

淺談「囚徒困局」現象的應用

李國偉

如果你去問一些數學家，數學是在作什麼？大部份可能會回答，數學是處理形與數的學問，它的組織方式包括了定義、定理、證明這些東西。但是人類有關數學性的活動，不僅僅是純數學家關心研究的那些題材，而是更廣泛的會與人類文化及生活息息相關。所以在一般我們認為可以產生數學問題的物理、化學題材外，一些人生現實、文化、社會的現象裡，也不是完全沒有數學的蹤影，端看我們如何用數學的工具來捕捉描述這些現象而已。例如：當我們在高速公路塞車的時候，如果大家規規矩矩的排在車道內，而有一個人違規行駛路肩，那他必定會佔到便宜。但是如果每個人都有相同的想法，也都付諸行動的話，則情況必然更為惡化，人人都沒有便宜可佔。這種情形其實經常出現在我們的生活裡，就是如果每個人守規矩做事，那麼第一個不守規矩的人會討到便宜；但是如果每個人都不守規矩的話，後果反而對人人均不利。這個現象的背後是否隱藏某些數學的訊息？為什麼我們要如此問呢？因為當一件事反覆的在各處發生時，它背後就很可能有某種結構存在；

而當這種特殊結構存在時，我們就有可能用數學工具把它形式化，再加以分析。講題裡所謂的「囚徒困局」就是一個這種類型的例子。

我們先來舉個例子說明為什麼叫「囚徒困局」：假使你你的朋友作了一些於法不容的事，結果被警察逮到關了起來。這時進來一位檢察官，他對你說如果你乖乖認罪，幫警方指證你的朋友的話，你將可獲無罪開釋。如果你不招，而你的朋友招供的話，你將會被判五年牢獄。假使你們兩個人都不招的話，因為還有相當的證據在警方手中，足夠判你們兩個人每人兩年。如果你們兩個分別都招供了，因為證據確鑿，對不起，每人都得坐四年牢。這樣的條件另一位檢察官正在跟你的朋友說明，請問在這種局面下，你是招也不招？

現在我們把這個問題分析一下，首先將這些利害關係製成一張表。其中的數字對，左邊數字的絕對值表示你的坐牢年數，右邊數字的絕對值表示你朋友的坐牢年數。因為坐牢總是件倒楣事，所以用負數代表滿意的程度，因此負得越多處罰得越厲害。

你的朋友	不 招	招
你自己		
不 招	-2, -2	-5, 0
招	0, -5	-4, -4

讓我們來分析一下，如果你的朋友招了，你若不招會關五年，你若招了會關四年，所以招供對你有利。倘若你的朋友不招，你若不招會關二年，你若招了就可以撒鴨子跑了，當然還是招供對你有利。總而言之，無論你的朋友招還是不招，你都會發現招供對你比較有利。你的朋友應該不比你更笨，相同的邏輯思考，也會使他選擇招供，於是你們就哥倆好的到牢裡四年。如果你們別那麼邏輯，但是有決不出賣朋友的江湖義氣，兩人都死不招認，反倒可以只關二年。這種「困局」是該歸咎於理性的邏輯嗎？

我們再舉一個例子，假使你要跟一位鑽石商作交易，但是因為某種不方便的理由，你們不能面對面做生意。你們約好分別把貨款與鑽石放在指定的地方，再各自去拿自己的目的物。在這種秘密的交易裡，誠信變成非常重要，否則豈不陪了夫人又折兵？不過誰不是先照顧自己的利益呢？倘若你放一袋廢紙，而不是一袋鈔票，對方如果也要詐，你至少不會損失；對方如果老老實實放了一袋鑽石，那你就大賺了一票。因此不論你的對手如何行動，你都是要詐比較有利。同樣這套邏輯，你的對手也不是不會理解，所以他也決定要詐，結果你們兩個誰也沒佔到什麼便宜，倒不如開始就誠實的交易，還彼此都能得到想得的東西。這種狀況我們也可用一張表來展示各人的滿意程度。

	合 作	耍 詐
合 作	2, 2	-1, 4
耍 詐	4, -1	0, 0

這張表與前面那張結構其實是一樣的，如果我們把第一張表的每一個數加4，就得到第二張表。我們可以更一般性的觀察下列的表：

	<i>C</i>	<i>D</i>
<i>C</i>	<i>R, R</i>	<i>L, T</i>
<i>D</i>	<i>T, L</i>	<i>P, P</i>

其中 *C* 表示合作 (cooperation), *D* 表示耍詐 (defection)。 *R, T, L, P* 是一些數值，分別表示報酬 (reward)、誘惑 (temptation)、損失 (loss)、懲罰 (punishment)。這些數值要滿足兩個條件：(1) $(T + L)/2 < R$; (2) $T > R > P > L$ 。第一個條件的作用使得你們不值得約好，彼此交替著一次合作一次耍詐，因為那樣不會比老老實實合作好。第二個條件才是關鍵，只要它成立，我們前面的推論方法都可以行得通，也就會帶入一種「囚徒困局」裡。

在以下的討論中，我們固定一張會導致「囚徒困局」的表。如果交易只進行一次，無論如何還是要詐比較不會損失。但是如果雙方確實有必要進行一連串的交易，事先並不知道何時會終了，那麼又該採取怎麼樣的交易策略呢？當然我們討論策略時，並不需要背負道德的包袱，也就是說交易的雙方都是絕對的以自我為中心。譬如交易進行若干次之後，如果你有可靠的消息說對方已經病重，

下次是他最後一次交易，那你會毫不猶豫的拿一袋廢紙去。

當然，策略有很多種，而且策略的好壞要看你是跟什麼樣的對手交易。譬如對方是一個永遠耍詐的人，那你只好毫不猶豫的跟他耍詐下去。如果你的對手是一位報復性強的人，也就是他首先會與你合作，一直等到你耍一次詐，他就跟你沒完沒了的耍詐下去。對付這種人，你應該先跟他合作，等你賺得差不多了，就詐他一次，吃他一票。自此之後，就陪他耍詐，保持不吃虧的情況。既然沒有永遠居優勢的策略，該如何決定合作還是耍詐，便成爲一個有趣的問題。

1979年美國密西根大學政治系教授 Robert Axelrod 作了一項研究，他找了十四位在對策論 (game theory) 發表過論文的教授，告訴他們要針對「囚徒困局」來一次策略的大競賽。也就是請他們各自送來一個策略的電腦程式，然後彼此捉對廝殺，每個人的目標是要攫取最多的分數。每一個參賽者的程式，對於對手的 *C* 或 *D*，可以回應以 *C* 或 *D*，而且能記憶住跟他交鋒的全部歷史，回應的辦法也允許採取隨機的方式。事實上在送進來的十四個程式之外，Axelrod 加了一個完全採用隨機步驟的程式叫 RANDOM。在這些送進來的策略中，有的只要兩行 BASIC 指令就寫完了，也有的用了六、七十行。每一個程式都要跟其他任一個程式 (還包括一個自己的複製品) 廝殺兩百回合。結果出人意料的，兩行的程式居然能獲得極高的分數。這是由資深對策論教授 Anatol Rapoport 所提出的，他的辦法是第一步一

定合作，之後作對手前一步作的動作。這個策略可名爲 TIT FOR TAT(以牙還牙)。

有一位名叫 Johann Joss 的人，他的策略與 TIT FOR TAT 非常相似，只是有十分之一的機會他會在對手合作過後，突然耍一次詐。如果把 JOSS 與 TIT FOR TAT 配對比賽，如表所列：

TIT FOR TAT	JOSS	
<i>C</i>	<i>C</i>	
<i>C</i>	<i>C</i>	
<i>C</i>	<i>C</i>	
·	·	
·	·	
·	·	
<i>C</i>	<i>C</i>	
<i>C</i>	<i>D</i>	← 1/10次的耍詐
<i>D</i>	<i>C</i>	
<i>C</i>	<i>D</i>	
·	·	
·	·	
·	·	
<i>C</i>	<i>D</i>	
<i>D</i>	<i>D</i>	← 1/10次的耍詐
<i>D</i>	<i>D</i>	
·	·	
·	·	
·	·	

一開頭兩者都採取合作的態勢，但 JOSS 的方法會有十分之一耍詐的機會。於是到了某一次交易，突然 JOSS 耍了一次詐，使得兩邊本來合作的態勢打破，而形成

CD, DC, CD, \dots 的交易關係。過了一陣，因為 JOSS 十分之一機會的耍詐，再一次的 D ，便使得兩邊交易關係完全陷入相互耍詐的死胡同。JOSS 這種方法在與各種不同的對手比賽時，因為他會不期然的突襲對方，往往導致像與 TIT FOR TAT 對陣時互信的瓦解。所以長期累積下來，JOSS 所獲得的點數就比不過 TIT FOR TAT。

在這一輪策略大賽中，像上面的例子提供了特別值得深思的課題。Axelrod 總結整個賽局的經驗說：

「從這場競賽裡獲得一項啓示，就是在各種力量競爭的環境中，如何降低回音效應是非常重要的。策略分析的高手必須深入三個層次：第一層是動作的直接反應，此項最爲明顯，因為耍詐一定贏過合作。第二層是間接的效應，也就是考慮對手會不會懲罰你耍的詐，絕大部份的參賽者都考慮了這一層。但是第三層要考慮在回應對手的耍詐中，有可能重複或擴大自己前面想佔人便宜的動作。因此以直接效應而言，耍一次詐可能有利，甚至把二次效應考慮進去也顯現不出弱點。真正的代價或許要在第三層效應時才付出，使得一次孤立的耍詐動作，因為回音效應而陷入彼此互懲的漩渦。參賽的好些策略沒有體會到這種長期影響，結果變成自己在懲罰自己，對手的作用只是使這種自我懲罰延遲幾步而已。...

分析策略賽局中，我們可以獲得不少啓示，知道應該如何去因應勢力競爭的環境。即使研究策略的專家，包括政治學、社會學、經濟學、心理學、數學的專家，也往往犯了過分強勢的錯誤，寬容心不夠，也太不相信對方善意的回應。」

Axelrod 認爲第一次競賽已經提供了足夠深思的議題，因此不妨來一次更大規模的競賽，看看大家能不能由第一次的教訓中，謀求更大的改進。於是他除了邀請第一次的參賽者外，更在電腦或賽局的專門雜誌上登廣告徵求競賽者。參賽的人都被告知第一次競賽的經驗教訓，而且如果有人以爲第一次的十四個策略中哪一個最爲有利，也可以用那個策略參賽。結果有六十多個參賽者，有趣的是只有 Rapoport 自己再送入 TIT FOR TAT。

經過一場捉對廝殺後，原本在前一輪中得分較高的那些策略方法，幾乎全都遭遇悽慘。唯獨 TIT FOR TAT 仍然高居領先地位，表面上看起來令人費解的是，當 TIT FOR TAT 與別人對壘時，頂多處於不吃虧的狀況，絕不會超越它的對手。但是從長期總體的效應來看，它卻能得到最高的分數。Axelrod 檢討了排名在前的策略的特長，歸納出下面幾個要素：

1. 絕不會主動先耍詐。
2. 不會對人家的善意回應毫無回饋，也就是說當對方耍了你一次詐之後，又再度對你作善意回應時，你不會因爲舊恨而跟他糾纏下去，你會用寬容的態度還以善意的回應。

3. 但是對人家的耍詐行動, 你會毫不遲疑的立即加以反擊。

第二次競賽更進一步啓發 Axelrod 可以用類似的方法作一次生態環境的模擬。開始時的參賽策略可以想像是一群原始的生命體, 經過一輪競賽後, 得分多的策略在下一輪中可以用較多的份數參賽。這種設計的意義是模擬演化過程裡, 較適者可以有較多的後代。

這種生態賽局經過上千輪的實驗後, 也有一些有趣的現象浮現。在前面說明的第二次競賽中, 也有一些經常主動耍詐攻擊人的策略, 會獲得相當高的分數。這些策略在生態賽局中的前兩百輪, 取得相當大的優勢, 數量很可觀的增加, 因為這個段落裡比較軟弱的策略被他吃定了。但是隨著弱肉被強食之後, 剩下的還有大量的像 TIT FOR TAT 的策略, 就不再有便宜可佔了。等到一千五百輪之後, 這些窮凶惡極的生物終於滅絕。

總結這些競賽的經驗, 會體認出 TIT FOR TAT 所以處處會取得優勢, 基本上它是一個會導引合作的策略。除了前面說的三種特性外, 它還有一個特性就是明確, 跟它交手的策略很快就能察覺到不容易剝削它。像是 RANDOM 那樣的策略, 因為太無章法可循, 對手只好把它當作一個非常不合作的敵人處理了, 因此它的積分一直是敬陪末座的。

Axelrod 寫了一本書叫做「合作的演化」, 把賽局導引出合作的例證, 再進行了深

入的考察。他發現在一片兇狠的生態環境中, 只要有一小撮誘導合作的生物, 就可以存活下去。而且它們如果具有像 TIT FOR TAT 的四種優點, 便可以在演化中漸漸取得上風, 大量的繁殖下去。而且有趣的是, 在一大堆惡漢中來了一小撮合作者, 它們能存活繁殖。但是一旦合作者取得上風, 就不在會被惡漢入侵破壞。

Axelrod 的研究有很大的啓示作用, 從較侷限的模擬例證中可體認出, 合作的產生並不需要有道德的動機, 但是一旦有了立足點, 從實際利益上合作者也會獲得好的報應。這種教訓運用到人類社會、政治、經濟的研究裡, 實在頗有耐人咀嚼的地方。

我們從簡單的「囚徒困局」模式出發, 瞭解一些有關選擇的難題。然後利用反覆操作「囚徒困局」的模式, 更引發了有關生態演化方面極具啓發性結果。雖然這種過程裡, 我們並沒有證明一大堆困難的定理, 其實它仍然是數學的精彩應用。數學提供了概念的架構與模擬的工具, 以及分析組織的方法。希望從這個例證中, 能使讀者了悟到數學應用的宏大範圍。

—本文作者任職於中央研究院數學研究所—