một giới hạn mà sẽ là hiển nhiên đối với họ rằng có một giới hạn liên quan đến việc người Nga sẽ để cho họ tiến bao xa nếu Mỹ vượt qua nó và rằng người Nga biết rằng Mỹ biết giới hạn đó, thì ngoài ra chẳng còn nơi dừng lại nào mà được cả hai phía ngầm thừa nhận. Trong những tình huống mà Liên Xô chấp nhận sự xâm nhập biên giới mà không có loại trả đũa đáng kể nào sẽ là việc thừa nhận rằng lãnh thổ Xô Viết là một nơi thuận tiện cho việc dần dần mở rộng chiến tranh. Do vậy ranh

giới chính trị là *hữu ích* là nơi để dừng lại, chứ không bắt buộc về mặt luật pháp; nó có ích cho cả hai bên khi thiếu vắng bất cứ một lựa chọn thay thế nào có thể nhận biết được một cách đơn giản, bởi vì cả hai bên đều quan tâm đến việc tìm kiếm giới hạn nào đó. Biên giới có một *tính độc đáo* làm cho nó trở thành một giới hạn thích hợp. Nó là một trong số ít những lằn ranh - có lẽ là lằn ranh duy nhất, nhưng chắc chắn là một trong số ít - mà người ta có thể vẽ trong một vùng mà có thể được ngầm thừa nhận bởi cả hai bên như một giới hạn địa lý "hiển nhiên" mà cả hai bên phải tuân thủ. Nó có một *sức mạnh gợi ý* ràng buộc, một công bố cần phải lưu tâm, và việc phủ nhận nó có thể dường như - khi thiếu vắng bất cứ một lựa chọn thay thế nào có thể nhận biết được một cách đơn giản - là sự phủ nhận bất cứ giới hạn nào khác.

Nhưng, nếu các nhận thức về ranh giới chính trị hay yếu tố dân tộc có vẻ vẫn mang tính chất pháp lý, và do vậy là có thực, thì hãy xem xét một số cách phân biệt khác mà có ý nghĩa trong quy trình giới hạn. Mỹ đã cung cấp khá nhiều thiết bị nhưng không cung cấp nhân lực cho cuộc chiến tranh ở Đông Dương. Mỹ đã cung cấp thiết bị, kỹ năng lãnh đạo và tư vấn cho binh lính Hy Lạp trong cuộc chiến tranh du kích, nhưng không cung cấp đội quân chiến đấu. Mỹ đã cung cấp sự hỗ trợ hải quân trực tiếp cho Đài Loan trong khu vực Vịnh Đài Loan. Người ta đã nghĩ rằng có thể Mỹ đã cung cấp sự hỗ trợ không lực cho quân Pháp ở Đông Dương, mà không xuất đầu lò diện trước người Trung Quốc và người Nga theo kiểu Mỹ "can dự" vào bằng cách đưa bộ binh tham chiến.

Một nhà kinh tế học có thể tranh luận - với cùng khả năng thuyết phục như những người lập luận rằng các loại vũ khí

tác động nhỏ có độ chính xác “cực cao” chỉ là một loại súng pháo khác – rằng thiết bị và nhân lực là những nguồn lực thay thế được trong một chiến dịch quân sự, rằng sự can thiệp không lực thì không “thật sự” khác biệt với sự can thiệp bằng bộ binh, rằng trí tuệ quân sự thì cũng quan trọng như cơ bắp chân đối với những chiến binh thiếu kĩ năng lãnh đạo và hoạch định. Cuộc tranh luận về việc định nghĩa lại các chức năng dịch vụ dưới ánh sáng của các loại vũ khí hiện đại, và về tính hữu dụng của việc định nghĩa các chức năng dịch vụ quân sự xét theo phương tiện di chuyển, gợi mở rằng một sự phân biệt không lực-bộ binh hay sự phân biệt hải quân-bộ binh chẳng lệ thuộc vào điều gì khác ngoài truyền thống. Nhưng điểm mấu chốt của tất cả vấn đề này là, trong việc giới hạn chiến tranh, truyền thống thật sự quan trọng.

Thực ra, điều chúng ta đang bàn luận trong việc phân tích chiến tranh hạn chế chính là truyền thống. Chúng ta đang bàn về tiền đề, quy ước và sức mạnh của gợi ý. Chúng ta đang bàn lý thuyết về luật bất thành văn – với những quy ước mà sự thừa nhận của chúng trong tổng thể là nhu cầu cùng chấp nhận lẫn nhau để tránh khỏi phải hủy diệt nhau, và sự thừa nhận của chúng trong mỗi một trường hợp riêng lẻ là nguy cơ vi phạm một quy tắc có thể làm sụp đổ nó, và rằng làm sụp đổ nó có thể dẫn đến một giới hạn cùng ít mong muốn hơn hoặc không còn giới hạn nào cả, và có thể làm suy yếu thêm các quy tắc vẫn chưa đổ vỡ bằng cách cung cấp chứng cứ rằng “thẩm quyền” của chúng không thể được công nhận.

Điều làm cho các vũ khí nguyên tử khác biệt chính là một truyền thống mạnh mẽ cho rằng chúng là khác biệt. Lý do – trong việc trả lời cho câu hỏi thông thường đại ngôn – tại sao

chúng ta không cầm cung tên trong chiến tranh mặt đất vì chúng, giống như vũ khí hạt nhân, cũng giết người và làm cho tàn tật, chính là vì có một truyền thống cho phép sử dụng cung tên, một kỳ vọng được cùng thừa nhận rằng chúng sẽ được sử dụng nếu việc sử dụng chúng là thích hợp. Không có một truyền thống như thế cho việc sử dụng vũ khí nguyên tử. Thay vào đó có một truyền thống cầm sử dụng chúng - một kỳ vọng được cùng thừa nhận rằng chúng không thể được sử dụng cho dù có các tuyên bố sẵn sàng sử dụng chúng, thậm chí cho dù có các lợi thế chiến thuật để sử dụng chúng.

Các truyền thống hay các quy ước không đơn giản là sự tương tự với những giới hạn trong chiến tranh, hay là một khía cạnh khác lạ của những giới hạn đó. Truyền thống hay tiền đề hay quy ước là cốt lõi của các giới hạn. Đặc điểm cơ bản của bất cứ giới hạn nào trong một cuộc chiến tranh giới hạn chính là đặc điểm tâm lý, tri thức hay xã hội của việc cùng được thừa nhận bởi cả hai bên như là có một loại thẩm quyền nào đó, loại thẩm quyền xuất phát chủ yếu từ việc nhận thức đơn thuần về sự thừa nhận chung, về một sự "mặc cả ngầm". Và một giới hạn đặc biệt có thêm thẩm quyền từ việc thiếu vắng sự tin tưởng mà một trong hai bên có thể đặt vào nơi mà các giới hạn thay thế khác có thể được tìm thấy nếu giới hạn đó không được tuân thủ. Cơ sở dằng sau giới hạn này mang tính chất chủ nghĩa pháp lý và nguy biện, chứ không mang tính pháp lý, đạo đức hay vật chất. Các giới hạn này có thể tương ứng với các khác biệt về pháp lý hay vật thể hay sự phân biệt về đạo đức. Thực vậy, chúng thường tương ứng với điều gì đó tạo cho chúng một tính độc đáo và định tính và cung cấp cho chúng một tiêu điểm nào đó để các kỳ

vọng hối tụ. Nhưng thẩm quyền nằm ở trong chính các kỳ vọng, chứ không phải ở trong sự việc mà các kỳ vọng được gắn kết vào.

Liệu các giới hạn về việc sử dụng các vũ khí nguyên tử, chứ không phải một giới hạn đặc biệt cấm sử dụng chúng, có thể được định nghĩa một cách hợp lý lại bị làm cho mơ hồ hơn, chứ không phải ít mơ hồ hơn, bởi tính chất đa dụng gia tăng của các vũ khí nguyên tử. Ngày nay người ta thừa nhận rộng rãi rằng có một cấp độ gia tăng khá liên tục tầm vóc hiệu quả của vũ khí hạt nhân có thể có, một biến thể khá liên tục về hình thức có thể được sử dụng, về phương tiện phóng, về các mục tiêu mà chúng có thể nhắm tới, và vân vân. Do vậy có vẻ như không có sự đứt mạch “tự nhiên” giữa một số cách sử dụng giới hạn nào đó và những cách sử dụng khác. Vậy thì nếu chúng ta hỏi đâu là lằn ranh chúng ta có thể vẽ ra nếu chúng ta mong muốn cách nào đó hạn chế tầm vóc của vũ khí, phương tiện phóng, cũng như các tình huống mà trong đó hay mục tiêu mà trên đó vũ khí nguyên tử có thể được sử dụng, câu trả lời là chúng ta - theo một nghĩa thuận về kỹ thuật - được tự do để vẽ lằn ranh ở bất cứ chỗ nào chúng ta muốn. Không có một lý do thuyết phục nào để vẽ lằn ranh đó ở bất cứ một lằn mức cụ thể nào thay vì ở một lằn mức khác. Nhưng đó chính là lý do tại sao rất khó để tìm ra một cơ sở cho bất cứ một lằn ranh cụ thể nào. Không có cấp độ sử dụng, hay kích cỡ của vũ khí, hay số dặm được xem là thích hợp hơn các cấp độ khác, các kích cỡ khác hay các khoảng cách khác mà lại cung cấp được một tiêu điểm cho sự kỳ vọng của cả hai bên. Các hạn chế theo chủ nghĩa pháp lý cần phải có tính chất

định tính và riêng biệt, thay vì định tính và liên tục. Đây không chỉ là vấn đề làm cho việc vi phạm dễ bị nhận biết hơn, hoặc làm cho việc tuân thủ dễ áp dụng hơn đối với các chỉ huy của riêng một bên; nó liên quan đến nhu cầu cần có bất cứ giới hạn ổn định nào để có được một đặc điểm hiển nhiên biểu tượng, ở mức mà vi phạm vào điều đó là một hành động công khai và mạnh mẽ có thể đặt cả hai bên trước nguy cơ rằng các giới hạn thay thế khác sẽ không dễ dàng tìm được.

Nhu cầu cần có các giới hạn dễ phân biệt về mặt định tính mà mang tính nhất độc đáo nào đó thì đặc biệt được nâng cao bởi yếu tố rắn; các giới hạn thường được tìm thấy qua một quy trình vận động và thương lượng ngầm. Chúng được tìm thấy bằng mưu mẹo hơn là bằng thương lượng công khai. Nhưng nếu hai bên phải tiến hành một "cuộc mặc cả" mà không có sự liên lạc công khai, thì giới hạn cụ thể buộc phải có một tính chất nào đó để phân biệt nó với một thể liên tục các giới hạn hay thế có thể có; nếu không sẽ ít có cơ sở cho sự tin tưởng của mỗi bên rằng bên kia cũng thừa nhận cùng một điều gì đó. Ngay cả một đường song song vĩ tuyến, hay một mũi giờ quốc tế, hay cực bắc, có thể có tính chất này khi không có điểm hay đường thẳng tự nhiên, hợp lý, "hiển nhiên" khác để cho các kỳ vọng hội tụ lại.

Một sự kiểm chứng luận điểm này liên quan đến các loại vũ khí hạt nhân có thể đặt ra vấn đề sau đây.<sup>1</sup> Giả sử bất cứ ai trong chúng ta tìm cách hợp tác để giành một giải thưởng: chúng ta phải ngồi xuống ngay lúc này, ngồi riêng biệt ra và không có bất cứ sự thỏa thuận trước nào, và viết ra một hạn chế được đề nghị về việc sử dụng các loại vũ khí hạt nhân,

<sup>1</sup> The New York Times, ngày 29-12-1957, trang 20.

nhiều hay ít chi tiết tùy ý, cho phép chúng ta có các giới hạn về bất cứ mô tả nào hấp dẫn đối với chúng ta - tầm vóc của vũ khí, việc sử dụng vũ khí, ai sẽ là người sử dụng chúng, tần suất sử dụng, bom sạch đối nghịch với bom bẩn, sử dụng tấn công đối nghịch với sử dụng phòng vệ, chiến thuật chống lại chiến lược, nhắm vào hay không nhắm vào các thành phố, có hay không có cảnh báo - để xem liệu tất cả chúng ta có thể viết ra *cùng* một sự chỉ định về giới hạn hay không. Nếu chúng ta thống nhất một cách hoàn hảo về các giới hạn mà chúng ta chỉ định, chúng ta giành được một phần thưởng; nếu các giới hạn của chúng ta là khác nhau, chúng ta không có phần thưởng nào. Chúng ta đang làm điều này chỉ để giành giải thưởng, để xem trong thực tế chúng ta có thể đồng ý ngầm về việc trình bày các giới hạn hay không, và để xem - đối với những người trong chúng ta có cố gắng tìm cách phối hợp ngầm các đề xuất của mình - các loại giới hạn nào có vẻ dễ có được thừa nhận chung ngầm. Chúng ta được phép đi đến các thái cực một mặt không có giới hạn nào cả, và mặt khác là không sử dụng vũ khí hạt nhân gì cả, và bất cứ lằn mức nào hay biến thể nào được xác định theo bất cứ cách nào chúng ta muốn.

Lập luận của tôi là có những giới hạn cụ thể - các giới hạn đơn giản, riêng biệt, định tính, "hiển nhiên" - mà dẫn dắt đến một sự lựa chọn có phối hợp; những người liệt kê các loại giới hạn khác, theo tôi phỏng đoán, có thể tìm thấy rất ít đồng minh hoặc không có ai cả có giới hạn trùng khớp với giới hạn của họ. (Trong bài toán này cần nhắc chủ yếu để lựa chọn bất cứ các giới hạn cụ thể nào là khả năng nếu chúng ta lựa chọn các giới hạn đó trong nỗ lực để trùng khớp chính xác với các

giới hạn của người khác, biết rằng họ đang cố gắng phối hợp sự chọn lựa của họ với sự chọn lựa của chúng ta, thì chúng ta sẽ thành công).

Tôi không tuyên bố rằng bài tập này chứng minh loại giới hạn nào có khả năng sở hữu tính ổn định hay tính thẩm quyền. Nó có thể hiện rằng các đặc tính nào đó của các giới hạn, đặc biệt là tính đơn giản của chúng, tính độc đáo, tính riêng biệt, tính nhạy cảm của định nghĩa định tính, và vân vân, có thể nhận được một ý nghĩa khách quan, một ý nghĩa mà ít ra là thích hợp với quy trình thương lượng ngầm. Nó gợi ý rằng một số loại giới hạn nào đó có khả năng cùng được kỳ vọng bởi hai bên, có khả năng tập trung sự kỳ vọng, và có khả năng được nhìn nhận là tách biệt về mặt định tính so với một thể liên tục các lựa chọn thay thế khác.

Kết luận đầu tiên rút ra từ chuỗi lập luận này là có một sự tách biệt giữa các loại vũ khí hạt nhân và vũ khí phi hạt nhân, một sự tách biệt tương ứng với quy trình hạn chế chiến tranh. Đó là một sự tách biệt mà ở một mức độ nào đó Mỹ có thể tăng cường hay làm suy yếu đi, làm rõ ràng hay làm mờ nhạt đi. Mỹ có thể tăng cường truyền thống, và nâng cao ý nghĩa tương trưng của sự tách biệt này, bằng cách nói và hành động hết sức nhất quán theo sự tách biệt đó; Mỹ có thể làm suy yếu sự tách biệt này - nhưng không dễ dàng hủy bỏ nó - bằng cách hành động như thể Mỹ không tin vào điều đó, bằng cách nhấn mạnh lập luận "chỉ là một loại vũ khí khác", và bằng cách thể hiện rõ ràng rằng trong thực tế Mỹ sẽ ít hối tiếc về việc sử dụng vũ khí hạt nhân. Việc Mỹ sẽ theo đuổi chính sách nào tùy thuộc vào việc liệu Mỹ có xem sự tách biệt giữa vũ khí hạt nhân và các loại vũ khí khác như là một giá trị mà Mỹ chia

sẽ với Liên Xô, một sự tách biệt hữu ích, một truyền thống mà giúp giảm thiểu bạo lực - hay thay vào đó như là một sự khó chịu, một trách nhiệm tuyên truyền, một rào cản ngoại giao, và là một sự ngăn cấm đối với hành động quyết định và sự phân cấp thẩm quyền. Những người tin tưởng rằng vũ khí hạt nhân cần phải được sử dụng sớm nhất khi thuận tiện, hay vào bất cứ khi nào sự thích hợp về mặt quân sự đòi hỏi, cũng vẫn cần phải nhìn nhận sự tách biệt này tồn tại để Mỹ có thể làm suy yếu nó trong thời gian chuyển tiếp.

Đây không chỉ là vấn đề các nước trung lập châu Á hay các đồng minh châu Âu của Mỹ cảm nhận như thế nào về sự tách biệt. Nó liên quan đến một mối quan hệ giữa Mỹ và Nga - một sự hiểu biết mà có thể tồn tại cho dù Mỹ có thích điều đó hay không. Nó liên quan đến việc liệu người Nga có nghĩ rằng Mỹ chia sẻ với họ một kỳ vọng ngầm rằng có một giới hạn chống lại việc sử dụng vũ khí hạt nhân. Vì lợi ích của việc hạn chế chiến tranh, Mỹ cần phải mong muốn người Nga hay người Trung Quốc không tin rằng việc sử dụng vũ khí hạt nhân ban đầu của Mỹ trong một cuộc chiến cục bộ là một thách thức đối với ý tưởng toàn bộ về các biện pháp giới hạn, hay là một lời công bố rằng Mỹ sẽ không bị ràng buộc bởi bất cứ một loại giới hạn nào. Mỹ cần mong muốn họ giải thích việc sử dụng vũ khí hạt nhân của Mỹ là nhất quán với khái niệm chiến tranh giới hạn và nhất quán với mong muốn của Mỹ được cộng tác ngầm trong việc phát hiện và thừa nhận các giới hạn. Mỹ cần mong muốn việc sử dụng vũ khí hạt nhân của Mỹ không bị quy kết về nội dung biểu tượng thái quá. Do vậy, nếu tôi lập luận đúng rằng một sự tách biệt có tồn tại theo một ý nghĩa thích hợp với việc hạn chế chiến tranh, và nếu

cho dù Mỹ muốn có được sự tự do tối đa để sử dụng vũ khí hạt nhân, thì vì lợi ích của việc hạn chế chiến tranh Mỹ phải hủy bỏ hay làm suy yếu ở mức tối đa có thể sự tách biệt này. (Chẳng hạn như một chương trình có chủ đích về việc sử dụng sớm và ở mức độ cao “chất nổ hạt nhân” trong các dự án chuyển dời đất, đặc biệt là tại các nước kém phát triển, có thể giúp làm suy yếu sự tách biệt này. Tương tự như thế là một chương trình huấn luyện các lực lượng thân thiện ở các nước kém phát triển về việc làm cách nào sống sót qua những vụ nổ hạt nhân, sử dụng một số vũ khí thật sự cho mục đích này trong nước họ). Nếu ngược lại Mỹ mong muốn nâng cao sự hiểu biết ngầm cùng với các đối thủ rằng vũ khí hạt nhân là một loại khác hẳn và tùy thuộc vào các điều kiện đặc thù nào đó, thì việc thỏa thuận về ngăn chặn thử nghiệm hạt nhân (hoặc thậm chí chỉ là việc thảo luận sâu về một thỏa thuận như thế) có lẽ sẽ đóng góp cho mục đích này.<sup>2</sup>

Một kết luận thứ hai là sự ngăn chặn cơ bản đối với việc sử dụng các loại vũ khí nguyên tử trong cuộc chiến tranh hạn chế có thể biến mất ngay khi sử dụng chúng lần đầu. Thật khó mà tưởng tượng được thỏa thuận ngầm rằng vũ khí hạt nhân là một chủng loại khác cũng sẽ có hiệu lực mạnh mẽ khi xảy ra một cuộc chiến tranh hạn chế *kế tiếp* sau khi chúng đã được sử dụng trong một cuộc chiến rồi. Do đó, có lẽ Mỹ không thể phớt lờ sự tách biệt và sử dụng vũ khí hạt nhân trong một cuộc chiến đặc biệt khi mà việc sử dụng chúng có thể mang

---

<sup>2</sup> Về ý nghĩa tương trưng của một thỏa thuận thử nghiệm, xem thêm Henry A. Kissinger trong “Thử nghiệm Hạt nhân và Vấn đề Hòa bình” (*Nuclear Testing and the Problem of Peace*), *Foreign Affairs*, 37:1-18 (tháng 10-1958), đặc biệt là các trang 12-13.

lại lợi thế cho Mỹ và sau đó dựa vào sự tách biệt với hy vọng rằng cả Mỹ và kẻ thù có thể thoái lui. Một khả năng hạn chế chiến tranh tiềm năng sẽ bị mất giá trị một cách đáng kể bất cứ lúc nào nếu Mỹ phá vỡ truyền thống và tạo ra một tiền lệ trái ngược. (Cũng có thể có một số giới hạn hay một số khái niệm riêng được Mỹ công nhận mà cần được khảo sát lại để xem liệu chúng có phải tự nguyên gốc là những sản phẩm phụ của một lệnh cấm hạt nhân được giả thiết và có thể biến mất cùng với nó hay không. Mỹ có thể muốn nhìn lại vai trò của các tàu hải quân chẳng hạn, một phần để dự đoán trước cách kẻ thù đối phó với chúng, một phần để tránh diễn giải sai các ý định của kẻ thù nếu kẻ thù đối phó với chúng một cách khác sau khi vũ khí hạt nhân đã được đưa vào sử dụng).

Một kết luận thứ ba là khi xảy ra việc sử dụng vũ khí hạt nhân lần đầu, Mỹ có lẽ ít nhất cũng phải quan tâm đến các xu hướng và các tiền lệ mà Mỹ đã thiết lập, và đến "vai trò hạt nhân" mà Mỹ đã áp dụng tương tự như cách mà Mỹ quan tâm đến các mục tiêu ban đầu của chiến tranh giới hạn. Chẳng hạn như, nếu các loại vũ khí hạt nhân được sử dụng cho việc bảo vệ Kim Môn, có lẽ hẳn là Mỹ quan tâm đến hậu quả đối với Kim Môn ít hơn là quan tâm đến đặc tính của việc tấn công bằng vũ khí hạt nhân, các tiền lệ mà nó thiết lập, vai trò mà Mỹ cố gắng đảm nhận, và vai trò mà đối thủ đảm nhận trong quá trình đó. Mỹ sẽ không chỉ sử dụng chúng một cách *bất thường* cho cuộc chiến nhỏ đang đề cập, mà quan trọng hơn là định dạng các cuộc chiến hạt nhân trong tương lai.

Thứ tư, chúng ta nên nhận rằng - ít ra là khi vũ khí hạt nhân được sử dụng lần đầu trong một cuộc chiến tranh hạn chế - đối thủ cũng sẽ thực sự tham gia vào ít nhất hai loại

hành động chiến tranh hạn chế khác nhau cùng một lúc. Một loại sẽ là đấu tranh hạn chế đối với các mục tiêu ban đầu; loại hai sẽ là thương lượng ngầm hoặc nghi binh về vai trò của các loại vũ khí hạt nhân. Để minh họa, trong việc liên hệ đến Kim Môn, Mỹ có thể quyết định sử dụng vũ khí hạt nhân. Thông thường người ta sẽ giả thiết rằng Mỹ có lẽ sẽ thực hiện điều này chỉ khi nào thật sự cần thiết để bảo vệ Kim Môn, và rằng Mỹ có lẽ sẽ sử dụng chúng theo một cách nhằm đạt được các mục đích về Kim Môn của họ. Nhưng, khi cân nhắc việc người Trung Quốc hay người Nga sẽ sử dụng chúng để đáp trả, có lẽ người Mỹ không nên chỉ lo lắng chủ yếu về việc đối thủ suy nghĩ việc sử dụng vũ khí hạt nhân của Mỹ sẽ có tác dụng gì cho việc xâm chiếm Kim Môn. Dường như đối với họ bản chất của việc họ "đáp trả" sáng kiến sử dụng hạt nhân của Mỹ là điều quan trọng hơn nhiều. Họ sẽ quan tâm đến việc không nhận lãnh vai trò quy phục, mà đòi hỏi một loại "bình đẳng" nếu không muốn nói là ưu thế trong vai trò hạt nhân của chính họ. Và, trừ phi Mỹ sẵn sàng cho một loại đối đầu quyết định nào đó mà trong đó Mỹ hoặc là được tất cả hoặc là mất tất cả, Mỹ cũng cần phải có thiện chí "thương lượng" (bằng hành động của mình) để có được các mục tiêu hạn chế xét theo ưu thế hạt nhân, các truyền thống và các tiền lệ của việc sử dụng hạt nhân, và các "luật chơi" mà Mỹ tạo ra cho các cuộc chiến tranh trong tương lai, tương tự như bất cứ loại mục tiêu nào khác trong chiến tranh hạn chế.

## VỀ VIỆC TỪ BỎ TÍNH ĐỐI XỨNG TRONG LÝ THUYẾT TRÒ CHƠI

Phần đầu tiên của phụ lục này lập luận rằng trò chơi thuần măc cả “bất chuyển động” như được phân tích bởi Nash, Harsanyu, Luce và Raiffa và một số tác giả khác<sup>1</sup> có thể là không tồn tại, hoặc nếu nó tồn tại, thì mang một đặc tính khác so với những gì đã được nêu ra một cách chung. Điểm xuất phát của lập luận này là ý nghĩa hoạt động của *thỏa thuận*, một khái niệm mà hầu như một cách bất biến bị bỏ mặc không được định nghĩa. Phần hai của phụ lục này lập luận rằng tính đối xứng trong kết quả của các trò chơi măc cả không thể được hỗ trợ xét trên khái niệm “các kỳ vọng hợp lý”; xuất phát điểm lập luận này là sự nhận dạng cần thiết về các kỳ vọng hợp lý.

---

<sup>1</sup> John F. Nash, “Vấn đề măc cả” (The Bargaining Problem), *Econometrica*, 18:155-162 (tháng 4-1950), và “Các trò chơi hợp tác hai người” (Two-Person Cooperative Games), *Econometrica*, 21:128-140 (tháng 1-1953); John Harsanyi, “Các phương pháp tiếp cận vấn đề măc cả trước và sau lý thuyết trò chơi: một thảo luận quan trọng về các lý thuyết của Zeuthen, Hicks và Nash” (Approaches to the Bargaining Problem Before and After the Theory of Games: a Critical Discussion of Zeuthen's, Hicks', and Nash's Theories), *Econometrica*, 24:144-157 (tháng 4-1956); R. Duncan Luce và Howard Raiffa, *Games and Decisions* (Trò chơi và các quyết định) (New York, 1957), trang 114”.

Một trò chơi có tổng lợi ích khác không (“hợp tác”) không ngầm - trò chơi mặc cả - không được *định nghĩa* theo ma trận lợi ích của nó; các hoạt động mà qua đó các lựa chọn được thực hiện vẫn cần phải được xác định. Thông thường các hoạt động này được phác họa qua việc liên hệ đến ý niệm về “các thỏa thuận ràng buộc” và ý niệm về tự do trao đổi thông tin trong quá trình đạt được thỏa thuận. Như vậy khi nói rằng hai người chơi có thể chia nhau 100 đô-la ngay khi họ có thể đồng ý về việc chia như thế nào, và rằng họ có thể thảo luận vấn đề đầy đủ với nhau, thì thường được xem là đủ để xác định trò chơi.<sup>2</sup>

Một trò chơi thuộc loại này thì mang tính đối xứng trong cấu trúc nước đi của nó, cho dù nó có thể bất đối xứng trong cấu hình của các lợi ích. Cả hai người chơi có các đặc quyền giống nhau về truyền đạt thông tin, về việc từ chối các đề nghị, và về việc đạt thỏa thuận. Nếu thay vì chia 100 đô-la hai người chơi lại đồng ý với nhau về các giá trị X và Y chứa đựng trong một phạm vi, hàm lợi ích có thể là không đối xứng nhưng cấu trúc nước đi là có. Để nhấn mạnh điểm này, thậm chí Harsanyi đã bổ sung một cách công khai định đè về các nước đi đối xứng: “Các bên mặc cả tuân theo các quy tắc (đối xứng) đồng nhất về hành vi (bất luận bởi vì họ tuân theo các

---

<sup>2</sup> Luce và Raiffa, trong thực tế, xác định các trò chơi hợp tác hai người bằng cách liên hệ đến một ma trận lợi ích và ba điều kiện sau đây: (1) Tất cả các thông điệp trước cuộc chơi do một người chơi trình bày được chuyển tải không bị bóp méo qua cho người chơi kia. (2) Tất cả các thỏa thuận đều mang tính chất ràng buộc, và chúng có thể thực thi bằng những quy tắc của trò chơi. (3) Việc đánh giá của một người chơi về các kết quả của trò chơi không bị gây nhiễu bởi các thương lượng trước trò chơi. *Games and Decisions* (Trò chơi và các quyết định), trang 114.

nguyên tắc thống nhất về hành vi hợp lý hoặc là bởi vì họ bị chi phối theo cùng các luật lệ tâm lý).<sup>3</sup>

Điều tôi muốn làm là xem xét quan niệm “thỏa thuận” này trên cơ sở giả thiết về *tính đối xứng hoàn hảo trong cấu trúc nước đi của trò chơi*, tập trung chú ý vào “các chi tiết pháp lý” của quá trình mặc cả. Chúng ta cũng phải xem xét ý nghĩa của sự “phi thỏa thuận.” Bởi vì bất cứ trò chơi được xác định rõ ràng nào cũng phải có quy tắc nào đó về kết thúc trò chơi, chúng ta trước tiên hãy xem xét các quy tắc kết thúc.<sup>4</sup>

Nếu chúng ta phải tránh việc thêm một chiêu kích hoàn toàn mới vào ma trận lợi ích của mình, dưới dạng các tỷ lệ khẩu trừ, chúng ta phải giả thiết rằng trò chơi được kết thúc đủ sớm sao cho không có những yếu tố như lãi suất được đưa vào trong bức tranh. Ngoài bản thân sự thỏa thuận, chúng ta không muốn phải xem xét *thời điểm* mà tại đó đạt được thỏa thuận. Đây không chỉ là vấn đề thuận tiện; trò chơi ngưng không còn là “bất chuyển động”, ngoại trừ trong những trường hợp rất đặc biệt, trừ phi chúng ta đưa ra quy định này. Bởi vì, nếu các ưu tiên về thời gian của những người chơi có định hình cách nào đó ngoại trừ dạng của một tỷ lệ khẩu trừ thống nhất liên tục, thì trò chơi chính nó thay đổi cùng với thời gian

<sup>3</sup> John Harsanyi, “Các phương pháp tiếp cận vấn đề mặc cả trước và sau lý thuyết trò chơi...”, (*Approaches to the Bargaining Problem Before and After the Theory of Games....*), *Econometrica*, 24:149 (tháng 4-1956).

<sup>4</sup> Mô hình được thảo luận ở đây là trừu tượng, nhân tạo và phi thực tế; nhưng nó vẫn có thuận lợi trong việc giúp kiểm tra liệu *ngay cả trong một mô hình nhân tạo trừu tượng* thì nó có kết quả trong việc làm thành định đê tính đối xứng hoàn hảo trong cấu trúc nước đi hay không, và xem tính bất đối xứng như một trường hợp đặc biệt, và tính đối xứng như trường hợp thông thường hay không.

và một người chơi có thể, trong thực tế, thay đổi chính trò chơi bằng cách không đạt thỏa thuận. Ý niệm về một tỷ lệ khấu trừ thống nhất liên tục thì có lẽ quá đặc biệt để có thể xem như một điều kiện *cần*, và dù sao cũng không được làm thành một định đề mở trong các mô hình đang được xem xét; do vậy chúng ta phải giả thiết rằng trò chơi cách này hay cách khác đã được hoàn tất.

Có lẽ cách đơn giản nhất để kết thúc một trò chơi là cho rung chuông tại một thời điểm được xác định từ trước. Có những cách thức khác, chẳng hạn như cho một trọng tài cứ vài phút lại tung viên (xí ngầu), và cho ngừng trò chơi khi ông ta tung được hai con số 1. (Chúng ta có thể cho trò chơi kết thúc sau khi một số lượt quy định đề nghị bị từ chối, nhưng điều này sẽ làm thay đổi đặc tính của trò chơi bằng cách làm cho một số loại liên lạc nào đó trở thành "các nước đi thực sự" khiến cho trò chơi trở nên khác với trước đó, và tất yếu dẫn chúng ta đến các chiến thuật như là việc hết quyền đề nghị).

Để đơn giản, giả sử rằng trò chơi sẽ được kết thúc tại một thời điểm được xác định trước đối với những người chơi, và để thuận tiện chúng ta hãy gọi khoảnh khắc cuối cùng là "nửa đêm." Nếu thỏa thuận tồn tại thì khi chuông nửa đêm vang lên, những người chơi phân chia những gì họ có được theo cách mà họ đã thỏa thuận; nếu thỏa thuận không tồn tại, những người chơi không nhận được gì.

Kế tiếp, chúng ta muốn nói đến điều gì khi đề cập đến "thỏa thuận"? Để đơn giản, hãy giả sử rằng mỗi người chơi giữ (hay có thể giữ) đề nghị "chính thức" hiện tại của mình được lưu lại theo một cách mà trọng tài có thể thấy được khi chuông

rung lên. Có lẽ anh ta giữ nó bằng cách viết nó lên một tấm bảng mà người chơi kia có thể thấy được; có lẽ anh ta giữ nó trong một phong bì kín để rồi trao cho trọng tài khi chuông rung lên. Có lẽ anh ta giữ nó bằng cách đánh vào một bàn phím riêng ghi nhận lại đề nghị hiện tại của mình đặt trong phòng trọng tài. Khi chuông rung lên, tấm bảng được chụp hình, phong bì được trao lại hay bàn phím được khóa, sao cho trọng tài chỉ cần kiểm tra hai bảng đề nghị “hiện tại” như chúng tồn tại ở vào thời điểm nửa đêm để xem chúng có tương thích hay không. Nếu chúng tương thích, phần có được sẽ được chia theo “thỏa thuận”. Nếu hai người chơi cùng đòi hỏi nhiều hơn mức có được, sự “bất hòa” tồn tại, và hai người chơi không nhận được gì. (Hãy hoãn lại một lát việc phán quyết về điều gì xảy ra nếu hai người chơi cùng yêu cầu ít hơn tổng số có được, liệu họ sẽ nhận được mức như họ yêu cầu hay sẽ không nhận được gì vì thiếu thỏa thuận thích hợp. Và, trong diễn biến kế tiếp, sẽ không thành vấn đề liệu một thỏa thuận đã hoàn tất đạt được trước nửa đêm - nghĩa là tính tương thích của các đề nghị hiện tại xảy ra trước nửa đêm - có kết thúc trò chơi hay không).

Có những cách thức khác để định nghĩa “thỏa thuận” xét theo các hoạt động mà qua đó nó được đạt tới hay được ghi nhận. Nhưng nếu chúng ta tuân theo khái niệm về một *cấu trúc nước đi đôi xứng hoàn hảo*, thì tôi nghĩ rằng nhìn chung chúng sẽ có một đặc tính mà tôi đang cố tách ra để có thể chú ý. Đặc tính đó là như thế này. Phải có một khoảng thời gian tối thiểu để cho một người chơi thực hiện, hay thay đổi, đề nghị hiện tại của anh ta. (Một lần nữa, để đơn giản, chúng ta hãy giả thiết là cùng một hoạt động có thể tạo ra một đề nghị

hay làm thay đổi nó, sao cho chúng ta luôn luôn có thể giả thiết rằng một “đề nghị hiện tại” có tồn tại). Thế thì phải có một khoảnh khắc thời gian quan trọng, một thời điểm giới hạn trước khi tiếng chuông nửa đêm vang lên, đó là khoảnh khắc cuối cùng mà tại lúc đó một người có thể bắt đầu các hoạt động ghi nhận lại đề nghị cuối cùng của mình. Nghĩa là, có khoảnh khắc cuối cùng nào đó trước khi chuông rung lên, vượt quá khoảnh khắc đó thì là quá trễ cho một người thay đổi đề nghị đang hiện hữu của mình. Theo các quy tắc của trò chơi và định đề về tính hợp lý, cả hai người chơi đều biết điều này. Và theo quy tắc của tính đối xứng, khoảnh khắc này phải tương đồng cho cả hai người chơi.

Từ điểm này dẫn đến đặc trưng có ý nghĩa sau đây. Đề nghị cuối cùng về mặt cơ học cũng như pháp lý mà một người chơi có thể đưa ra được là đề nghị mà anh ta nhất thiết phải đưa ra mà không biết đề nghị cuối cùng của người chơi kia là gì. Và đề nghị cuối cùng mà một người chơi có thể đưa ra là đề nghị mà người chơi kia có khả năng là không thể đối phó trong quá trình của trò chơi. Trước cái thời điểm áp chót đó, không một đề nghị nào có được tính chất cuối cùng. Và tại khoảnh khắc cuối cùng đó các người chơi hoặc là thay đổi hoặc là không thay đổi các đề nghị hiện tại của mình, và bắt cứ điều gì họ thực hiện thì được thực hiện trong sự hoàn toàn không hay biết gì về điều người kia đang thực hiện, và mang tính chất cuối cùng.<sup>5</sup>

---

<sup>5</sup> Một cách tinh cù, lập luận này không bị ảnh hưởng bởi việc giả thiết rằng một người chơi có thể thay đổi đề nghị của mình “một cách tức thời” chừng nào mà chúng ta vẫn giữ quy tắc đối xứng rằng cả hai cùng có thể làm điều đó “một cách tức thời như nhau” khi tiếng chuông kết thúc vang lên.

Điều này hẳn phải đúng. Nếu một trong hai người có thể liếc mắt xem đè nghị cuối cùng của người kia kịp lúc để có thể làm được điều gì đó về việc đó, hoặc nếu một trong hai người có thể cho người kia cơ hội liếc xem đè nghị cuối cùng của mình kịp thời để cho người kia đổi phó, thì đó không phải là - và được biết là không phải là - đè nghị cuối cùng.<sup>6</sup>

Nhưng lúc này chúng ta đã đi đến một kết luận quan trọng về trò chơi mặc cả nước đi đổi xứng hoàn hảo. Kết luận đó là nó nhất thiết dẫn đến, tại khoảnh khắc áp chót cụ thể nào đó, một trò chơi mặc cả *ngầm* (bất hợp tác). Và mỗi người chơi biết điều này.

Cách thức nhiều thông tin nhất để mô tả trò chơi không phải là những người chơi phải đạt tới một thỏa thuận công khai vào lúc mà tiếng chuông cuối cùng vang lên hoặc cùng nhau biết trước các phần thưởng. Nó là cách mà họ phải đạt được thỏa thuận công khai vào một khoảnh khắc áp chót cụ thể (và được xác định rõ) - khi “chuông báo động” rung lên - *nếu không thì phải chơi một biến thể ngầm của cùng trò chơi*.

Mỗi người chơi phải được giả thiết là biết điều này và có thể, nếu anh ta muốn, đơn giản bằng cách tránh thỏa thuận công khai, chọn lựa để chơi trò chơi ngầm thay thế. Do vậy, nếu chúng ta giả thiết (vào lúc này) rằng trò chơi ngầm có một giải pháp được nhìn nhận rõ ràng, và rằng giải pháp đó có

---

<sup>6</sup> Có một giả thiết cơ học ở đây rằng trong quá trình đưa ra một đè nghị, một người có thể ngưng lại và khởi động lại. Trường hợp này thì phức tạp hơn một chút nếu một đè nghị được khởi đầu một phút rưỡi trước nửa đêm nhất thiết là đè nghị cuối cùng bởi vì quá trình không thể khởi động lại cho tới khi một phút đã trôi qua và vào lúc đó thời điểm quan trọng đã bị trôi qua. Trường hợp này sẽ được xem xét lại ở bên dưới.

*hiệu quả*, mỗi người chơi có một chiến lược hành vi thuần tối thiểu hóa tối đa trong giai đoạn trước đó. Một trong hai người có thể thực hiện giải pháp ngầm này bằng cách thoái lui khỏi thỏa thuận cho đến khi chuông báo động rung lên. Không ai trong hai người có thể giành được điều gì tốt hơn từ một đối thủ duy lý chỉ bằng việc mặc cả bằng lời.

Từ điểm này dẫn đến việc giải pháp của trò chơi hợp tác phải đồng nhất với giải pháp của trò chơi ngầm tương ứng (nếu trò chơi sau có một giải pháp có thể dự đoán được và có hiệu quả). Nó phải là như thế, bởi vì trò chơi ngầm xuất hiện như một phần tiếp theo cơ học không tránh khỏi của trò chơi hợp tác.

Tại điểm này có vẻ như đặc điểm hợp tác của trò chơi là không thích hợp. Các người chơi thật sự không cần xuất hiện cho đến 11 giờ 59 phút; thật ra họ hoàn toàn không cần phải xuất hiện. Việc truyền đạt thông tin trước khi chơi và khả năng đạt tới các thỏa thuận ràng buộc, có mục đích nhằm mô tả đặc tính trò chơi, có thể chứng minh là không thích hợp; không tồn tại trò chơi hợp tác như một trò chơi khác biệt với trò chơi ngầm.<sup>7</sup>

---

<sup>7</sup> Trong bài báo của mình vào năm 1953 mang tựa đề "Các trò chơi hợp tác hai người", Nash trình bày một mô hình mà mang tính chất ngầm một cách rõ ràng trong giai đoạn cuối cùng của nó. Mỗi quan hệ của mô hình này với trò chơi hợp tác mang tính chất gợi ý khám phá: nó nhằm giúp phát hiện ra điều gì có thể tạo ra "các kỳ vọng hợp lý" (và từ đó là kết quả hợp lý được chỉ định) trong trò chơi hợp tác tương ứng. Lý luận của bài báo này là mối quan hệ đó có xu hướng mang tính chất cơ học hơn là tính chất tri tuệ nếu một cấu trúc nước đi đối xứng được tuân theo một cách nghiêm ngặt, và rằng với tính đối xứng nghiêm ngặt thì rất khó nếu không nói là không thể xác định được trò chơi tương ứng không ngầm vốn là chủ đề cuối cùng của đề tài nghiên cứu.

Nhưng kết luận này không được đảm bảo. Trước hết, một trò chơi ngầm có thể không có một giải pháp hiệu quả được dự báo một cách lạc quan.<sup>8</sup> Hơn thế nữa, một số chi tiết nào đó của trò chơi hợp tác mà có thể đã có vẻ như vô thưởng vô phạt từ quan điểm thương lượng công khai có thể ảnh hưởng đến đặc tính của trò chơi ngầm; tương tự như thế, việc liên lạc trước khi chơi mà không có hiệu ứng ràng buộc lên chính các người chơi cũng có thể ảnh hưởng đến đặc tính của trò chơi ngầm. Ví dụ, hãy xem xét biến thể sau đây của trò chơi hợp tác.

---

<sup>8</sup> Cần phải nhấn mạnh rằng các giải pháp trò chơi mặc cả mà (giống như các giải pháp của Nash và Harsanyi) tùy thuộc vào một điểm zero được nhìn nhận rõ ràng - nghĩa là, tùy thuộc vào một kết quả không mơ hồ đóng vai trò chủ đạo khi thiếu vắng một thỏa thuận công khai - không thể nhất thiết được áp dụng cho một trò chơi hợp tác được dựa trên một ma trận các chọn lựa. Một ma trận (có lẽ trừ phi tất cả lợi ích là zero ngoại trừ trên đường chéo) không có một điểm zero được xác định bởi các quy tắc. Do vậy không có "hình thức bình thường" bao gồm một vùng lồi và điểm zero đi kèm trừ phi có sẵn một lý thuyết hoàn toàn tương xứng mà có thể "giải quyết" trò chơi ngầm (và thực hiện được như thế theo một cách mà các người chơi công nhận). Người ta có thể, theo Luce và Raiffa (chẳng hạn, trang 137) xem "các mức độ an toàn" (giá trị cực đại) của các người chơi như điểm zero. Nhưng điều này hoặc là mang tính vô doán tùy tiện, hoặc dựa trên giả thuyết mà, nếu để tự mình, những người chơi có thể thành công trong việc thực hiện không tốt hơn mức này trong trò chơi ngầm. Giả thuyết sau, đặc biệt là khi có những điểm hiệu quả thuần chiến lược (như trong trò chơi của Braithwaite, và như trong ma trận của Luce-Raiffa được thảo luận trong chú thích 18 bên dưới), là một giả thuyết yếu mà có thể bị bác bỏ về mặt thực nghiệm. Nó cho rằng những người chơi duy lý không có khả năng tương quan hóa các chiến lược mà không trao đổi với nhau, trong khi thực tế đây là điều mà họ thường có thể thực hiện ngay cả khi có những sở thích mâu thuẫn nhau. (Điểm này được bàn lại trong chú thích 18). Tính chất mơ hồ tiềm năng của điểm zero là vấn đề giữa Harvey Wagner và John Harsanyi trong tác phẩm của Wagner có tựa đề "Trả lời về vấn đề mặc cả" (*Rejoinder on the Bargaining Problem*), *Southern Economic Journal*, 24:480-482 (tháng 4-1958).

Thay vì nói rằng những người chơi có thể phân chia một bộ phần thưởng họ có thể đạt tới thỏa thuận về việc phân chia hoàn toàn, chúng ta hãy nói rằng những người chơi này có thể phân chia một bộ phần thưởng ở *phạm vi* mà họ đã đạt được thỏa thuận về một sự phân chia. Họ có thể phân chia một phần của các phần thưởng có được như họ đã thỏa thuận vào lúc mà chuông rung lên. Nếu, ví dụ như có một trăm vật phẩm riêng biệt và các người chơi khi chuông rung đã đạt thỏa thuận về việc chia tám mươi trong số đó, phần hai mươi vật phẩm trở lại chủ trò trong khi phần tám mươi đã đạt được thỏa thuận sẽ được chia theo thỏa thuận.<sup>9</sup>

Bây giờ, trong trường hợp mặc cả công khai (hợp tác), nếu chúng ta đã kết luận rằng có một giải pháp hiệu quả cho trò chơi này - nghĩa là, những người chơi trong thực tế sẽ đạt được một thỏa thuận trọn vẹn - hẳn là có lẽ chúng ta sẽ xem việc tái lập công thức cho bài toán là chuyện không lô-gic. Trong thực tế, việc tái lập công thức chỉ đề cập rằng việc mặc cả nên theo hình thức mỗi người chơi viết ra mức tổng thể yêu cầu của mình và rằng các hoạt động thỏa hiệp sẽ có dạng mỗi người chơi loại bỏ bớt các khoản trong danh mục yêu cầu của anh ta, và sẽ đạt được thỏa thuận hoàn toàn khi không còn khoản nào bị tranh chấp trên các danh mục yêu cầu. Nhưng,

---

<sup>9</sup> Trong trường hợp về một vật thể duy nhất có thể phân chia như tiền bạc, quy tắc tương ứng có thể là họ phân chia tiền theo các đề nghị của mình sau khi nhả cái đã loại trừ đi phần "trùng lắp." Mỗi người chơi nhận được ở mức mà người kia đã ngầm chấp thuận cho anh ta; nếu một người đòi hỏi 65% số tiền vào cuối trò chơi, còn người kia đòi 55%, thì người thứ hai đã được chấp thuận 35% và người đầu tiên được chấp thuận 45%; những con số này nằm bên ngoài phạm vi tranh chấp và tạo thành "thỏa thuận."

khi chúng ta xem xét trường hợp ngầm, trò chơi bị thay đổi mạnh mẽ bởi việc tái lập công thức này. Trò chơi ngầm lúc này có một cấu trúc động cơ tréo ngoe. Chẳng có lý do thích hợp nào để cho một trong hai người chơi đòi hỏi ít hơn toàn thể phần thưởng; mỗi người đều biết điều này và biết rằng người kia cũng biết điều này. Chẳng có động cơ nào khuyến khích giảm bớt yêu cầu của mỗi người bởi vì bất cứ tranh chấp nào về phần còn dư cũng không khiến một người chơi phải trả giá nhiều hơn mức mà anh ta sẽ bị mất nếu anh ta giảm bớt mức yêu cầu của mình để loại trừ tranh chấp. Điểm cân bằng duy nhất này mang lại con số zero cho cả hai người chơi. Như thế trò chơi biến thể này, mà có vẻ khác đi một cách vô lý, thì khác biệt mạnh mẽ so với trò chơi ban đầu; nhưng nó không có vẻ như thế cho đến khi chúng ta đã nhận dạng được trò chơi ngầm ở điểm đến cuối cùng là một ảnh hưởng vượt trội.<sup>10</sup>

Hãy lấy một ví dụ khác. Giả sử rằng có 100 vật phẩm để phân chia và rằng, mặc dù chúng có thể thay thế được nếu xét về giá trị, thỏa thuận phải đề cập chính xác các vật phẩm riêng nào sẽ thuộc về người chơi nào. Nếu các quy tắc đòi hỏi

<sup>10</sup> Có vẻ dường như chúng ta có thể rút ra được một sản phẩm phụ từ phân tích ở đây, nghĩa là, việc quan sát rằng nhằm thiết lập một trò chơi "thật sự" hợp tác (không ngầm), thì định nghĩa ở khía cạnh pháp lý của thỏa thuận phải làm sao cho trò chơi ngầm cuối cùng trở thành tréo ngoe, để cho các người chơi phải đạt được thỏa thuận ràng buộc trước khi chuông báo động vang lên, nếu không sẽ thua hết. Nhưng vẫn có một vấn đề. Các người chơi lúc này chính họ phải định nghĩa "thỏa thuận" cho các mục đích của thỏa thuận của chính họ trước khi có tiếng chuông cuối cùng. Nếu điều này giống như định nghĩa trước đó của chúng ta, tất cả những gì mà họ hoàn thành là làm cho trò chơi hợp tác tréo ngoe trở thành một trò chơi tốt đẹp, ngắn hơn một phút, tương đương với một trò chơi ngầm ngắn hơn hai phút so với trò chơi nguyên thủy.

rằng thỏa thuận đầy đủ và trọn vẹn phải được đạt tới, thì trong trò chơi ngầm các người chơi phải phụ thuộc vào khả năng của họ không chỉ để phân chia tổng giá trị các vật phẩm theo cách phối hợp mà còn để phân loại 100 vật phẩm riêng thành hai đồng theo cách thống nhất. Nếu, sau đó, một trong hai người chơi đã yêu cầu các vật phẩm *cụ thể* trị giá 80% tổng giá trị và người chơi kia đã từ chối, người ban đầu có một lợi thế trong trò chơi ngầm. Đề xuất hiện có duy nhất cho việc chia 100 vật phẩm là mức *cụ thể* 80 của một người chơi mà sẽ làm anh ta hài lòng. Các cơ may họ phối hợp đồng nhất trong bất cứ việc phân chia nào đối với 100 vật phẩm, bình đẳng hay bất bình đẳng giữa họ, có thể là quá nhỏ đến mức vì lý do thỏa thuận họ buộc phải chấp nhận đề xuất hiện có duy nhất cho dù nó là không công bằng. Như thế việc trao đổi trước khi chơi có một ý nghĩa chiến thuật ở chỗ nó có thể tác động tới phương tiện phối hợp một khi giai đoạn ngầm của trò chơi đã được đạt tới.

Nếu lúc này, trong khi xem xét các hàm ý chiến thuật của điểm cuối cùng này, chúng ta đòi hỏi quy tắc hành vi đối xứng, chúng ta phải kết luận rằng nếu một trong hai người chơi mở miệng để nói át đi điều mà người chơi kia chuẩn bị nói, anh ta sẽ luôn luôn thấy người chơi kia cũng đang mở miệng, cả hai cùng biết rằng nếu một trong hai người nói ra thì người kia cũng sẽ được nhìn thấy đang nói, không ai có thể nghe được, và cứ tiếp tục như thế. Nói cách khác, như một kết luận trước đã được thừa nhận giả thiết về tính đối xứng hoàn toàn của hành vi có vẻ như là bước đạo đầu của loại hành động mà dường như có thể làm phong phú cho trò chơi ở giai đoạn trao đổi trước trò chơi.

Nhưng đến lúc này chúng ta chắc chắn đã đầy trò chơi nước đi đổi xứng hoàn hảo xa đến mức nó đáng được như thế.<sup>11</sup> Chúng ta có thể tiếp tục phân tích trò chơi này ở mức

<sup>11</sup> Một chi tiết có thể đáng theo đuổi, phù hợp với một chủ thích trước đó. Giả sử rằng cần một phút để qua ra hay thay đổi một đề nghị và (ngược lại với mô hình trước đó) và rằng quy trình ghi nhận một đề nghị mới, một khi đã được bắt đầu, không thể bị chặn lại trước khi nó được hoàn tất. Theo thủ tục này, bắt cứ đề nghị nào được xướng lên trong phút kế tiếp của phút cuối cùng của trò chơi thì là đề nghị cuối cùng của một bên. Nếu đề nghị cuối cùng này *không thể* được truyền đạt cho người kia trước khi kết thúc phút đó, thì trò chơi chủ yếu vẫn như cũ; “đồng thời” ở đây có nghĩa là trong phạm vi một phút của lẫn nhau cho các mục đích thực tế, và một lần nữa không ai trong hai người có thể thấy được đề nghị cuối cùng của người kia trong khi anh ta đề xướng đề nghị của mình, bắt chấp việc các đề nghị được khởi xướng vào lúc nào trong phút cuối cùng. Nhưng giả sử mỗi người chơi gắn vào một tấm bảng thấy được mà được khóa lại trong một phút trong khi đề nghị được ghi nhận, sao cho người chơi kia có thể thấy được đề nghị của người này trong vòng một vài giây mặc dù người này không thể đề xướng một sự thay đổi cho đến khi việc trì hoãn của phút đó hết đi. (Và giả sử rằng không ai trong hai người có thể làm cho chính mình mất khả năng rõ ràng trong việc nhìn thấy đề nghị của người kia một khi nó được ghi nhận theo cách đó). Trong trường hợp này, nếu hai đề nghị trong phút cuối cùng đó không diễn ra đồng thời, người chơi nào đi sau sẽ thực hiện đề nghị cuối cùng của mình trong sự nhận biết hoàn toàn về đề nghị của người kia; và bởi vì cơ hội duy nhất của anh ta để giành bắt cứ thắng lợi nào là chấp nhận nó, anh ta phải chấp nhận bắt cứ điều gì mà người kia đề nghị. Như vậy, “đi sau” thua nếu người đi trước biết rằng người kia đang chờ đợi. Lúc này chúng ta có một trò chơi mà có thể được mô tả như sau: các người chơi cứ rẽ rẽ trong khoảng 23 giờ 58 phút và rồi chơi một trò chơi kéo dài một phút, trò chơi này cho phép mỗi người chơi có một và chỉ một đề nghị mà anh ta có thể đưa ra bắt cứ lúc nào trong phút đó. Trong thực tế, trò chơi này đưa ra ba chiến lược cho một người chơi, bao gồm (1) giả thiết rằng người kia sẽ chờ đợi, và đòi hỏi 99%; (2) giả thiết rằng cả hai sẽ đưa ra đề nghị đồng thời, và đòi hỏi bắt cứ mức độ nào được gợi ý bởi trò chơi ngầm; (3) chờ đợi. Nếu cả hai chờ đợi, trò chơi vẫn còn được tiếp tục chơi. Nếu có một con số có hạn nhưng lần chờ đợi tiệm ngắn, chúng ta có các chiến lược chờ đợi-một-lần-rồi-de-nghi-99%, chờ đợi-một-lần-de-nghi-giải-pháp-ngầm, chờ đợi-hai-lần-rồi-de-nghi-99%, chờ đợi-hai-lần-

chi tiết hơn, xem xét các sự việc như các giải pháp thay thế trong việc chấm dứt trò chơi hay định nghĩa "thỏa thuận", và tiếp tục như thế. Tuy nhiên, có vẻ như là sẽ đáng giá hơn khi nêu ra tại điểm này vấn đề liệu trò chơi "bất động" hoàn hảo hay trò chơi "nước đi đổi xứng" có lợi hơn để nghiên cứu. Liệu có phải trò chơi không kỳ thí, nước đi đổi xứng là một trò chơi "tổng quát", một trò chơi mà tách khỏi "các trường hợp đặc biệt" hay không? Hay nó là một trường hợp đặc biệt và giới hạn mà trong đó các khía cạnh thú vị nhất của trò chơi hợp tác đã biến mất?

Cần nhấn mạnh rằng cái thay thế có kết quả cho tính đối xứng không phải là giả thiết về tính bất đối xứng, mà chỉ là tính *phi đối xứng*, thừa nhận cả tính đối xứng và tính bất đối xứng như những khả năng xảy ra mà không cam kết lựa chọn cái nào như một kết luận từ trước.

Một minh họa có thể hữu ích. Giả sử chúng ta cần phải phân tích trò chơi trong đó có 100 đô-la ở cuối con đường dành cho người chơi nào đến được đó trước. Loại trò chơi kỹ năng này không khó để phân tích: khoản tiền dành cho người đi nhanh nhất, không kể các tai nạn và các yếu tố ngẫu nhiên. Chúng ta có thể dự đoán hành vi thích hợp (chạy) và kết quả (tiền cho người nhanh nhất). Các lần hòa sẽ thỉnh thoảng xảy ra; nhưng chúng sẽ xảy ra ở cuối chặng đua và không được

---

đề-nghị-giải-pháp-ngầm, và tiếp tục như thế. Trò chơi này ("siêu trò chơi ngầm" bao gồm tất cả các chiến lược để chơi trò chơi một phút) chính là trò chơi *đó*; và nó có, nếu chúng ta chấp nhận nó, "đáp số theo nghĩa chặt" của nó bao gồm tất cả các chiến lược (tất cả các khoảng thời gian của những lần chờ đợi) mà kết thúc bằng những yêu cầu tương ứng với kết quả của trò chơi ngầm. (Đối với định nghĩa về đáp số theo nghĩa chặt trong một trò chơi ngầm giữa hai người, xin xem Phụ lục C).

nhìn nhận từ lúc bắt đầu khởi hành. Chúng ta cần có một quy tắc phụ để giải quyết các trường hợp hòa, nhưng nó không cần thiết phải thống lĩnh hoặc là trò chơi, hoặc là sự phân tích.

Hãy xem xét cùng một trò chơi được chơi trong một đám đông dân chúng trong đó mỗi người đều có thể chạy nhanh một cách chính xác như bất cứ người nào khác, và mọi người đều biết điều đó. Bây giờ chuyện gì xảy ra? Mỗi một cuộc đua đều kết thúc bằng trường hợp hòa, do vậy quy tắc phụ đóng vai trò quyết định tất cả. Nhưng bởi vì trường hợp hòa là một kết luận từ trước, thế thì tại sao họ phải chạy cho phiền toái?

Trò chơi hợp tác nước đi đối xứng hoàn hảo có vẻ giống cuộc đi bộ đua đó một chút. Mặc cả trong một trường hợp này thì cũng vô ích như chuyện động chân trong trường hợp kia. Mỗi một người chơi đều biết từ trước rằng tất cả mọi nước đi và mọi chiến thuật đều được xem như bị vô hiệu hóa bởi các tiềm năng đối xứng dành sẵn cho đối thủ của mình. Các yếu tố thú vị mà chúng ta có thể đưa vào trong trò chơi mặc cả là vô nghĩa nếu tính đối xứng hoàn hảo, và việc cả hai người chơi chấp nhận nó như là không thể tránh khỏi, được áp đặt lên trò chơi bởi chính định nghĩa của nó.

Vậy thì chúng ta nên thêm gì vào trò chơi để làm nó phong phú nếu giả thiết về tính đối xứng bị loại bỏ? Có nhiều "nước đi" thường có sẵn, nhưng không nhất thiết là có sẵn một cách bình đẳng cho cả hai người chơi trong tình huống trò chơi thật sự. "Các nước đi" sẽ bao gồm cam kết, đe dọa, hứa hẹn; phá rối hệ thống liên lạc; viện dẫn sự trừng phạt kèm theo các lời hứa, cam kết, và đe dọa; sự trừng phạt chuyển tải thông tin xác thực, sự tự định dạng; và việc đưa vào chi tiết bối cảnh

mà có thể kèm chế các kỳ vọng, đặc biệt là khi việc liên lạc không hoàn tất. Các “nước đi” như thế được thảo luận chi tiết trong các Chương 2-5.

Để minh họa, giả sử trong trò chơi hợp tác trước đó có một cửa xoay cho phép một người chơi rời chỗ nhưng không cho phép quay trở lại; đề nghị hiện tại của anh ta trong khi anh ta đi ngang qua cửa xoay vẫn nằm ở trên các tập sách cho đến khi chuông rung lên. Lúc này chúng ta có một phương tiện mà từ đó một người chơi có thể đưa ra một đề nghị “cuối cùng”, một “cam kết”. Bất cứ ai có thể lưu lại một đề nghị thuận lợi cho chính anh ta và được người kia biết tới, và rời cẩn phòng, thì có chiến thuật thắng cuộc. Dĩ nhiên bất cứ ai trong họ cũng đều có thể chiến thắng; nhưng điều này có thể có nghĩa là chúng ta kết thúc với một sự việc giống như cuộc đua, và người ở gần cửa xoay nhất thắng cuộc. Bằng cách phân tích chiến thuật, và cách sắp xếp thể chế hay vật chất, chúng ta có thể xác định ai có thể sử dụng cửa xoay trước tiên.

Cần phải lưu ý rằng chúng ta vẫn chưa chuyển trò chơi chiến lược thành một trò chơi kỹ năng bằng cách cho họ chạy đua đến cửa xoay. Vẫn đúng là một người chỉ có thể thắng cuộc khi anh ta tiến đến cửa xoay trước qua việc hợp tác của người kia, qua việc hạn chế việc lựa chọn chiến lược của người kia. Anh ta không thắng cuộc về mặt pháp lý hay thể lý bằng cách đi qua cửa xoay; anh ta thắng cuộc về mặt chiến lược. Anh ta làm cho người kia lựa chọn sao cho có lợi cho mình. Đó là một chiến thuật trong một trò chơi chiến lược, dù là việc sử dụng nó có thể tùy thuộc vào kỹ năng hay lợi thế về vị trí.

Chúng ta thậm chí có thể đưa một thể loại nào đó của tinh

đối xứng vào trong trò chơi lúc này, mà không phải hủy bỏ nó. Chúng ta có thể tung đồng xu để xem ai ở gần cửa xoay nhất khi trò chơi bắt đầu, hoặc để cho người chơi có vị trí tương tự và có vận tốc tương tự nhưng với các yếu tố ngẫu nhiên để xác định ai đến cửa xoay trước nhất. Mặc dù trò chơi này lúc này là *không phân biệt*, thì *kết quả* cũng vẫn sẽ là *bắt đối xứng* bởi vì mỗi người chơi có một động cơ để chạy đến cửa xoay, để lại phía sau một đề nghị hiện hành mà anh ta ưa thích.<sup>12</sup>

Chúng ta có thể bao gồm rủi ro "hòa" nào đó, đặc biệt là khi có hai cửa xoay và các người chơi có thể đi qua hai cửa đó đồng thời. Điều này tạo ra "tính đối xứng" như một khả năng thú vị, nhưng không phải như một kết luận từ trước. Thế bế tắc cũng như dự báo về thế bế tắc trở thành các khả năng thú vị nếu các hành động và cấu trúc thông tin thực ra mang tính chất dẫn dắt đến các trường hợp hòa. Nhưng, với tính chất phi đối xứng như triết lý của chúng ta, chúng ta không cần phải bận tâm đến khả năng hòa.

Lại nữa, nếu một người chơi có thể đưa ra một đề nghị và phá hủy việc liên lạc, anh ta do vậy có thể thắng cuộc trong trò chơi ngầm diễn ra tiếp theo bằng cách đã cung cấp đề nghị hiện có duy nhất mà cả hai người chơi có thể tập hợp vào lúc họ hết sức cần phối hợp các sự lựa chọn của mình về sau trong giai đoạn ngầm cuối cùng. Để chắc chắn, chúng ta có thể xem xét

<sup>12</sup> Có thể tranh luận ở điểm này rằng giá trị kỳ vọng của trò chơi vẫn có thể được phân chia đối xứng giữa hai người chơi, và rằng nhà phân tích do đó vẫn có thể xem trò chơi như là đối xứng xét theo các kết quả bình quân. Nhưng nếu làm như thế, nhà phân tích hạn chế mình ở mức hiểu biết tối thiểu về trò chơi và cách mà trò chơi được chơi.

điều gì xảy ra khi các năng lực phá hủy môi dây liên lạc đang hiện có là như nhau, và các người chơi phải nhìn nhận rằng họ có thể đồng thời phá hủy liên lạc mà không phải nhận thông điệp qua lại; nhưng trường hợp thú vị này có vẻ như là một trường hợp đặc biệt, không phải là trường hợp chung.

Nói tóm lại, trò chơi hợp tác “tĩnh” (không có nước đi) hoàn hảo hay “nước đi-đối xứng” không phải là một trường hợp tổng quát có kết quả, mà là một trường hợp hạn chế có thể thoái hóa thành một trò chơi ngầm thông thường. Trò chơi hợp tác thì phong phú và có ý nghĩa khi “các nước đi” được thừa nhận; và phần lớn ý nghĩa của các nước đi sẽ biến mất nếu tính chất đối xứng hoàn toàn trong tính sẵn có của các nước đi đó đối với hai người chơi được khắc ghi vào ngay trong định nghĩa của trò chơi. Chính các nước đi mới thú vị, chứ không phải trò chơi không có các nước đi; và chính tính bất đối xứng tiềm năng của các nước đi làm cho chúng trở nên thú vị nhất.

Tính đối xứng không chỉ thường được áp vào cấu trúc nước đi của các trò chơi, mà còn được viện dẫn như một đặc tính đáng tin cậy trong kết quả của trò chơi hay trong hành vi hợp lý mà kết quả phải phù hợp với. Lý thuyết của Nash về trò chơi hợp tác hai người đưa ra một cách rõ ràng định đe về tính đối xứng, và tương tự là lý thuyết của Harsanyi. Định đe về tính đối xứng chắc chắn là thích hợp, nó thường cho phép một người chơi tìm ra một “đáp số” cho một trò chơi và để - nếu anh ta muốn - trong phạm vi của toán học. Có ít khái niệm có tiềm năng một cách tương tự mà có thể cạnh tranh với nó để làm các cơ sở cho việc giải quyết một trò chơi.

Nhưng lý lẽ chứng minh cho định đê về tính đối xứng từ trước đến nay không chỉ là việc nó dẫn đến những kết quả dễ chịu; nó đã được chứng minh trên nền tảng rằng đối nghịch của tính đối xứng sẽ có xu hướng đối nghịch với tính hợp lý của hai người chơi. Đây chính là nền móng mà tôi muốn tấn công.

Điều tôi chuẩn bị tranh luận là, mặc dù tính đối xứng phù hợp với tính hợp lý của các người chơi, người ta cũng không thể thể hiện rằng tính bất đối xứng là không phù hợp với tính hợp lý của người chơi, trong khi việc đưa tính đối xứng vào *định nghĩa* của tính hợp lý lại làm bật ra câu hỏi. Do vậy tôi mong muốn đưa ra điều mà tôi nghĩ là *một* lập luận ủng hộ các kết quả đối xứng, một lập luận có xu hướng làm cho tính chất đối xứng chỉ là một trong nhiều ảnh hưởng tiềm năng lên kết quả mà không có đòi hỏi phiến diện đối với thành tích xuất sắc nào.

Các tuyên bố rõ ràng về mối quan hệ giữa tính đối xứng và tính hợp lý đã được đưa ra bởi Harsanyi. Ông ta nói, "Vấn đề mặc cả có một giải pháp xác định hiển nhiên trong ít nhất một trường hợp đặc biệt: nghĩa là, trong các tình huống mà hoàn toàn có tính chất đối xứng liên quan đến hai bên mặc cả. Trong trường hợp này, hoàn toàn tự nhiên khi giả thiết rằng hai bên sẽ có xu hướng chia sẻ tổng lợi ích một cách đều nhau bởi vì không bên nào sẵn sàng để cho bên kia các điều kiện tốt hơn là mức mà bên kia sẽ để cho anh ta."<sup>13</sup> Trong một đê tài sau đó ông ta đê cập đến tiên đê hiển nhiên về tính đối

<sup>13</sup> Harsanyi, 147. Ông ta nói tiếp rằng, "Ví dụ như, mỗi người sẽ kỳ vọng rằng hai tập đoàn độc quyền duy nhất với cùng các chức năng về chi phí, quy mô, điều kiện thị trường, nguồn lực vốn, nhân lực, v.v... sẽ đạt thỏa thuận để trao cho nhau mức lợi nhuận bình đẳng."

xứng như là “định đê cơ bản” và nói rằng “Một cách trực giác giả thiết khẳng định cho tiên đê này là một người mặc cả duy lý sẽ không kỳ vọng một đối thủ duy lý trao cho anh ta các nhượng bộ lớn hơn mức mà chính anh ta sẽ thực hiện dưới cùng các điều kiện.”<sup>14</sup>

Lúc này, công thức trực giác này bao hàm hai định đê. Trước tiên, đó là một người mặc cả sẽ không nhượng bộ nhiều hơn mức mà anh ta sẽ kỳ vọng có được nếu chính anh ta ở trong vị trí của người kia. Thứ hai, cơ sở duy nhất cho sự kỳ vọng của anh ta về điều mà anh ta sẽ nhượng bộ nếu anh ta ở trong vị trí của người kia chính là nhận thức của anh ta về tính chất đối xứng.

Công thức trực giác, hay thậm chí là một công thức cẩn thận theo các tiêu chí tâm lý học, về điều mà một người chơi duy lý “kỳ vọng” trong tương quan với một người chơi duy lý khác, đặt ra một vấn đề mô tả thuần về khoa học. Cả hai người chơi, với tính duy lý, phải nhìn nhận rằng loại kỳ vọng

---

<sup>14</sup> Việc trích toàn văn đáng được đưa ra: “Điều mà lý thuyết Zeuthen-Nash về mặc cả chủ yếu đề nghị thực hiện là chỉ ra đâu là những kỳ vọng mà cả hai người mặc cả duy lý có thể ấp ú một cách nhất quán trong tương quan với các chiến lược của nhau nếu họ biết rõ về các chức năng sử dụng của nhau. Định đê cơ bản của lý thuyết này là một tiên đê về tính chất đối xứng, chỉ ra rằng các hàm xác định các chiến lược tối ưu của hai bên xét theo các dữ liệu (hay, một cách tương đương, các hàm xác định các phân thường cuối cùng của hai bên) có cùng một hình thức toán học, dĩ nhiên loại trừ yêu tố rằng các biến thể gắn liền với việc hai bên phải trao đổi qua lại. Một cách trực giác, giả thiết nhấn mạnh cho tiên đê này là một người mặc cả duy lý sẽ không kỳ vọng một đối thủ duy lý dành cho anh ta những nhượng bộ lớn hơn mức mà anh ta sẽ nhượng bộ dưới các điều kiện tương tự.” (Harsanyi, “Mặc cả trong việc bỏ qua các chức năng sử dụng của đối thủ” (Bargaining in Ignorance of the Opponent’s Utility Function), Cowles Foundation Discussion Paper số 46, ngày 11-12-1957, được trích dẫn với sự cho phép của tác giả).

"hợp lý" duy nhất họ có thể có được là một kỳ vọng được chia sẻ hoàn toàn về một *kết quả*. Nó có lẽ không hoàn toàn chính xác - như một mô tả về hiện tượng tâm lý - khi nói rằng một người kỳ vọng người kia nhượng bộ một điều gì đó hay chấp nhận một điều gì đó; sự sẵn sàng của người thứ hai trong việc nhượng bộ hay chấp nhận chỉ là một cách biểu lộ về điều mà anh ta muốn người thứ nhất nhượng bộ hay chấp nhận, mà đến lượt nó lại là điều anh ta muốn người thứ nhất kỳ vọng người thứ hai kỳ vọng người thứ nhất kỳ vọng, và tiếp tục như thế. Để tránh một "sự thêm vào vô hạn" trong quy trình được mô tả, chúng ta phải nói rằng cả hai cảm thấy một kỳ vọng được chia sẻ về một *kết quả*, "sự kỳ vọng" của một người là một niềm tin rằng cả hai cùng xác định *cùng* một kết quả như được gợi ý trong tình huống, do vậy hầu như là không thể tránh khỏi. Cả hai người chơi, trong thực tế, chấp nhận một thẩm quyền chung - sức mạnh của trò chơi để đưa ra kết quả riêng thông qua khả năng tri thức của người chơi trong việc nhận thức nó - và điều họ kỳ vọng là cả hai cùng nhận thức về một kết quả chung.<sup>15</sup>

---

<sup>15</sup> Xem xét theo cách này, quy trình tri thức để đạt tới "các kỳ vọng hợp lý" trong trò chơi mặc cả có trao đổi thông tin đầy đủ thì hầu như đồng nhất với quy trình tri thức để đạt tới một chọn lựa phối hợp trong trò chơi ngầm. Các kết quả thực sự có thể khác nhau bởi vì các bối cảnh trò chơi có thể khác nhau, với các chi tiết gợi ý khác nhau; nhưng bản chất của hai kết quả dường như hầu như đồng nhất bởi vì cả hai thuộc vào một thỏa thuận được đạt tới bởi sự đồng ý *ngầm*. Điều này là đúng bởi vì thỏa thuận rõ ràng được đạt đến trong trò chơi có liên lạc đầy đủ thì tương ứng với các kỳ vọng *tiền nghiệm* mà được đạt tới (hay trong lý thuyết có thể đã được đạt tới) cùng nhau nhưng độc lập bởi hai người chơi trước khi sự mặc cả bắt đầu. Và nó giống như một thỏa thuận ngầm theo nghĩa rằng cả hai có thể có các kỳ vọng hợp lý chắc chắn chỉ khi nào cả hai ý thức rằng cả hai chấp nhận kết quả được chỉ định từ trước là *cái kết quả* mà cả hai cùng biết rằng cả hai cùng kỳ vọng.

Trong các giới hạn này, phần đầu tiên (rõ ràng, trong giả thuyết Harsanyi có thể được nói lại: rằng có, trong bất cứ tình huống trò chơi mặc cả nào (với thông tin hoàn hảo về các lợi ích), một kết quả cụ thể mà một trong hai người chơi duy lý có thể nhìn nhận rằng bất cứ người chơi duy lý nào cũng sẽ nhìn nhận kết quả đó là "đáp số" được chỉ định. Phần thứ hai (ngụ ý) của giả thuyết này là kết quả cụ thể được nhìn nhận như thế được quyết định bởi tính chất đối xứng toán học. Cái đầu tiên chúng ta có thể gọi là định đề "lời giải-hợp lý"; chính cái thứ hai mới tạo ra định đề "tính đối xứng."

Câu hỏi lúc này là liệu định đề về tính đối xứng có được xuất phát từ tính hợp lý của các người chơi - tính hợp lý nơi các kỳ vọng của họ - hay phải dựa vào các cơ sở nền tảng khác. Nếu nó dựa vào các cơ sở khác, chúng là gì và sự hỗ trợ của chúng mạnh mẽ như thế nào?

Để theo đuổi câu hỏi đầu tiên, liệu tính đối xứng có thể được rút ra từ tính hợp lý nơi các kỳ vọng của những người

---

Có một sự lượng giá ở điểm này. Với thông tin đầy đủ về các hệ giá trị của lẫn nhau và một bộ đồng nhất các lợi ích được phân chia, có thể có vô số các đáp số tương đương, tất cả đều mang đến cùng một giá trị cho hai người chơi, nhưng không có khó khăn nào trong việc thống nhất một sự lựa chọn trọng tài trong bộ không khác biệt này. Nhưng mặc cả ngầm thường đòi hỏi một cấp độ phối hợp cao hơn, nghĩa là một sự lựa chọn có phối hợp ngay cả trong số các cách phân chia tương đương phần thu được. Vì vậy việc thương lượng về một đường ranh giới trong một vùng lãnh thổ đồng nhất khác với việc cất cử đồng thời quân đội đến tiến chiếm các vị trí thể hiện các tuyên bố (như trong Câu hỏi 6 ở chương 3); các tuyên bố như thế có thể trùng lắp nhau và gây ra phiền toái thậm chí dù các giá trị đất đai là nhất quán. Như vậy vẫn đề phối hợp khác đi; và không có một sự bảo đảm tiền nghiệm rằng kết quả cho trò chơi ngầm (hay cho các trò chơi với việc liên lạc không hoàn hảo ở mức độ nào đó, và vân vân) sẽ nằm trong bộ các kết quả tương đương cho một trò chơi hoàn toàn công khai.

chơi, chúng ta có thể xem xét chung tính hợp lý của hai người chơi và thắc mắc liệu một kết quả phi đối xứng cùng được kỳ vọng có đối nghịch với định đè về tính đối xứng không. Nếu hai người chơi tin tưởng một cách chắc chắn là họ chia sẻ, và có chia sẻ, kỳ vọng về một kết quả cụ thể nào đó, và kết quả đó là không đối xứng theo nghĩa toán học, liệu chúng ta có thể chứng minh rằng những kỳ vọng của họ là không hợp lý, và rằng định đè về tính đối xứng là mâu thuẫn hay không? Một cách cụ thể, giả sử hai người chơi có thể có 100 đô-la để chia nhau ngay khi họ đồng ý một cách rõ ràng về việc chia như thế nào; và họ hoàn toàn sẵn sàng đồng ý rằng A sẽ có được 80 đô-la và B sẽ có 20 đô-la; và chúng ta biết rằng các số tiền trong trường hợp cụ thể này là tương ứng về tỷ lệ với các lợi ích, và những người chơi cũng biết. Liệu chúng ta có thể chứng minh rằng những người chơi là không hợp lý chăng?

Chúng ta phải cẩn thận không làm cho tính đối xứng trở thành một phần của *định nghĩa* về tính hợp lý; làm như thế sẽ phá hủy tính tương thích thực nghiệm của lý thuyết này và chỉ làm cho tính đối xứng trở thành một tiên đè độc lập. Chúng ta phải có một định nghĩa đáng tin về tính hợp lý mà không đe cập đến tính đối xứng và chỉ ra rằng tính bất đối xứng trong các kỳ vọng mặc cả sẽ là không nhất quán với định nghĩa đó. Vì mục đích hiện tại của chúng ta, chúng ta phải giả thiết rằng hai người chơi đã chọn 80 đô-la và 20 đô-la theo thỏa thuận và xem liệu chúng ta có thể xác định mọi loại sai lầm tri thức, các kỳ vọng sai lầm, hay về tính tư lợi rối loạn, nơi một hay cả hai người chơi, trong việc không chọn lựa được một điểm đối xứng của họ.

Một cách cụ thể, đâu là “sai lầm” của B trong việc nhượng bộ 80 đô-la cho A? Anh ta đã kỳ vọng – anh ta có thể nói với chúng ta, và giả sử rằng chúng ta có phương tiện để kiểm tra tính trung thực của anh ta (một sự giả định khiêm tốn nếu thông tin đầy đủ về các lợi ích đã được giả thiết rồi!) – rằng A sẽ “yêu cầu” 80 đô-la; anh ta đã kỳ vọng A kỳ vọng có được 80 đô-la; anh ta đã biết rằng A biết rằng anh ta biết điều này, và tiếp tục như thế. A kỳ vọng có được 80 đô-la, biết rằng B sẵn sàng về mặt tâm lý, bởi vì anh ta, B, biết rằng A kỳ vọng một cách chắc chắn rằng B sẵn sàng, và tiếp tục như thế. Nghĩa là, cả hai cùng biết – họ nói với chúng ta – và cả hai biết rằng cả hai cùng biết, rằng kết quả không thể tránh khỏi sẽ là 80 đô-la cho A và 20 đô-la cho B. Cả hai đều chính xác trong mỗi kỳ vọng. Các kỳ vọng của mỗi bên nhất quán nội tại và nhất quán với các kỳ vọng của bên kia.

Chúng ta có thể bối rối về việc *làm thế nào* họ có thể đạt tới các kỳ vọng như thế; nhưng sự kỳ lạ cũng lôi cuốn sự khâm phục chẳng kém gì sự khinh thường. Định đè “lời giải-hợp lý” được xác nhận một cách đẹp đẽ; trò chơi dường như đã chỉ thị một kết quả cụ thể mà cả hai người chơi đều nhận thức rõ. Nếu, ở điểm này, chúng ta cảm thấy rằng chính chúng ta đã không cảm nhận được cùng một kết quả, chúng ta có thể kết luận rằng một trong bốn giả thuyết là sai: (1) định đè lời giải-hợp lý, (2) tính hợp lý của A và B, (3) tính hợp lý của chính chúng ta, và (4) định dạng (trong tất cả các khía cạnh quan trọng) của trò chơi mà chúng ta chơi một cách nội quan với trò chơi mà A và B vừa chơi. Nhưng chúng ta không thể, trên cơ sở chứng cứ, tuyên bố rằng cái thứ hai là sai – tức là tính hợp lý của A và B.

Lưu ý rằng nếu B đã cố đòi 50 đô-la, hoặc nếu A đã đồng ý yêu cầu 50 đô-la, tuyên bố là hợp lý và tranh luận trên cơ sở tự tin vào một kỳ vọng được chia sẻ về kết quả đó, cả hai người chơi hẳn sẽ là “sai làm” và chúng ta không thể phân biệt, trên cơ sở chứng cứ, ai là không hợp lý hoặc cả hai là không hợp lý. Trừ phi chúng ta làm cho tính chất đối xứng thành định nghĩa của tính hợp lý, chúng ta chỉ có thể kết luận rằng ít nhất một trong hai người chơi là không hợp lý, hoặc rằng định đe lời giải-hợp lý không đúng vững. Điều chúng ta có được tốt nhất là một điều kiện cần duy nhất cho tính không thích hợp chung của cả hai người chơi; chúng ta không có điều kiện đủ, và không có được điều kiện cần cho một người chơi duy nhất.

Chúng ta cũng không thể lột trần họ ra nếu chúng ta hỏi họ làm thế nào họ đạt tới các kỳ vọng của mình. Bất cứ cơ sở nào mà nhất quán cũng sẽ thích ứng, bởi vì bất cứ cơ sở nào mà mỗi người tự tin kỳ vọng người kia chấp nhận cũng sẽ là cơ sở mà anh ta không thể bỏ qua. Các câu chuyện nhất quán là tất cả những gì mà họ cần; và nếu họ nói rằng một dấu hiệu trên tấm bảng nói rằng A 80 đô-la, B 20 đô-la, hoặc nói rằng họ thấy trong một bản tin rằng hai người chơi khác là A' và B' chia nhau 80 đô-la-20 đô-la, và rằng họ nhận thấy rõ ràng đây là biểu hiện rõ ràng cho cả hai người về điều cần kỳ vọng - rằng đây là kết quả “có thể kỳ vọng” duy nhất - chúng ta không thể nói rằng họ sai và chứng minh là họ không hợp lý. Họ có thể là không hợp lý; nhưng chứng cứ không chỉ ra điều đó.

Tuy nhiên, có một cơ sở để phủ nhận lý luận hiện tại của tôi. Bởi vì tôi chưa thực sự áp dụng một bài kiểm tra độc lập

về tính hợp lý cho hai người chơi, trao cho họ trò chơi để chơi, và quan sát việc phân chia 80:20 mà tôi vừa đề cập, mà mới chỉ đưa nó ra như một khả năng xảy ra để xem liệu nó có mang ngụ ý về tính hợp lý *nếu* nó xảy ra, thì người ta có thể phản ứng rằng nó có thể không xảy ra. Và lập luận sẽ dựa vào vấn đề phối hợp; nó sẽ diễn tiến như sau.

Nếu hai người chơi cùng kỳ vọng *một cách tiền nghiệm* cùng một kết quả, và nhìn nhận chắc chắn nó là kỳ vọng *chung* của họ, họ phải có khả năng trí tuệ để lựa chọn một điểm cụ thể chung. Nếu tổng số 100 đôla có thể được phân chia cho đến từng đồng xu nhỏ nhất, thì có đến 9.999 cách phân chia thích hợp để xem xét, một trong số đó sẽ được lựa chọn ra một cách đồng thời nhưng tách biệt bởi cả hai người chơi là các kỳ vọng của họ về kết quả. Nhưng làm thế nào hai người có thể phối hợp sự lựa chọn của họ để lấy ra một trong số 9.999, theo nghĩa rằng sự kỳ vọng của họ tập trung vào hay hội tụ vào sự lựa chọn đó, bất chấp tỷ lệ 9.999 trên 1 chống lại họ? Câu trả lời hẳn phải là họ sử dụng một thủ thuật hay manh mối nào đó, hay một thiết bị phối hợp tự động ra cho họ. Họ phải, một cách ý thức hay không ý thức, sử dụng một thủ tục chọn lựa mà dẫn đến các kết quả duy nhất. Phải có gì đó giúp phân biệt điểm mà họ chọn - nếu không phải trong lập luận ý thức của họ, thì ít nhất trong sự phân tích ý thức của chúng ta - từ một thể liên tục các kết quả lựa chọn có thể.

Bây giờ, liệu có thể có được rằng hai người chơi duy lý, thông qua bất cứ điều gì khác với sự trùng hợp ngẫu nhiên đơn thuần hay ma thuật, tập trung sự chú ý của họ vào cùng một kết quả cụ thể và mỗi người có thể chắc chắn "một cách

hợp lý" rằng người kia tập trung vào cùng một kết quả với cùng một sự đánh giá rằng nó được kỳ vọng hỗ tương không? Và nếu như thế, thì họ có thể làm như thế bằng cách nào?

Câu trả lời là họ có thể, như đã thể hiện trong Chương 3. Họ có thể dùng bất cứ phương tiện nào có sẵn: bất cứ manh mồi nào, gợi ý nào, bất cứ quy tắc loại suy nào mà dẫn đến một sự lựa chọn không mơ hồ hay một khả năng cao về sự lựa chọn có phối hợp. Và một trong số các quy tắc này, hay manh mồi này, hay gợi ý này, là tính đối xứng về toán học.<sup>16</sup>

Trong một trò chơi mà tuyệt đối không có các chi tiết nào khác ngoài cấu trúc toán học của nó, trong đó không có vấn đề bối cảnh vô ý nào có thể khiến cho nó được đánh giá bởi một người chơi như là một sự việc cũng có thể được người chơi kia đánh giá, thì có thể không có gì để xử lý ngoại trừ một thể liên tục các con số. Và tất cả các con số có thể được phân loại tùy thuộc vào liệu chúng có tương ứng với các cách phân chia

---

<sup>16</sup> Cơ sở tri thức cơ bản, hay giả thuyết hoạt động, cho các người chơi hợp lý trong trò chơi này dường như là cơ sở cho rằng quy tắc nào đó phải được sử dụng nếu sự thành công cần phải lớn hơn sự trùng hợp ngẫu nhiên, và rằng quy tắc tốt nhất sẽ được tìm thấy, bất chấp tính hợp lý của nó là gì, thì do vậy sẽ là một quy tắc về tính hợp lý. Cơ sở này sẽ hỗ trợ cho chàng hạn như mô hình của Nash xem một trò chơi ngầm "không trọn tru" là giới hạn của một trò chơi "trọn tru" khi mà sự trọn tru tiến đến zero. Trong khi quan điểm này về trò chơi không trọn tru là không cần thiết về mặt lôgic ở bất cứ nghĩa nào, nó lại là một quan điểm có tính gợi ý mạnh mẽ mà, khi thiếu vắng bất cứ lý do căn bản tốt hơn nào cho việc tập hợp vào một điểm duy nhất, có thể tập trung sự chú ý của những người chơi có nhu cầu về một sự lựa chọn chung. Quy trình giới hạn cung cấp manh mối cho việc chọn ra một trong rất nhiều đến vô hạn điểm cân bằng mà thực sự tồn tại trong trò chơi không trọn tru. Dĩ nhiên, cơ sở này hỗ trợ một cách bình đẳng bất cứ thủ tục nào khác mà tạo ra một ứng viên cho việc lựa chọn từ trong số rất nhiều đến vô hạn các lựa chọn tiềm năng.

đối xứng hay bất đối xứng không. Nếu tất cả các con số ngoại trừ một con số thể hiện một sự phân chia bất đối xứng, thì tính đối xứng toán học không thôi cũng là một quy tắc đủ và là một quy tắc cực kỳ hữu ích trong việc phối hợp cho một sự chọn lựa chung. Và có thể thiết lập một trò chơi trong một cách thuần khiết như vậy, xoá bỏ lai lịch của các người chơi cũng như tất cả các chi tiết bối cảnh, mà hầu như không có cơ sở rõ ràng nào cho việc phối hợp trừ phi các yếu tố không thuần khiết len vào.<sup>17</sup>

Nói cách khác, tính đối xứng toán học có thể tập trung các sự kỳ vọng của hai người chơi có lý trí bởi vì nó có - với các đặc điểm giả thiết khác của trò chơi, chẳng hạn như thông tin đầy đủ về các hệ thống sử dụng của lẫn nhau - cung cấp một phương tiện cho việc phối hợp các kỳ vọng. Việc nó có phải

---

<sup>17</sup> Trong quan điểm này, lý thuyết của Nash (dẫn đến lời giải sản phẩm-degree thỏa dụng-tối đa) là một câu trả lời cho sự kiện rằng ngay cả trong lĩnh vực toán học vẫn có sẵn quá nhiều loại độc nhất hay đối xứng để cung cấp một quy tắc mơ hồ cho việc tuyển chọn, dẫn đến nhu cầu viện dẫn các tiêu chí đáng tin cậy (các tiên đề) đủ để làm sản sinh ra một sự tuyển chọn mơ hồ. Lý thuyết của Braithwaite có thể được mô tả theo cùng cách thức. Việc hai đáp số mâu thuẫn nhau ngữ ý rằng các nhà toán học có thể không có một khiếu thẩm mỹ toán học chung ở mức đủ để làm thỏa mãn phần đầu trong định đè của Harsanyi, nghĩa là, để phối hợp các kỳ vọng của họ cho cùng một kết quả. (R. B. Braithwaite, *Theory of Games as a Tool for the Moral Philosopher* (Lý thuyết trò chơi như một công cụ cho nhà triết học luân lý) [Cambridge, England, 1955]; Giải pháp của Braithwaite được mô tả trong Luce và Raiffa, *Games and Decisions* (Trò chơi và các quyết định), 145<sup>m</sup>). Cách xây dựng bài toán của Braithwaite thành bài toán chủ trò một người, và việc tái lập công thức của Luce và Raiffa cho lý thuyết của Nash xét theo phán đoán hơn là chiến lược (các trang 121-154), dường như nhấn mạnh rằng *sự phối hợp tri thức* nằm ở tiêu điểm của lý thuyết. Một kết quả theo chủ nghĩa pháp lý đòi hỏi tính chất hợp lý hóa *nào đó* của một kết quả độc nhất; tính nguy biện đơn thuần là có ích nếu các kết quả thay thế là rỗng.

là một phương tiện tiềm năng hay không có thể tùy thuộc vào các giải pháp thay thế có sẵn.

Việc có các phương tiện khác để phối hợp, kể cả một số phương tiện có giá trị hơn ý niệm về tính đối xứng, dường như được thể hiện một cách dồi dào trong các thí nghiệm ở Chương 3. Do vậy ở góc độ thể hiện, có thể thiết lập các trò chơi trong đó tính đối xứng về toán học cung cấp tâm điểm cho các kỳ vọng được phối hợp, và có thể ở góc độ thể hiện thiết lập các trò chơi trong đó khía cạnh khác nào đó của trò chơi tập trung các kỳ vọng. (Các khía cạnh khác này thông thường không nằm ở trong cấu trúc toán học của trò chơi mà là một phần của "nội dung chủ đề"; nghĩa là, chúng thường tùy thuộc vào "việc đánh dấu" những người chơi và các chiến lược, là thuật ngữ mà Luce và Raiffa sử dụng được đề cập trong Chương 4).

Tôi không có cơ sở nào để lập luận với sức nặng nào, và với tỷ lệ phần trăm bao nhiêu các trò chơi thú vị, thì tính đối xứng về toán học lấn lướt "các kỳ vọng hợp lý." Nhưng tôi nghĩ rằng chỗ đứng của định đề về tính đối xứng bị thay đổi về mặt định tính bởi việc thừa nhận rằng tính đối xứng có các đối thủ cạnh tranh trong vai trò hội tụ các kỳ vọng. Bởi vì, nếu tin rằng các kỳ vọng của những người chơi duy lý có thể được làm cho trở nên nhất quán bởi một giá trị toán học nào đó của hàm lợi ích, thì tính đối xứng dường như có tiếng nói không thể tranh cãi, đặc biệt là nếu có thể tìm ra một định nghĩa độc nhất về tính đối xứng mà đáp ứng một số tiêu đề hấp dẫn nào đó. Nhưng nếu người ta phải thừa nhận rằng những yếu tố khác - những yếu tố không nhất thiết là một phần trong cấu trúc toán học của hàm lợi ích - có thể thực hiện điều mà tính đối xứng thực hiện, thì không có lý do *tiền nghiệm* nào để giả thiết rằng

điều mà tính đối xứng thực hiện là 99% hay 1% của công việc. Sự thu hút của tính đối xứng không còn là tính chất toán học nữa, nó mang tính chất nội quan; và việc tranh luận thêm bị giới hạn ở sự thu hút riêng của các công cụ tập trung cụ thể đối với nhà lý thuyết trò chơi *như một người chơi*, nếu không thì ở quan sát thực nghiệm.

Vì vậy, một lý thuyết chuẩn tắc về các trò chơi, một lý thuyết chiến lược, tùy thuộc vào sự phối hợp tri thức, có một yếu tố mà nó mang tính chất thực nghiệm; nó tùy thuộc vào việc người ta có thể phối hợp các kỳ vọng của họ như thế nào. Do đó, nó tuỳ thuộc vào kỹ năng và vào bối cảnh. Người chơi duy lý phải đặt chính mình trước câu hỏi thực nghiệm về việc làm thế nào, trong bối cảnh cụ thể của trò chơi của anh ta, hai người chơi duy lý có thể đạt được sự phối hợp ngầm các lựa chọn của họ, nếu anh ta cần phải tìm ra trong trò chơi một cơ sở cho việc chia sẻ một kỳ vọng *tiền nghiệm* về kết quả với đối tác của mình. Việc đồng nhất tính đối xứng với tính hợp lý tùy thuộc vào giả thiết rằng có các quy trình tri thức nào đó mà những người chơi duy lý không có khả năng nắm bắt, cụ thể như việc phối hợp các lựa chọn trên cơ sở bất cứ điều gì khác với tính đối xứng toán học, và rằng những người chơi duy lý có lẽ biết điều này. Đó là một câu hỏi thực nghiệm liệu những người chơi thích hợp có thể thật sự thực hiện điều mà một lý thuyết như thế đã phủ nhận rằng họ có thể thực hiện và do vậy nên bỏ qua các nguyên tắc chiến lược được sản sinh từ một lý thuyết như thế.<sup>18</sup>

<sup>18</sup> Điều thú vị là trong việc đòi hỏi một kết quả đối xứng cho một trò chơi ngầm có vẻ là đối xứng, Luce và Raiffa loại bỏ hai ứng viên có nhiều hứa hẹn nhất. Họ xem xét (trong *Games and Decisions*) một ma

trận, và ghi nhận rằng nó có các điểm cân bằng chiến lược thuận trong các ô phía trên bên trái và phía dưới bên phải.

	I	II
i	1 3	-1 -1
ii	-1 -1	2 1

Các yếu tố này bị loại trừ trên cơ sở rằng “bất cứ tính hợp lý hóa nào tôi đặt cho i hay ii thì, do tính đối xứng của tình huống, có một tính hợp lý hóa tương tự cho người chơi 2, và do vậy có vẻ như là không thể tránh khỏi rằng cả hai chúng ta cùng thua.” (Tôi đã thay thế i và ii cho việc đánh dấu họ). Họ sau đó xem xét một cặp chiến lược tối đa hóa tối thiểu, mà không thỏa mãn bởi vì chúng không đưa ra một điểm cân bằng, và một chiến lược tối thiểu hóa tối đa mà họ nhận thấy còn kém hơn. Nhưng câu hỏi quan trọng là liệu có phải hai người chơi vừa duy lý vừa có tinh tưởng tượng cũng hoàn toàn bất lực như cách mà Luce và Raiffa có khẳng định không. Liệu các người chơi có thể liên hệ qua lại các chiến lược của mình mà không cần trao đổi thông tin không? Đây là một câu hỏi thực nghiệm; các thí nghiệm trong Chương 3 đưa ra một câu trả lời khẳng định, hoặc ít nhất cũng chỉ ra rằng trong các trường hợp đặc biệt câu trả lời có thể là có. Ngay lập tức có thể dường như là khó cho họ để phối hợp trên một cặp chiến lược phi đối xứng. Nhưng phần khó khăn nhất chính là việc nhìn nhận rằng họ phải làm thế; câu hỏi làm điều đó như thế nào sau đó trở thành một vấn đề thực tiễn. Họ phải cùng nhau tìm kiếm một cách ngầm hiểu một manh mối cho việc phối hợp lựa chọn của họ. Dĩ nhiên, một kết quả phi đối xứng trong trận ở trên là một kết quả phân biệt; nó buộc một trong hai người chơi chấp nhận phần thường nhỏ hơn so với người kia vì những lý do mà có vẻ thuận do sự cố hay do ngẫu nhiên. Nhưng chúng ta phải giả thiết rằng một người chơi duy lý có thể gó ép mình để chấp nhận phần nhỏ hơn nếu manh mối chỉ theo hướng đó. Chỉ có một manh mối mang tính phân biệt có thể chỉ ra một sự lựa chọn có phối hợp; phủ nhận tính phân biệt là phủ nhận cơ sở rằng một manh mối có thể được cùng tìm thấy và cùng được thực hiện vì lợi ích của một kết quả mà cùng tốt hơn nhiều so với bất cứ kết quả đối xứng nào. Luce và Raiffa kết luận đê tài thảo luận của họ về trò chơi cụ thể này với lời nhận xét rằng “mặc dù trò chơi có vẻ như vô thường vô phạt này có chứa đựng một số yếu tố đối xứng, vẫn khó để có thể khai thác chúng.” Nhưng chìa khóa thực sự cho trò chơi có vẻ vô thường vô phạt này là nó có thể, đặc biệt là khi được trình bày trong một bối cảnh, có chứa đựng một số yếu tố *bất đối xứng*; và mục đích là phải khai thác được chúng. Xem thêm từ trang 449.

Một trò chơi nội quan, mà có thể đưa vào thực nghiệm, có thể minh họa cho điểm này. Hãy tưởng tượng các kết quả tiềm năng của một trò chơi mà có bao gồm tất cả các điểm ở trên hay trong phạm vi nào đó trong góc phần tư phía trên bên phải liên quan đến một cặp các tọa độ hình chữ nhật.

Cho dù chúng ta có bị cuốn hút mạnh mẽ vào định đề tính đối xứng hay không, cho dù chúng ta có bị đặc biệt cuốn hút vào tính đối xứng cụ thể trong đáp số của Nash hay không, chúng ta hãy đặt chính mình vào một khung trạng thái tinh thần phù hợp với việc chấp nhận "điểm Nash" như là một kết quả hợp lý của một trò chơi mặc cả công khai.<sup>19</sup> Nay giờ hãy xem xét một số biến thể của trò chơi này.

---

<sup>19</sup> Lời giải đề xuất bởi J.F. Nash cho các trò chơi mặc cả trong đó cả hai người chơi đều có sự hiểu biết đầy đủ về các hệ thống lợi ích của mình cũng như của lẫn nhau (các đánh giá chủ quan) là kết quả mà làm tối đa hóa sản phẩm của các lợi ích của hai người chơi. Nếu tất cả các kết quả khả dĩ đều được đặt trên một đồ thị mà các cặp tọa độ của nó đo lường các lợi ích mà hai người chơi rút ra từ chúng, đáp số là một điểm độc nhất trong phạm vi phía trên bên phải của vùng. (Điểm này là độc nhất bởi vì nếu có hai điểm, cả hai điểm có thể được nối kết với nhau bằng một đường thẳng trình bày các kết quả thay thế, có thể đạt được bằng cách kết hợp, với các lợi thế khác nhau, các khả năng xảy ra của hai kết quả ban đầu; và các điểm nằm trên đường thẳng sẽ tạo ra các sản phẩm cao hơn cho các lợi ích của hai người chơi. Nói cách khác, vùng được giả thiết là một vùng lồi bởi lý do có khả năng tạo ra các khả năng kết hợp, và một vùng lồi có một điểm sản phẩm-độ thỏa dụng-tối đa duy nhất, hay còn gọi là "điểm Nash").

Một đặc điểm để phân biệt "đáp số" đặc biệt này là nó độc lập với tỷ suất trao đổi giữa các thang độ thỏa dụng của hai người chơi. Nói cách khác, nó là bất biến trong tương quan với bất cứ khối lượng cố định nào mà chúng ta có thể gắn vào độ thỏa dụng riêng của từng người chơi. Và nó cũng đáp ứng một số điều kiện khác, đáng chú ý là bao gồm cả điều kiện rằng với mỗi cặp bất kỳ các khối lượng cố định (hay bất cứ tỷ lệ trao đổi nào) liên quan đến thang độ thỏa dụng của hai người chơi mà tạo ra một vùng đối xứng, thì điểm giữa ở phía trên bên phải là đáp số.

---

(Nó là đáp số duy nhất đáp ứng được tất cả các điều kiện đã được nêu; Nash chỉ ra rằng bất cứ đáp số nào đáp ứng các điều kiện của ông hẳn sẽ dẫn đến kết quả mà dựa đến sản phẩm tối đa cho độ thỏa dụng của hai người chơi). Với mục đích lúc này, chúng ta có thể xem yêu cầu về tính đối xứng như là tính chất chung của đáp số, và nghĩ về các điều kiện khác (các tiên đề) có mục đích giúp cho việc lọc khái niệm thô về tính đối xứng thành một điểm mà một đáp số độc nhất được bảo đảm. Xem các phần tham khảo trước đó (trang 442) liên hệ tới Nash, Harsanyi và Luce và Raiffa; cũng xin xem thêm phần minh họa hoàn hảo của lý thuyết Nash, với phần phê bình, bởi Robert Bishop trong tác phẩm *Lời giải Nash về Độc quyền tay đôi và hai đơn vị độc quyền*. Và để áp dụng "điểm Nash" vào lý thuyết phán đoán, xem Layman E. Allen trong *Các trò chơi mắc cá: Một ứng dụng được đề xuất về lý thuyết các trò chơi cho mắc cá tập thể* (Games Bargaining: A Proposed Application of the Theory of Games to Collective Bargaining), *Yale Law Journal*, 65:660 (tháng 4-1956).

Một cách tinh cù, có thể xứng đáng để nhấn mạnh rằng lý thuyết Nash không chỉ là một lý thuyết mà không cần một phương tiện để so sánh các thang độ thỏa dụng của hai người chơi - một lý thuyết mà, độc lập với các cách so sánh độ thỏa dụng giữa con người, có thể thích hợp với chúng. Nói đúng hơn, bởi vì nó sử dụng tính chất phán quyết của tỷ suất trao đổi độ thỏa dụng như là một nguyên tắc căn bản, lý thuyết này phải được nhìn nhận là *tùy thuộc* vào khả năng không thể so sánh vốn có của độ thỏa dụng. Nếu các thang độ thỏa dụng nơi hai người chơi về mặt nguyên tắc có thể so sánh được, mặc dù khó khăn, lý thuyết Nash sẽ dường như không phải là một phương tiện hấp dẫn để loại trừ các so sánh khó khăn. Nếu về nguyên tắc độ thỏa dụng là có thể so sánh được, sẽ có ít giá trị trong một lý thuyết mà dựa vào, trong việc đạt đến một lời giải, nguyên tắc về tính chất không thể so sánh. Và trong khi các cơ sở khái niệm ngày nay về lý thuyết trò chơi và lý thuyết kinh tế có vẻ không tương thích với so sánh độ thỏa dụng giữa con người, thì ý niệm về *trọng tài* có thể là không như thế. Lý thuyết kinh tế nhận thấy thuận lợi trong việc sử dụng một ý niệm về độ thỏa dụng mà lầm cho lý thuyết về giá trị sử dụng tương ứng với lý thuyết lựa chọn, sao cho người ta có thể đạt được "kinh tế học phúc lợi" như là một sản phẩm phụ miễn phí của một lý thuyết về lựa chọn kinh tế. Nhưng nếu người ta muốn kết luận trước về sự tương ứng này, vì các mục đích rút ra được các nguyên tắc về trọng tài, người ta có thể hoặc bị dẫn đưa vào một nỗ lực nhằm đo lường "độ thỏa dụng" theo một cách tâm sinh lý nào đó, hoặc phải thiết lập theo chủ nghĩa pháp lý một số quy ước nào đó để thực hiện một sự so sánh - một quy ước mà, mặc dầu mang tính chất vô đoán, tương thích với mục đích xã hội của trọng tài.

Trước tiên, chúng ta sẽ chơi cùng một trò chơi dưới dạng ngầm của nó. Mỗi người trong hai chúng ta sẽ chọn một giá trị dọc theo trục của riêng mình, và nếu điểm kết quả nằm trên hay trong phạm vi, chúng ta sẽ có được các số lượng (các giá trị sử dụng) được biểu thị bằng các tọa độ mà chúng ta lựa chọn. Tôi phỏng đoán rằng, trong nǎo trạng mà tôi đã yêu cầu - một nǎo trạng mà đã làm cho điểm Nash trở nên hấp dẫn đối với chúng ta trong trò chơi mặc cả-công khai - chúng ta có lẽ nên chọn điểm Nash. Không cần phải hỏi chính xác tại sao, chúng ta hãy tiếp tục tới một biến thể khác của trò chơi. Biến thể này cũng ngầm; nhưng nó khác ở chỗ chúng ta không nhận được gì trừ phi điểm mà các tọa độ của nó đã được chúng ta lựa chọn nằm *chính xác* trên giới hạn. Chúng ta không nhận được gì trừ phi chúng ta dốc cạn hết tất cả những gì đạt được. Sự thận trọng không đưa chúng ta đi tới đâu cả; mỗi người phải chọn đúng như người khác kỳ vọng anh ta làm. Tôi đề nghị rằng trong nǎo trạng hiện tại của chúng ta chúng ta phải lấy điểm Nash.

Cuối cùng, hãy xem xét một biến thể khác nữa. Chúng ta được cho xem biểu đồ của trò chơi mà vừa được chơi và được cho biết rằng lúc này chúng ta là những đối tác hoàn hảo, cùng thắng và cùng thua với nhau. Ý thức được sự việc là trò chơi lúc này của chúng ta có mô hình của một trò chơi mặc cả, chúng ta sắp sửa phải lựa chọn, không có trao đổi với nhau, những tọa độ của một điểm nằm chính xác trên đường giới hạn. Nếu chúng ta làm được, cả hai chúng ta cùng thắng giải - cùng giải thưởng bất chấp điểm mà chúng ta đã thành công trong việc cùng nhau lựa chọn - và nếu chúng ta thất bại trong việc lựa chọn điểm nằm trên đường ranh giới, chúng ta không nhận

được gì. Trong trò chơi thuần phối hợp này, tôi một lần nữa phỏng đoán rằng chúng ta có lẽ (chắc chắn sẽ) trong khuôn khổ hiểu biết lúc này của mình sẽ lựa chọn điểm Nash.

Tại sao? Đơn giản là vì chúng ta cần sự hợp lý hóa nào đó mà dẫn đến một điểm độc nhất; và trong bối cảnh này, tính tương đồng trong mặc cả cung cấp điều đó. Trừ phi có một góc ngoặt (mà dù gì cũng rất có thể là điểm Nash); hay một điểm trung tâm đơn giản khi mà đường giới hạn là một đường thẳng hay một đường cung vòng tròn (mà một lần nữa lại trùng khớp với điểm Nash); hay một hình thức gọi ý đặc biệt nào đó mà đường như chỉ hướng đến một điểm đặc biệt; hoặc trừ phi có một sự không thuần khiết (chẳng hạn như một dấu chấm trên đường ranh giới, từ một lỗi máy in, hay một điểm duy nhất mà các số tọa độ của nó là các số nguyên, và tương tự như thế), chúng ta có thể bị dẫn dắt tìm kiếm một định nghĩa “độc nhất” về tính đối xứng để dựa vào. Và tính đối xứng thể loại Nash thì đáng tin như bất cứ điều gì mà tôi có thể nghĩ tới - không đơn giản như một số điều (như chỗ giao nhau với một đường chéo 45 độ từ gốc của đồ thị và các chỗ tương tự như thế), nhưng ít mơ hồ trên chính mức độ phức tạp của nó.

Và, nếu điểm Nash lôi cuốn chúng ta một cách mạnh mẽ trong trò chơi mặc cả, nó hẳn phải như thế bởi vì chúng ta tự tin rằng nó lôi cuốn một cách tương tự đối với đối tác của chúng ta, người mà chúng ta tin rằng đến lượt mình cũng ý thức rằng quan điểm của chúng ta trùng khớp. Nó do đó phải lôi cuốn chúng ta trong trò chơi thuần phối hợp như một điểm độc nhất mà đối tác sẽ nhận ra hiển nhiên một cách hiển nhiên.

Điều này chứng minh hay gợi ý điều gì? Không phải là tôi đang lập luận ủng hộ cho điểm Nash. Thay vào đó tôi đang

lập luận rằng sự lôi cuốn của điểm Nash đối với một nhà lý thuyết trò chơi (như một người chơi nội quan) có thể là sự đảo ngược của thứ tự mà tôi vừa đi qua. Có thể nó là phẩm chất trung tâm của điểm Nash trong trò chơi thuần phối hợp - tính hữu ích rõ ràng của một khái niệm về tính đối xứng được xác định độc nhất, khi mà không có các yếu tố ô tạp phi toán học để sử dụng - mà làm cho nó có một sức tác động kiểm soát trong một biến thể ngầm trên đường giới hạn và hợp tác cao độ của trò chơi; điều đó đến lượt mình lại làm cho nó trở thành một hướng dẫn đáng tin cậy trong một biến thể vùng giới hạn ngầm ít đòi hỏi hơn của trò chơi; và điều đó đến lượt nó lại làm cho mất nhuệ khí bất cứ người chơi nào trong trò chơi mặc cả công khai là người mà có thể hy vọng rằng các kỳ vọng có thể tập trung vào bất cứ nơi nào khác.

Nói cách khác, bằng cách nêu thành định đề *nhu cầu phối hợp các kỳ vọng*, chúng ta dường như có một cơ sở lý thuyết cho điều gì đó như các tiên đề Nash. Điều mà một lý thuyết giống như lý thuyết Nash cần có là cơ sở rằng một giải pháp có tồn tại; nó là hiện tượng có thể quan sát nơi sự phối hợp ngầm mà cung cấp chứng cứ thực nghiệm mà (đôi khi) các kỳ vọng hợp lý có thể được tập trung ngầm vào một kết quả độc nhất (và có lẽ có hiệu quả), và điều đó khiến người ta giả thiết rằng có thể có một điều tương tự trong một trò chơi mà không cung cấp gì khác ngoài các giá trị toán học để xử lý. Lý thuyết Nash là sự minh chứng cho giả định này - sự minh chứng hoàn bị nếu nó áp đảo mọi giải pháp toán học cạnh tranh xét theo thẩm mỹ toán học. Tuy nhiên, tiêu điểm có được từ đó bị giới hạn trong vũ trụ toán học, mà vũ trụ toán học đó không tương đồng với vũ trụ của lý thuyết trò chơi.

## PHỤ LỤC C

### GIẢI THÍCH LẠI VỀ KHÁI NIỆM ĐÁP SỐ CHO CÁC TRÒ CHƠI “BẤT HỢP TÁC”

Trò chơi thuận lợi ích chung, hay trò chơi phối hợp, có thể cung cấp thêm cái nhìn về lập luận phía sau một số khái niệm lời giải nào đó trong lý thuyết trò chơi, đặc biệt là khái niệm về *đáp số theo nghĩa chặt* cho trò chơi “bất hợp tác”. Bằng cách nói “lập luận nằm phía sau các khái niệm này”, tôi muốn nói đến lập luận được quy cho những người chơi duy lý mà đối với họ những khái niệm đó có sức lôi cuốn.<sup>1</sup>

		I	II
I	I	1	0
	II	0	3
		II	0
		0	3

H. 25

<sup>1</sup> “Bất hợp tác” là tên truyền thống dành cho trò chơi không có trao đổi công khai. Thật không may là nó có thể gợi ý rằng sự hợp tác là không có khi không có trao đổi. Như đã chỉ dẫn trong các Chương 3 và 4, sự hợp tác – hỗ tương qua lại và được công nhận bởi mỗi bên – là một yếu tố quan trọng, thậm chí là một yếu tố thống lĩnh, trong nhiều trò chơi ngầm có tổng lợi ích khác không.

Các trò chơi ngầm thể hiện trong hình 25 và 26 được xem là có một *giải pháp theo nghĩa chặt*. (Trong hình 26 một sự lựa chọn chiến lược thứ hai hay thứ ba cho mỗi người chơi tạo thành đáp số). Định nghĩa của một đáp số, theo Luce và Raiffa, là như sau: “Một trò chơi bất hợp tác được xem là có một *đáp số theo nghĩa chặt* nếu: (1) Có tồn tại một cặp cân bằng trong số những cặp chiến lược cùng được thừa nhận. (2) Tất cả các cặp cân bằng được cùng thừa nhận là có thể thay đổi qua lại và tương đương.”<sup>2</sup>

	I	II	III
i	1 1	0 0	0 0
ii	0 0	3 3	3 3
iii	0 0	3 3	3 3

H. 26

Một *cặp cân bằng* là một cặp chiến lược cho người chơi sao cho mỗi chiến lược là chiến lược tốt nhất của một người (hay tốt như bất cứ chiến lược nào khác) mà có thể kết thành cặp với chiến lược của người kia. Một cặp chiến lược *cùng được thừa nhận* là một cặp mà không cùng bị lấn át bởi một cặp khác; nghĩa là, nó tạo ra một cặp kết quả mà cả hai cùng

<sup>2</sup> *Games and Decisions* (Trò chơi và các quyết định), từ trang 107. Khái niệm đáp số đặc biệt này thì na ná nhưng vẫn khác biệt với khái niệm đề xuất bởi J. F. Nash vào năm 1951. Để so sánh một số khái niệm đáp số có liên quan, xin xem Chương 5 trong Luce và Raiffa, và J. F. Nash trong *Trò chơi bất hợp tác* (Non-cooperative Games), *Annals of Mathematics*, 54:286-295 (1951).

không kém hơn các kết quả trong ô nào khác. Các cặp cân bằng là *tương đương* nếu, đối với từng người chơi riêng rẽ, chúng tạo ra các kết quả tương đương. Các cặp cân bằng có thể *thay đổi lẫn nhau* nếu tất cả các cặp được hình thành từ các chiến lược tương ứng cũng là các điểm cân bằng. (Do đó chúng là tương đương và có thể trao đổi qua lại chỉ khi tất cả các cặp được hình thành từ các chiến lược tương ứng cũng tương đương). Như vậy, các cặp chiến lược (ii, II), (iii, III), (ii, III) và (iii, II) trong Hình 26 chỉ thị các cặp tương đương, thay đổi qua lại được, cân bằng cùng được thừa nhận.

Luce và Raiffa, ngay sau định nghĩa này, đưa thêm bình luận sau, mà có thể xem như là điểm xuất phát của chúng ta: “Điều kiện thứ hai ngăn chặn sự *lẫn lộn* trong trường hợp các cặp cân bằng cùng được thừa nhận phi độc nhất.” (Tôi cho in nghiêng).

Chính xác là vấn đề *lẫn lộn* này, hay *sự mơ hồ*, nằm ở tiêu điểm của trò chơi phối hợp trong Chương 3. Trò chơi trong Hình 27 không có một *đáp số theo nghĩa chặt*. Các chiến lược thứ hai và thứ ba cho hai người chơi thì không thay đổi lẫn nhau được và không tương đương - chúng không tạo ra các cặp tương đương ở tất cả bốn tổ hợp. Không có sự khác biệt về lợi ích giữa hai người chơi trong lựa chọn chiến lược của họ; chỉ có nguyên nhân cho sự lẩn lộn. Trong Hình 25 họ biết chính xác những chiến lược nào để chọn lựa; trong Hình 26 họ biết nhiều như mức họ cần biết; trong Hình 27 họ không biết. Sự thất bại trong việc phối hợp trong Hình 27 khiến họ nhận được con số không cho mỗi bên, và không có một manh mối để phối hợp, họ có thể được giả thiết là có cơ may 50-50 trong việc giành được 3 cho mỗi bên, cho một giá trị kỳ vọng là 1.5.

	I	II	III
i	1	0	0
ii	0	3	0
iii	0	0	3

H. 27

Tại sao (ii, II) lại là đáp số gợi ý trong hình 25, chứ không phải là (i, I)? Một câu trả lời ngay lập tức là phần thưởng cho (ii, II) thì nhiều hơn là cho (i, I). Nhưng đây chỉ là một phần của câu trả lời. Một phần khác lộ diện nếu chúng ta xem hình 28, mà giống như Hình 25 trong thứ tự ưu tiên nhưng khác về các sức mạnh tuyệt đối của sự ưu tiên. Trong Hình 28, có vẻ như là điều quan trọng không phải là đạt được 10 hơn là 9, nhưng 9 hay 10 còn hơn là số zero. Nói một cách nôm na, hai cặp cân bằng là gần tương đương nhưng không thể thay đổi qua lại; và mặc dù những người chơi có thể ít bận tâm về việc liệu họ nhận được 9 hay 10, họ lại rất quan tâm đến việc làm sao không phải nhận con số 0. Lợi ích chính của họ là làm sao tránh được “sự lỗ lợn.”

	I	II
i	9	0
ii	0	10

H. 28

Họ cần tìm kiếm mạnh mồi nào đó, hay quy tắc nào đó, hay sự chỉ dẫn để phối hợp các sự lựa chọn của họ. Trong một trò chơi ở mức trùu tượng như ma trận trong hình 28, có ít yếu tố hướng dẫn họ ngoài những con số; và giữa các quy tắc thay thế về việc lựa chọn cặp nhỏ hơn hay cặp lớn hơn, thì cặp thứ hai có lẽ có tính thuyết phục cao hơn. Chúng ta có thể đặt câu hỏi đối với các người chơi sẽ đáng giá thêm bao nhiêu nếu có thêm một đồng đôla gắn liền với (ii, II) bằng cách so sánh nó với (i, I); liệu nó có đáng giá thật nhiều với tính chất của một thiết bị báo hiệu, và chỉ đáng giá một chút khi xét về khoản tiền tăng thêm. Chính sự khác nhau giữa 9 và 10 làm cho việc phối hợp các lựa chọn trở thành có thể. Trong hình 29, nếu chúng ta giả thiết rằng họ không thể tìm thấy quy tắc nào cho sự phối hợp, thì giá trị kỳ vọng của họ được nghĩ là 5 cho mỗi người.

	I	II
i	10 10	0 0
ii	0 0	10 10

H. 29

(Thật sự trò chơi trong Hình 29 nếu được trình bày trong ma trận như đã cho có thể không gây ra khó khăn. Các kết quả thực nghiệm trong Chương 3 ngụ ý rằng nó không cần thiết như thế. Một ma trận cụ thể cho phép có các sự tách biệt trái-phải, trên-dưới, trước-tiên-sau cùng-ở giữa.. Vì mục đích lúc này của chúng ta, chúng ta phải giả thiết rằng các chiến

lược đến với các người chơi theo dạng như thế và với việc đánh dấu những người chơi duy lý không có khả năng về mặt tri thức để yêu cầu chúng một cách không mơ hồ. Một trò chơi hoàn toàn dễ hiểu hay không mạnh mồi vô cùng khó hiểu thì được coi là phải có đầy sự đánh dấu và một bộ lợi ích đối xứng hoàn hảo. Một cách ngẫu nhiên, một trò chơi ngầm với các chiến lược nhiều vô hạn rõ ràng là không có dạng “thuần”; một số vô hạn các chiến lược chỉ có thể được trình bày cho những người chơi bằng phương tiện của một công thức tạo lập, và bất cứ công thức tạo lập nào cũng có xu hướng tạo ra cho những người chơi phương tiện nào đó để yêu cầu các chiến lược).

Tình huống có thể không khác biệt nhiều nếu chúng ta giả thiết rằng cặp chiến lược (ii, II) được gạch dưới, in đậm và có các mũi tên chỉ hướng về nó; hoặc có một chú thích nói rằng trong trường hợp lẩn lộn thì chủ trò gợi ý lựa chọn (ii, II). Điều mà những người chơi cần là dấu hiệu *nào đó* để phối hợp các chiến lược của họ; nếu họ không thể tìm ra nó trong cấu hình toán học của các phần thường, họ có thể tìm kiếm nó ở bất cứ cấu hình nào khác. Và chiến lược có thể xảy ra theo cách như thế, hay với việc đánh dấu hay với ý nghĩa rộng như thế, để cung cấp một cơ sở tiềm năng cho việc sắp xếp hay phân loại chúng mà những người chơi duy lý thấy là hữu ích.<sup>3</sup>

---

<sup>3</sup> Loại hình “tính duy lý” hay kỹ năng tri thức được đòi hỏi trong những trò chơi này thì cũng giống như điều được đòi hỏi trong việc giải các câu đố. Một câu đố là một bối cảnh trong đó người ta được mời gọi để tìm kiếm một manh mối, các quy tắc là mạnh mẽ phải không quá khó để tìm kiếm mà cũng không quá dễ. (Người ta ít nhất phải có thể nhận ra rằng anh ta lẽ ra phải nhận ra nó, khi nó được giải thích cho anh ta). Một câu đố chủ yếu là một bài toán của hai người; phương pháp luận

Thế thì gợi ý của phụ lục này là một giá trị quan trọng thích hợp cho một “đáp số theo nghĩa chặt” – một lý do tại sao những người chơi duy lý có thể lựa chọn nó – là khả năng ra hiệu, một phương tiện liên lạc ngầm, mà có sẵn cho hai người chơi để hỗ trợ cho sự hợp tác ngầm của họ khi mà thất bại trong phối hợp các chọn lựa là nghiêm trọng. Dĩ nhiên đây không phải là giá trị có ý nghĩa duy nhất của một đáp số như thế; nhưng nó có thể là một phần quan trọng làm cơ sở cho việc một người chơi lựa chọn nó.

Một cách thức khác để minh họa điểm này là chúng ta có thể, trong các trò chơi giống như những trò chơi đã được trình bày trong sách này, quy định việc sắp xếp liên lạc với các chi phí liên lạc nào đó và phân tích các trò chơi để xem liệu việc liên lạc có đáng với chi phí đó và các thông điệp nào được gửi qua những kênh nào sẽ tạo ra “đáp số”. Các “mạnh mồi” được thảo luận trong sách này khi đó sẽ có vẻ là sự truyền đạt quá tự do đến nỗi không thể tận dụng; và một câu hỏi thực nghiệm là sự tự do liên lạc nào mà một người chơi duy lý có lẽ có thể tìm thấy và công nhận. Giống như những hạn chế về thẩm mỹ hay cú pháp trong một ngôn ngữ giúp loại bỏ các cắt xén bóp méo trong một thông điệp được truyền đạt kém, thì các hạn chế về thẩm mỹ hay nghệ thuật sân khấu, các hạn chế về ngữ biện hay hình học, có thể giúp loại bỏ sự mơ hồ trong một tình huống đòi hỏi sự lựa chọn phối hợp ngầm.

---

của lời giải tùy thuộc vào việc một người khác đã cài đặt một thông điệp mà theo phán đoán của anh ta là khó tìm ra nhưng không quá khó. Về nguyên tắc, một người không thể lập ra hay giải các câu đố mà không có kinh nghiệm thực nghiệm; người ta không thể suy luận *tiền nghiệm* liệu một đối tác duy lý có thể hiểu được một gợi ý không. “Lý thuyết gợi ý” là một phần thực nghiệm vốn có của lý thuyết trò chơi.

	I	II	III	IV
i	10 10	0	0	0
ii	0	10 10	0	0
iii	0	0	9 9	0
iv	0	0	0	10 10

### H. 30

Điểm này có thể được đẩy đi xa hơn nữa. Hãy xem xét trò chơi trong hình 30. Một lần nữa hãy giả thiết rằng các chiến lược xảy ra theo một cách khiến cho việc sắp xếp chúng trở thành không thể về mặt tri thức đối với những người chơi duy lý, một cách cụ thể, không theo hình thức của một ma trận vuông đặc biệt, không được đánh dấu bằng các con số hay chữ, hoặc - nếu chúng được đánh dấu - với các tên được đảo lộn cho hai người chơi. Từ đây sẽ xuất hiện tình hình là nếu không có phương tiện phối hợp tốt hơn nào được nhận ra, thì "đáp số" có thể là cặp chiến lược (ii, III) với kết quả là 9 cho mỗi bên. Đây là điều ít được mong muốn nhất trong số các điểm cân bằng, nhưng nó có được tính độc nhất trong khi những điểm cân bằng kia gây ra sự lẩn lộn; nó cung cấp một manh mối để phối hợp các chọn lựa. Xét theo riêng cấu trúc *kết quả* (nghĩa là không đưa vào "việc đánh dấu", các ma trận được thực tiễn sẵn, hay bất cứ chi tiết nào khác bên ngoài cấu trúc thuần định lượng của trò chơi), khó nhận thấy đáp số này thè kém thuyết phục hơn nhiều, so với đáp số trong Hình 31,

mặc dù đáp số sau đáp ứng định nghĩa của Luce-Raiffa còn đáp số trước mâu thuẫn với nó.<sup>4</sup>

	I	II	III	IV
i	9 0	0	0	0
ii	0	9 0	0	0
iii	0	0	10 10	0
iv	0	0	0	9 0

H. 31

Trò chơi trong Hình 32 và 33, cả hai đều không có một đáp số theo nghĩa chặt, có vẻ thể hiện cùng một điểm. Nó “trông có vẻ như là” những người chơi có lập luận để chọn (ii, II) trong Hình 33. Một lập luận có thể là, khi thiêu vắng bất cứ cách thức nào để nhận biết có nên nhắm tới (i, I) hay (ii, II), người ta cần phải xem xét sự bảo đảm nào họ có thể dựa vào. Người chọn hàng ngang sẽ không nhận được gì nếu anh ta chọn sai hàng ngang bên trên, anh ta nhận được 5 nếu anh ta chọn sai hàng ngang bên dưới, “sai” nghĩa là anh ta thất bại trong việc liên hệ với đối tác của mình để đạt được 10. Anh ta sau đó có thể chọn hàng ngang bên dưới, lập luận rằng anh ta làm thế bởi vì ít nhất anh ta cũng nhận được 5

<sup>4</sup> Chứng cứ thực nghiệm cho những trò chơi này và những trò chơi tương tự có thể sẵn sàng đạt được cho chính anh ta bởi bất cứ độc giả nào muốn theo đuổi tiếp điểm này.

nếu anh ta không nhận được 10, và các cơ may của anh ta để nhận được 10 không xấu đi vì sự lựa chọn này. Có lẽ đây là tất cả mà "tính hợp lý" đòi hỏi nơi anh ta; nhưng có thể là sâu sắc hơn nếu lý luận như sau.

	I	II
i	10 10	0 0
ii	0 0	10 10

H. 32

	I	II
i	10 10	5 0
ii	5 0	10 10

H. 33

"Chỉ so sánh (i, I) và (ii, II) đối tác của tôi và tôi không có cách nào phối hợp sự lựa chọn của chúng tôi. Tuy nhiên, hẳn phải có cách thức nào đó, vậy chúng ta hãy tìm kiếm nó. Một nơi khác duy nhất để tìm kiếm là trong các ô (ii, I) và (i, II). Chúng có cung cấp cho chúng ta gợi ý nào chúng ta cần để phối hợp nhằm đạt được 10 mỗi bên không? Có, chúng có; chúng dường như "chỉ về hướng" (ii, II). Chúng cung cấp hoặc là một lý do, hoặc là một lời biện minh cho sự tin tưởng hay giả bộ rằng (ii, II) là tốt hơn (i, I); bởi vì chúng ta cần một lời biện minh, nếu không phải là một lý do, để giả bộ, nếu không phải để tin tưởng, rằng một trong các cặp cân bằng là tốt hơn, hay tách biệt hơn, hay nổi bật hơn, hay đáp ứng hơn cặp kia, và bởi vì tôi không tìm thấy quy tắc hay chỉ dẫn bắt buộc nào để đi theo hay manh mối nào để theo đuổi, chúng ta như vậy có thể đồng ý sử dụng quy tắc này để tư tưởng gấp nhau."

Trong trường hợp này những người chơi không chọn các chiến lược thứ hai của họ bởi vì 5 được ưa chuộng hơn là 0. Họ không có kỳ vọng nghiêm túc nào về việc chọn 5. Họ đang

sử dụng cấu hình của các số 5 và 0 như một *manh mồi* cho việc phối hợp các hành động. Nó có ích cho các người chơi - và mỗi người nhận ra rằng người kia cũng nhận ra nó là hữu ích - để ghi nhận đâu là nơi có các số 5, nhưng chỉ như một bước trong quy trình phối hợp các ý định. Xu hướng đối với ma trận trong hình 33 để "hội tụ" vào (ii, II) thì về nguyên tắc cũng tương tự như nếu ma trận được in ra có các mũi tên hướng đến góc phía dưới bên phải, các mũi tên không có vai trò lô-gic hay tính thẩm quyền khác với sức mạnh gợi ý và do đó là khả năng phối hợp các kỳ vọng.<sup>5</sup>

<sup>5</sup> Giả sử rằng một người chơi thật sự chọn ii hay II, có thể cũng đáng để tìm ra một cách thức hoạt động để phân biệt giữa các động cơ lựa chọn đó, thậm chí nếu chỉ để đảm bảo rằng khái niệm đó là có hoạt động. Như giữa hai động cơ đã được đề cập - động cơ "bảo đảm" và động cơ "manh mồi-phối hợp" chúng ta có thể phân biệt như sau. Chúng ta đề nghị với một người chơi các trò chơi thay thế như trong hình 33 mà chỉ khác biệt trong việc thay thế các giá trị từ 0 đến 9 bằng các giá trị 5 trong ma trận đó, để mặc các giá trị 10 và 0 yên một chỗ. Sau đó chúng ta yêu cầu anh ta "đánh giá" các trò chơi cho chúng ta - chỉ ra mức giá mà anh ta sẽ trả để có được cơ hội chơi trò chơi với một đối tác có thật và các khoản tiền thường có thật. (Trong một cách thay thế khác, chúng ta hỏi anh ta về mức giá mà anh ta sẽ trả để có được đặc quyền chơi các biến thể khác thay cho biến thể có kết quả là 5). Nếu câu trả lời của anh ta là khá vô tình đối với các dạng khác nhau trong kết quả đặc biệt đó chứng nào mà nó còn mang giá trị dương, và nếu tuy nhiên anh ta gắn cho trò chơi với kết quả dương nào đó một giá trị cao và gắn một điều gì đó như một giá trị kỳ vọng kiểu chiến lược ngẫu nhiên cho trò chơi với các kết quả là các con số 0 như trong hình 32, chúng ta có thể kết luận rằng các kết quả phía dưới bên trái và phía trên bên phải chủ yếu hấp dẫn đối với anh ta như những tín hiệu mà thôi. Nếu ví dụ như anh ta trả 9,95 đô-la cho một cơ hội để chơi trò chơi trong Hình 33 (có lẽ ngụ ý rằng một kỳ vọng 90% Cột sẽ chọn II), 8,65 đô-la cho trò chơi với giá trị 5 được thay thế bởi 1 (ngụ ý một kỳ vọng 85% chọn II), và 9,95 đô-la cho trò chơi với giá trị 5 được thay thế bằng 9 (ngụ ý một kỳ vọng 95% chọn II), và cuối cùng trả giá 5 đô-la cho trò chơi như trong Hình 32 (ngụ ý một kỳ vọng ngẫu nhiên giữa I và II), chúng ta có thể kết luận rằng hàm lợi ích, hay giá trị đối với người chơi, phía trên bên phải và phía dưới

## LỢI ÍCH XUNG ĐỘT

Bây giờ chúng ta có thể xem xét trường hợp phối hợp trộn lẫn với xung đột. Hình 34 và 35 minh họa các trò chơi có các điểm cân bằng, hai trong số chúng được cùng thừa nhận, không có “một đáp số theo nghĩa chặt” bởi vì các cặp cân bằng là không tương đương mà cũng không thể thay đổi qua lại được.

Vấn đề phối hợp trong điểm đầu tiên của hai điểm thì rõ ràng là “không thể giải quyết” trong hình thức thuần túy tương của nó, nghĩa là không có việc đánh dấu các chiến lược; có vẻ quá lầm là một cơ hội ngẫu nhiên trong việc đạt được một trong hai kết quả (có hiệu quả) cùng được thừa nhận<sup>(6)</sup>.

---

bên trái thì chủ yếu là hàm về manh mối phối hợp. Nếu thay vào đó anh ta trả giá bằng các khoản tiền mà ngu ý các khả năng giữa I và II là bất biến hoặc gần như thế, trong tương quan với các kết quả phía trên bên phải và phía dưới bên trái, và đặc biệt là nếu anh ta trả bằng một giá trị trung bình số học, thì việc diễn giải về sự bảo đảm sẽ được chỉ dẫn. (Lưu ý rằng các tinh từ “phía trên bên phải” và “phía dưới bên trái” chỉ là cách viết ở đây của tác giả; chúng không có ý nghĩa nào đối với người chơi bởi vì chúng ta đang xem xét trường hợp các chiến lược *không đánh dấu*, mà không thể được thể hiện trong một ma trận vuông, hoặc với các tên như “i” và “ii” - hoặc, nếu chúng có đánh dấu, thì hẳn phải được đánh dấu bởi một quy trình ngẫu nhiên tách biệt với quy trình ngẫu nhiên mà cung cấp tên hay các vị trí cho người chơi kia. Một cách cụ thể, Hàng phải không biết được liệu ma trận của Cột trông có giống như Hình 33 hay không, hoặc thay vào đó để cho các cột dọc thay đổi lẫn nhau với các kết quả giá trị thấp ở phía trên bên trái và phía dưới bên phải).

<sup>6</sup> Xem chú thích 18, Phụ lục B liên quan đến thảo luận về một ma trận tương tự khi cơ sở của sự trừu tượng thuần được giảm nhẹ.

Điểm thứ hai có thể là không phải không thể giải quyết được. Mỗi người chơi nên chấp nhận điểm cân bằng “tốt thứ nhì” của anh ta hơn là thất bại trong việc phối hợp; họ có một lợi ích chung trong việc hợp tác để tìm ra một manh mối cho sự lựa chọn chung. Tại sao lại không sử dụng manh mối chứa trong các ô khác, mà dường như chỉ hướng tới (ii, II)?<sup>7</sup>

		I	II
		4	0
i	6	0	
	0		6
ii	0		4

H. 34

Đối với một trong hai người chơi đây không phải là kết quả thuận lợi nhất, nhưng những người ăn xin không có quyền chọn lựa khi vận may lên tiếng. Có manh mối nào khác không? Có thể là *công bằng* như nhau khi sử dụng mặt trái của manh mối này; cũng như sẽ là công bằng như nhau nếu các mũi tên chỉ hướng về (ii, II) và hướng ra khỏi (i, I), xem các chiếc lông chim ở đuôi mũi tên như những dấu hiệu hơn là các đầu mũi tên. Nhưng sự công bằng không thể giúp ích; thực vậy nó làm cho sự phối hợp trở nên không thể. Nếu tất cả các manh mối đều đáng tin như nhau theo chiều ngược lại, chúng ta rơi trở về sự lẩn lộn. Chỉ có một manh mối *phân biệt* có thể chỉ ra một sự chọn lựa có phối hợp, phủ nhận tính phân biệt là phủ

<sup>7</sup> Trò chơi trong hình 35 còn có một điểm cân bằng khác, bao gồm một chiến lược hỗn hợp 80:20 cho Hàng và hỗn hợp 40:60 cho Cột. Nó mang lại kết quả 3,6 cho mỗi bên, và do vậy cũng bị lấn át bởi các ô phía trên bên trái và phía dưới bên phải.

nhận tiền đề rằng một manh mối có thể tìm thấy và hành động phối hợp để đạt đến một kết quả hữu hiệu khi sở thích xung đột.<sup>8</sup>

Một lần nữa ở đây, các manh mối có tiềm năng nhất có thể là những manh mối mà chúng ta thừa nhận khi chúng ta vượt qua tính chất toán học của ma trận lợi ích. Nếu chúng ta đang lái xe đến cùng một ngã tư từ các đường thẳng góc trên một sa mạc nơi không có hệ thống pháp luật nào xác định quyền ưu tiên đường, và không thích và không tin tưởng lẫn nhau và nhìn nhận rằng không có nghĩa vụ đạo đức nào giữa hai chúng ta, người đang tiến đến phía bên trái của người kia tuy nhiên vẫn có thể đi chậm lại để cho người kia băng qua trước, để tránh phải dừng khẩn cấp tại ngã tư; và người lái xe kia có thể dự liệu được điều này.<sup>9</sup> Hệ thống ưu tiên quy ước thiếu sức mạnh pháp lý hay đạo đức; nhưng nó rất thích hợp khi sự phối hợp được cần đến, đến mức mà người bị phân biệt đối xử có thể nhường bước cho nguyên tắc của nó, nhìn nhận rằng anh ta cần phải biết ơn một trọng tài, mặc dù nó phân biệt đối xử anh ta, và cũng nhìn nhận rằng anh ta bị mắc kẹt bởi việc người kia chấp nhận dấu hiệu và kỳ vọng rằng cả hai

<sup>8</sup> Sức mạnh của các dấu hiệu tương tự cùng được nhìn nhận đường như nằm phía sau khái niệm về "sự vượt trội tâm lý" được sử dụng bởi Luce và Raiffa để thảo luận về sự hấp dẫn trong một số trò chơi nào đó thuộc loại một điểm cân bằng cùng không được chấp nhận. Xem *Games and Decisions* (Trò chơi và các quyết định), các trang 109-110. Cũng xem thêm chú thích 18, Phụ lục B bình luận về một trò chơi tương tự.

<sup>9</sup> Một vấn đề xung đột lợi ích thuộc loại này - hai chiếc xe hơi cùng tiến tới trên một nơi chật hẹp bất đối xứng trên đường từ các hướng ngược nhau - được bao gồm trong bản câu hỏi mô tả trong Chương 3. Các kết quả xác nhận cho nguyên tắc chung, nhưng được loại bỏ cho vắn tắt khỏi Chương 3.

sẽ cùng tuân thủ. Qua cách lập luận này, như đã được phát triển trong Chương 3, trò chơi trong Hình 34 có thể giải quyết được khi được trình bày trong dạng ma trận *đặc biệt* đối với cả hai người chơi (nghĩa là được trình bày giống y như đã thể hiện trong Hình 34), hoặc khi cặp chiến lược thắng cuộc được *đánh dấu* như là “mặt sấp” và “mặt ngửa”, i, ii, I, và II, và tiếp tục như thế.

### SỰ THAO TÚNG BỞI MỘT BÊN THÚ BA

Một cách tình cờ, tất cả các trò chơi này đòi hỏi sự phối hợp, cả các trò chơi với các sở thích có xung đột cũng như các trò chơi với các sở thích trùng hợp với nhau, có thể tùy thuộc một cách đáng kể vào sự kiểm soát hay tầm ảnh hưởng của một nhà trung gian. Nếu chúng ta cho một người chơi thứ ba sức mạnh để đưa ra các thông điệp cho hai người chơi ngầm ban đầu, anh ta sẽ ở trong một vị trí tốt để giúp đỡ hai người chơi kia; anh ta thậm chí cũng ở trong một vị trí tốt để giúp cho chính anh ta nếu anh ta có được một thu hoạch mà tùy thuộc vào cặp chiến lược mà hai người chơi ban đầu lựa chọn. Một người trung gian tốt bụng làm cho một trò chơi thuần lợi ích chung trở nên dễ dàng đến mức vụn vặt; một người trung gian có khả năng phân xử về tính công bằng trong một trò chơi như trò chơi trong Hình 34;<sup>10</sup> một người trung gian có lợi

---

<sup>10</sup> Hãy nhớ lại bài toán số 8 ở Chương 3, bao hàm khoản tiền bị mất và tìm lại được, và một nhà trung gian tự chỉ định.

ích mạnh mẽ làm “người thứ ba” trong trò chơi trong hình 36, nơi mà các con số trong ngoặc đơn là phần thưởng cho người trung gian (hay một người độc quyền thông tin) mà ở vị trí cung cấp các chỉ dẫn - mang tính chất gợi ý không thô thiển, chứ không mang tính thẩm quyền - cho hai người chơi kia.

	I	II	III	IV
i	(2) 5	0	0	0
ii	0	7 (3) 9	0	0
iii	0	0	6 (4) 5	0
iv	0	0	0	9 (1) 7

H. 36

### *GIẢI THÍCH CÁC LỢI ÍCH*

Như một điểm cuối cùng, có thể lưu ý rằng, đối với lối lập luận được phát triển ở đây, sẽ không thành vấn đề nếu chúng ta giải thích các lợi ích như những thực thể có thể đo lường một cách khách quan, như tiền bạc hay các hàng hóa đồng dạng, hay như “độ thỏa dụng” theo một ý nghĩa lúc này đã là quen thuộc trong lý thuyết trò chơi. Nó không tùy thuộc vào sự hiểu biết của mỗi người chơi về các sức mạnh của các ưu tiên của người kia, chừng nào mà các lợi ích danh nghĩa được biết đến. (Nếu cả các giá trị khách quan lẫn các độ thỏa dụng được biết đến, và không có tỷ lệ tương ứng với nhau, các “dấu

hiệu" có thể mất đi sức mạnh nào đó; vấn đề lẩn lộn hay sự mơ hồ sẽ bị làm trầm trọng thêm).

## SỐ NGƯỜI CHƠI

Việc thảo luận ở đây cho đến lúc này đã xem xét các trò chơi hai người, ngoại trừ một phần xem xét ngắn về một người chơi thứ ba mà có thể đóng vai trò không ngầm. Nhưng vấn đề này có thể được mở rộng tới bất cứ số lượng người chơi nào, với các phần thưởng tùy thuộc vào hoặc là sự lựa chọn nhất trí hoặc là một loại lựa chọn đa số hay số nhiều hoặc là các kiểu liên minh thành công (tương tự đôi chút với các quy tắc về quy trình câu hỏi điều tra thực sự được mô tả ở Chương 3). Lúc này vấn đề về sự mơ hồ có thể trở nên nghiêm trọng hơn, và khía cạnh phối hợp của trò chơi có thể trở nên thậm chí thích hợp hơn đối với cơ sở lý luận về một "đáp số". Có lẽ là chính trong phạm vi của các trò chơi có nhiều hơn hai người mà lý thuyết về sự phối hợp trở nên thích hợp nhất, các trò chơi bao hàm công thức về các liên minh. Việc nghiên cứu các dấu hiệu và các kênh liên lạc trong việc hình thành liên minh có vẻ như là một cơ sở gấp gáp hiệu quả cho lý thuyết trò chơi và xã hội học.

## KẾT LUẬN

Nói tóm lại, lý thuyết trò chơi phối hợp gợi ý rằng “đáp số theo đúng nghĩa” của một trò chơi có tổng lợi ích khác không ngầm phi triệt tiêu chỉ được hiểu một phần, và trong một số trường hợp được hiểu rộng lớn hơn, bằng cách đề cập tới các phẩm chất báo hiệu của nó. Bởi vì các nguồn báo hiệu khác có thể hiện hữu *ngay cả trong cách tạo lập thuần toán học của trò chơi*, các phẩm chất đặc biệt của “đáp số theo đúng nghĩa” chỉ là một trong số nhiều yếu tố quyết định tiềm năng của một “lời giải hợp lý.” Việc các dấu hiệu có thể được đánh giá như thế nào một phần là một vấn đề thực nghiệm, chứ không phải chỉ là một vấn đề suy luận tiền nghiệm.



# CHIẾN LUỢC XUNG ĐỘT

THOMAS C. SCHELLING

---

*Dịch: Vũ Tiến Phúc - Nguyễn Văn Thắng*

*Hiệu đính: Ts. Trần Vinh Dự - Ts. Nguyễn Quang Thắng  
Đại học Tổng hợp Texas - Austin, Hoa Kỳ*

Chủ trách nhiệm xuất bản: Ts. Quách Thu Nguyệt

Biên tập: Xuân Nguyễn

Bìa: Thùy Trinh

Sửa bản in: Hoàng Yến

Kỹ thuật vi tính: Thanh Hà

## NHÀ XUẤT BẢN TRẺ

161B Lý Chính Thắng - Quận 3 - Thành phố Hồ Chí Minh

ĐT: 9316289 - 9316211 - 8465595 - 8465596 - 9350973

Fax: 84.8.8437450 - E-mail: nxltre@ hcm.vnn.vn

Website: <http://www.nxltre.com.vn>

## CHI NHÁNH NHÀ XUẤT BẢN TRẺ TẠI HÀ NỘI

20 ngõ 91, Nguyễn Chí Thanh, Quận Đống Đa - Hà Nội

ĐT & Fax: (04) 7734544

E-mail: vanphongnxltre@ hn.vnn.vn

---

Khổ 14,5x20,5cm. Số: 373-2007/CXB/19-56/Tre. Quyết định  
xuất bản số: 1083A/QĐ - Tre, ngày 13 tháng 11 năm 2007.  
In 1.500 cuốn, tại Xí nghiệp In Nguyễn Minh Hoàng.  
Địa chỉ: 100 Lê Đại Hành, P.7, Q.11, TP.HCM. ĐT: 8555812.  
In xong và nộp lưu chiểu tháng 11 năm 2007.

... Cuốn sách "Chiến lược xung đột" (The Strategy of Conflict) mà bạn đang cầm trên tay là sự tổng kết quá trình nhiều năm nghiên cứu và trải nghiệm của Schelling. Không phải ngẫu nhiên mà cuốn sách được hình thành và ra đời trong nỗi ám ảnh của chiến tranh lạnh, trong bối cảnh cuộc chạy đua hạt nhân giữa hai cường quốc thế giới lúc bấy giờ là Liên Xô và Mỹ ngày càng trở nên gay cấn. Schelling đã rất ngạc nhiên (và có phần thích thú nữa) khi phát hiện ra rằng ưu thế thương lượng (hay mặc cả) của mỗi bên chịu ảnh hưởng bởi nhiều yếu tố, đặc biệt là các lựa chọn khả hữu cũng như khả năng chủ động thay đổi những lựa chọn này của bản thân họ và của đối thủ. Bằng các công cụ toán học, ông đã chứng minh được điều tưởng chừng là nghịch lý rằng, trước một tình huống xung đột với đối thủ, một bên có thể chiếm ưu thế bằng cách "phế bỏ" một số lựa chọn đang có của mình. Kết luận này làm chúng ta nhớ lại câu "đặt vào chỗ chết thì sống, đặt vào chỗ mất thì còn" trong binh pháp Tôn Tử hay tích "qua sông đốt thuyền" của Hàn Tin. Trong thời đại ngày nay, những chiến lược này vẫn tiếp tục được vận dụng một cách linh hoạt.

... Không chỉ quan tâm tới xung đột, Schelling còn nghiên cứu các quá trình thiết lập môi trường cho sự tin cậy và cam kết chiến lược, nhờ đó sự hợp tác trong dài hạn có thể được duy trì. Trong quá trình hợp tác này, các bên cũng đã nhận thấy về lâu về dài họ cùng có thể được lợi nếu như ban đầu họ chịu hy sinh một phần quyền lợi của mình. Những phân tích này của Schelling đã giúp giải thích một phổ rộng lớn những hiện tượng thường gặp, từ chiến lược cạnh tranh của công ty cho tới việc ủy thác quyền quyết định về chính trị hay ngoại giao.

... Có lẽ không phải ngẫu nhiên mà Viện khoa học Hoàng gia Thụy Điển đã nhận định rằng, những đóng góp khoa học của Schelling đã góp một phần quan trọng trong việc giải quyết xung đột và nỗ lực để tránh chiến tranh. Hiểu rõ xung đột là cơ sở để này nở sự hợp tác



Trích Lời giới thiệu sách của TS. Vũ Thành Tư Anh

Chương trình Giảng dạy Kinh tế Fulbright

© Copyright 1960 by The President and Fellows of Harvard College.  
Bản tiếng Việt do Nhà xuất bản Trẻ độc quyền xuất bản theo Hợp đồng  
nhượng quyền với Harvard University Press tháng 3. 2006.

Chiến lược xung đột

8 934974 069485

Giá : 79.000 đ