

Evolutionary Game Theory

進化穩定的純策略均衡
進化穩定的混合策略

1

進化博弈論

- 廣泛應用於生物學、經濟學、社會學，甚至人工智慧。
- 核心思想是，成功的策略（產生更好回報的策略）在人群中變得更加普遍，模仿自然選擇的過程。

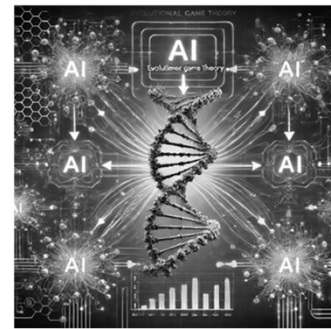
4

演化 vs 革命 vs 內捲化

Evolution vs Revolution vs Involution

2

Evolutionary game theory can be applied to AI, inspired by biological genes



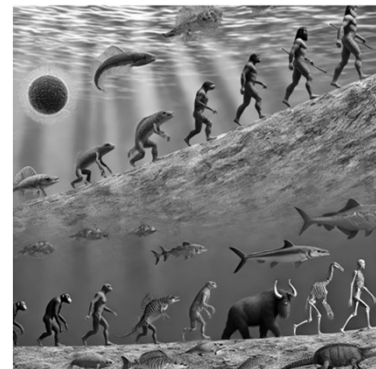
5

Darwin with his book "On the Origin of Species,"



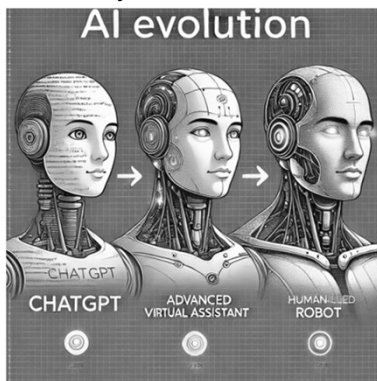
3

Evolution



6

the stages from 'ChatGPT' to 'Advanced Virtual Assistant' and finally to the 'Human-like Robot.'



7

演化博弈論的著名例子

- Hawk-Dove Game (Conflict and Aggression)
- Prisoner's Dilemma (Cooperation vs. Defection)
- Battle of the Sexes (Gender Roles and Coordination)

10

Survival of the fittest in the Jungle



8

演化博弈論的著名例子

- Public Goods Game (Social Cooperation)
- Rock-Paper-Scissors (Cyclic Dominance)
- Replicator Dynamics (Strategy Spread in Populations)
- Stag Hunt (Trust and Collaboration)

11

"Survival of the Fittest" in the Arctic,



9

進化 vs 革命 Evolution vs Revolution

- 進化是穩步前進的過程，漸進的變化為前進鋪平了道路。
- 革命是一種巨大的轉變，它會瞬間重塑景觀，挑戰規範並催生新時代。

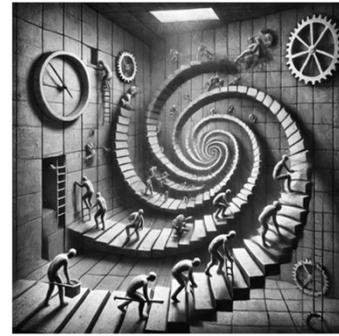
12

Revolution



13

involution, illustrating the feeling of being trapped in an endless, fruitless cycle.



16

革命 (Revolution)

- 革命是指突然的、徹底的或徹底的改變。
- 它涉及對現有秩序或系統的重大且常常是突然的背離。
- 政治革命，如法國大革命或俄國革命，涉及政府和社會結構的重大劇變。
- 技術革命，如工業革命，標誌著技術和工業的快速進步。

14

involution, capturing the repetitive and exhausting cycle associated with it



17

內捲化 (Involution)

- 美國人類學家 (Clifford Geertz) 前往印尼的爪哇島，發現當地人千百年來一直維持著刀耕火種的原始農業型態，生活方式和世界觀也同樣保持著千百年前的狀態
- 他們日復一日、年復一年地長期停留在簡單重複、沒有進步的輪迴狀態。這種現象取名「內捲化(Involution)。

15

演化 Evolution

- 演化是隨著時間的推移逐漸持續變化或發展的過程。
- 它涉及一系列小的、漸進的步驟，從而導致累積的轉變。
- 生物進化描述了物種透過自然選擇等過程一代又一代地逐漸發展。
- 社會或文化演化是指人類社會、價值觀和制度隨著時間的推移而逐漸發生的變化。

18

文化演化 cultural evolution

- 就像生物進化一樣，文化演化涉及思想、習俗和社會行為轉變的過程，有時是由於環境變化、技術進步或不同群體之間的相互作用。
- 從這個意義上說，文化隨著人們傳承知識、適應新環境或根據過去的經驗進行創新而演變。這一持續的過程使社會能夠發展出複雜的語言、多樣化的藝術形式和各種社會結構，以適應隨著時間的推移而不斷變化的需求和價值觀。

19

革命、進化、內捲的差別

- 雖然這三個概念都涉及變化，但革命的特徵是突然和徹底的轉變，進化涉及漸進和持續的發展，而內捲則代表導致衰退或倒退的向後或向內的運動。

22

社會演化 Society evolution

- 社會演化是指人類社會隨著時間的推移逐步發展和改變。
- 這種演變涉及社會結構、制度、技術、文化規範和經濟體系的改變。與代際遺傳變化驅動的生物演化不同，社會演化主要由文化創新、環境適應、經濟需求和社會互動所驅動。

20

《自私的基因》

- 以基因為中心的觀點，兩個個體的基因相關性越強，他們的合作行為就越有意義。
- 一個演化支系(lineage, 族群)會進化以最大化其包容性適應性 (inclusive fitness)，globally（而不是由特定個體 a particular individual）傳遞基因的拷貝數。結果，族群 (populations) 將趨向於演化穩定的策略 (evolutionarily stable strategy)。

23

政治體制的演化 the evolution of political systems

- 權力分立代表了政治體系的重大演變，旨在確保治理內部的製衡。
- 這個概念將政府權力劃分為不同的部門，通常是行政、立法和司法，以防止任何一個部門變得過於強大並維護民主原則。
- 這個制度是為了因應歷史上中央集權的濫用而發展起來的，是現代民主治理的基本要素。

21

《自私的基因》

- 「迷因」meme，是類似於基因的人類文化進化單位 (a unit of human cultural evolution)，可「自私」的複製，也可能在不同的意義上塑造人類文化。

24

利他的 altruistic

- 表現出幫助他人或為他人帶來好處的願望，即使這會對自己造成不利影響

25

間接互惠 (Indirect reciprocity)

- 我幫助你，別人幫助我。
- 間接互惠合作是建立聲譽 (reputation building)。

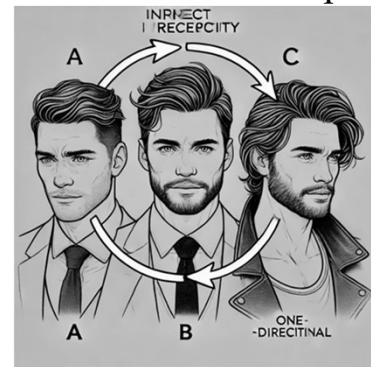
28

互惠利他主義 (Reciprocal altruism)

- 一個生物體的行為方式會暫時降低其適應性，同時增加另一個生物體的適應性，並期望其他生物體稍後會以類似的方式行事。

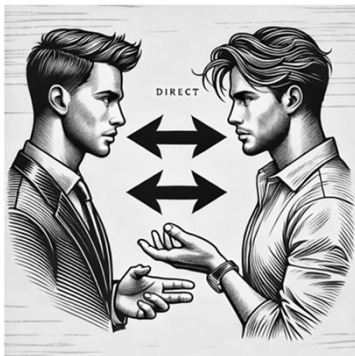
26

間接互惠 Indirect Reciprocity



29

Direct Reciprocity between two individuals



27

天鵝的團隊合作

Teamwork of Swans

- 天鵝在冬季遷徙時以 V 形編隊飛行。
- 他們作為一個團隊能夠獲得 80% 的效率。
- 每隻天鵝都只需要付出 20% 的努力，就可以排成隊形飛行。

30

A group of swans



31

擠成一團如何幫助企鵝保持溫暖？
How does huddling help penguins stay warm?

- 皇帝企鵝 (Emperor penguins) 生活在地球上最寒冷的氣候中，受著高達-60攝氏度的寒風。
- 它們的黑色羽毛可以吸收太陽的熱量，它們體內的脂肪可以隔熱，但是沒有什麼可以取代一個好的企鵝群。

34

a group of swans flying in the sky



32

擠成一團如何幫助企鵝保持溫暖？
How does huddling help penguins stay warm?



35

天鵝的團隊合作

Teamwork of Swans

- 必須付出100%努力的領頭天鵝與前面的天鵝交換位置，
- 他們自願承擔艱難的領導工作，每隻天鵝都佔據領先地位。
- 如果天鵝跌倒或受傷，幾隻天鵝會留下來照顧受傷的天鵝

33

擠成一團如何幫助企鵝保持溫暖？
How does huddling help penguins stay warm?

- 企鵝群的中心，可以達到高達攝氏 37 度（華氏 98.6 度）的溫度。
- 在中心的企鵝不斷在群中移動，以免過熱，而在外面的企鵝向內移動以獲得溫暖。

36

虎鯨殺死了一隻成年藍鯨
Killer Whales killed an adult blue whale

- 虎鯨（Orca），又稱殺人鯨：Killer whale）通過它們獨特的語言和行為，共同完成狩獵等重要任務。

37

A group of killer whales attacking a blue whale in the ocean



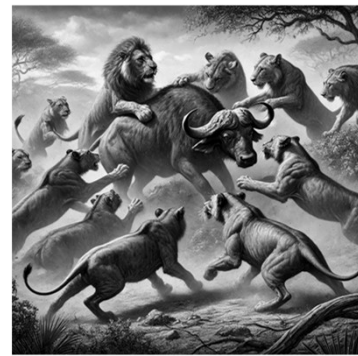
40

A group of killer whale



38

a group of lions attacking a buffalo in the wild



41

killer whales working together to hunt a seal



39

紅杉（Sequoia）

- 紅杉（Sequoia）是世界上最高的針葉樹，能長到115公尺高，主要分佈於美國加利福尼亞州。
- 加州紅杉都成群結隊，長成一片森林，在地底下，它們的根彼此緊密相連，形成一片根網，有的可達上千頃。

42

giant sequoia trees



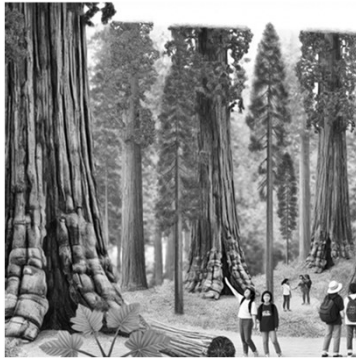
43

適者生存

- 倖存的不是最強大或最聰明的人，而是能夠最適應變化的人。” - 查爾斯·達爾文— Charles Darwin

46

giant sequoia trees with a group of tourists



44

適者生存

- 生存不只是體力的問題，它關乎適應的能力、忍耐的力量以及進化的智慧

47

紅杉 (Sequoia)



45

隨著時間的推移，COVID-19 逐漸變得更加溫和

- 病毒進化與適應：
 - 病毒會隨著時間的推移自然變異，通常以更容易傳播的方式發生變異。當病毒透過保持宿主的移動和與他人的互動（而不是讓他們患上重病並隔離他們）來傳播時，它往往會演變為變得不那麼嚴重。
 - 較溫和的 COVID-19 可以更有效地傳播，因為症狀較輕的人更有可能外出並與他人互動，從而幫助病毒傳播。

48

隨著時間的推移，COVID-19
逐漸變得更加溫和

- 人群免疫力：
 - 透過疫苗接種努力和對先前感染的自然免疫力，全球人群形成了部分免疫力。這種免疫力意味著，雖然人們仍然可能被感染，但他們的免疫系統通常已經做好了更好的準備，從而降低了疾病的嚴重程度。

49

進化穩定策略 (ESS)

(evolutionarily
stable strategy)

52

隨著時間的推移，COVID-19
逐漸變得更加溫和

- 變異體的自然選擇：
 - 引起較不嚴重疾病的變異體具有生存優勢，因為它們更容易在保持活躍的宿主中傳播。
 - 較嚴重的病毒株會導致人們病重並導致隔離，但其傳播效率往往較低。這種選擇壓力自然有利於較溫和的菌株。

50

進化穩定策略 (ESS)
(evolutionarily stable strategy)

- 在賽局論和行為生態學，進化穩定策略 (ESS) 是一種策略，是不能被任何突變策略入侵。

53

隨著時間的推移，COVID-19
逐漸變得更加溫和

- 醫學進步和適應：
 - 治療和預防措施也有助於減少疾病的整體影響。隨著更好的治療方案、更多關於控制症狀和預防性健康行為的知識，隨著時間的推移，感染的後果已經不那麼嚴重了。

51

進化穩定策略 (ESS)
(evolutionarily stable strategy)

- ESS是一個納許均衡，是“進化”穩定。
- 它足以防止突變的策略成功地入侵。

54

達爾文理論

(Darwinian Theory)

- 一個行為模式(A mode of behavior)只有在沒有其他模式是比較成功的產生後代才會存在。
- 如果採取策略 a 的適應度(fitness) 超過採取策略 a' 的適應度，採取策略 a 較快產生後代。

55

單態純策略均衡

Monomorphic Pure Strategy Equilibrium

- 有機體的行為模式很高概率(high probability)是其父代(parent) 的行為模式，而很低概率(low probability)的是隨機選擇的“突變”(“mutant)行為模式。

58

單態純策略均衡

Monomorphic Pure Strategy Equilibrium

- 每個有機體(organism)定期產生後代（繁殖是無性的）。

56

單一型態純策略均衡

- 一個有機體採取策略 a ，當它的對手採取策略 a' 。
- $u(a, a')$ 為適應性 (fitness), 後代的預期數量)。

59

a single-celled organism in an aquatic environment,



57

一般對稱賽局

General Symmetric Game

	$S \ (1-x)$	$T \ (x)$
$S \ (1-x)$	a, a	b, c
$T \ (x)$	c, b	d, d

60

一般對稱賽局
General Symmetric Game

- 假設在群體中：
- 非常小的 x 採取策略 T，
- $1-x$ 採取策略 S。

61

General一般對稱賽局
General Symmetric Game Game

- 如果對於 $x > 0$ 的所有足夠小的值，S 是演化穩定的：
- $a(1-x) + bx > c(1-x) + dx$

64

一般對稱賽局
General Symmetric Game

- 一個玩家在概率為 $1-x$ 的情況下，它遇到 S 的另一個玩家，得到 a 的收益；
- 而在概率 x 的情況下，它遇到 T 的玩家，得到了 b 的收益。
- 採取 S 策略的玩家的預期收益是 $a(1-x) + bx$ 。

62

一般對稱賽局
General Symmetric Game

- 當 x 變為 0 時，左側變為 a ，右側變為 c 。
- 如果 $a > c$ ，則一旦 x 足夠小，則左側較大。
- 如果 $a = c$ ，則當 $b > d$ 時，則左側較大。

65

一般對稱賽局
General Symmetric Game

- 採取 T 策略的玩家的預期收益是 $c(1-x) + dx$

63

一般對稱賽局
General Symmetric Game

- 如果
- (i) $a > c$ ，或
- (ii) $a = c$ 且 $b > d$ 。
- 在兩位玩家和兩個策略的對稱賽局，S 是進化上穩定策略。

66

$a=1, b=3, c=0, d=2$
由於 $a > c$ ，D是進化穩定的

	D	C
D	1, 1	3, 0
C	0, 3	2, 2

67

狩獵野兔是一種進化穩定策略

$a=3, b=3, c=0, d=4, (a > c)$

	Hare	Stag
Hare	3*, 3*	3, 0
Stag	0, 3	4*, 4*

70

$a=2, b=0, c=3, d=1$
由於 $a < c$ ，所以C不是進化穩定的

	C	D
C	2, 2	0, 3*
D	3*, 0	1*, 1*

68

狩獵雄鹿是一種進化穩定策略

$a=4, b=0, c=3, d=3 (a > c)$

	Stag	Hare
Stag	4*, 4*	0, 3
Hare	3, 0	3*, 3*

71

囚徒困境具有唯一的對稱納許均衡 (D, D)，並且納許均衡是嚴格的。因此，D是唯一的進化穩定策略。

	C	D
C	2, 2	0, 3*
D	3*, 0	1*, 1*

69

鷹派 vs 鴿派 Hawk-Dove

- 相同物種的兩隻動物爭奪一種資源 resource（例如食物或良好的築巢地點），其價值（以“適應度” fitness為單位）為 $v > 0$ 。
- （也就是說， v 表示資源帶來的預期後代數量的增加。）

72

鷹派 vs 鴿派 Hawk-Dove

- 每隻動物可以是攻擊性(鷹派, aggressive)也可以是消極被動(鴿派, passive)。
- 如果兩隻動物都具有攻擊性，它們會戰鬥直到一隻受到嚴重傷害；獲勝者獲得資源而不會遭受任何傷害，而失敗者則遭受 c 的損失。每隻動物獲勝的可能性均等，因此每隻動物的預期收益為 $v / 2 + (-c) / 2$ 。

73

鷹派 vs 鴿派 Hawk-Dove

- 如果 $v > c$ ， A 是優勢策略
- 有唯一的納許均衡 (A , A)
- (A , A) 是嚴格的納許均衡，因此 A 是唯一的演化穩定策略。

76

鷹派 vs 鴿派 Hawk-Dove

- 如果兩隻動物都是消極被動，那麼每隻動物都將以 $\frac{1}{2}$ 的概率獲得資源，而無需打架 (No Cost)。
- 最後，如果一種動物具有攻擊性，而另一種則是消極被動，那麼侵略者無需戰鬥即可獲得資源

74

鷹派 vs 鴿派 Hawk-Dove

- 如果 $v = c$ ，具有唯一的納許均衡 (A , A)。
- 但是均衡不是嚴格的：對於對手選擇 A ，無論選擇 A 還是 P ，玩家都可以獲得相同的收益。
- 但 (A , A) 是對稱，唯一的演化穩定策略。

77

鷹派 vs 鴿派 Hawk-Dove

	A	P
A	$\frac{1}{2}(v-c), \frac{1}{2}(v-c)$	$v, 0$
P	$0, v$	$\frac{1}{2}v, \frac{1}{2}v$

75

鷹派 vs 鴿派 Hawk-Dove

- 如果 $v < c$ ，則博弈在純策略中不具有對稱的納許均衡： (A, A) 和 (P, P) 都不是納許均衡。
- 因此，在這種情況下，博弈沒有進化穩定的策略。

78

鷹鴿博弈中進化穩定的混合策略

- 鷹鴿博弈有兩個不對稱納許均衡：(D,H) 和 (H,D)。
- 鷹鴿博弈無進化穩定策略，但有進化穩定的混合策略。

79

收益 Payoffs

- 獲得資源 $v = +50$
- 失去資源 0
- 輸掉戰鬥（受傷） $c = -100$
- 鴿派的成本 -10 (cost for posture)

82

鷹派 vs 鴿派 Hawk-Dove

- 鷹派兇猛，只有在遇到困難時才會撤退。
- 鴿派會威脅但不會傷害。
- 如果鷹派與鴿派打鬥，鴿派很快就會逃跑，不會受傷。

80

鷹派 vs 鴿派 Hawk-Dove

	Hawk	Dove
Hawk	$(1/2)(50) + (1/2)(-100) = -25$	+50
Dove	0.	$\frac{1}{2}(50) - (10) = 15$

83

鷹派 vs 鴿派 Hawk-Dove

- 如果鷹派遇到鷹派，它們就會一直戰鬥到其中之一受重傷或死亡為止。
- 如果鴿派遇到鴿派，它們會互相對視做樣子(posture)，直到其中之一厭倦或後退。

81

什麼是最好的策略 What is the best strategy

	Hawk (f(H))	Dove(1-f(H))
Hawk(f(H))	$(-25, -25)$	$50^*, 0^*$
Dove (1-f(H))	$0^*, 50^*$	15, 15

84

鷹派頻率

Hawk Frequency $f(H)$

- Payoff of hawk strategy is $25f(H) + (1-f(H)) = 50 - 75f(H)$
- Payoff of dove strategy is $15(1-f(H))$
- $50 - 75f(H) = 15 - 15f(H)$
- $60f(H) = 35$
- $f(H) = \frac{35}{60} = \frac{7}{12} = 0.583$

85

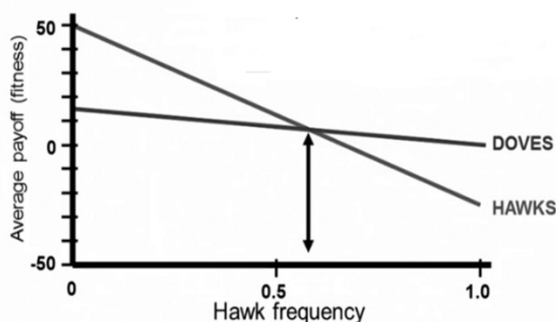
平衡的鷹派頻率

Equilibrium Hawk Frequency is 0.583

- 有一個平衡的鷹派頻率
- There is an equilibrium hawk frequency
- 鷹派玩家和鴿派玩家沒有最好的“純粹”策略
- No best “pure” for Hawk and Dove
- 鴿派玩家可以入侵“全鷹派玩家”
- An all hawk population can be invaded by doves.

88

$$50 - 75f(H) = 15 - 15f(H)$$



86

平衡的鷹派頻率

Equilibrium Hawk Frequency is 0.583

- 鷹派玩家可以入侵“全鴿派玩家”。
- An all dove population can be invaded by hawks.
- 策略的適應性是“依頻率而定”
- Fitness of strategies is “frequency dependent”

89

平衡的鷹派與鴿派收益

- 隨著時間的推移，鷹派與鴿派的比例保持穩定。
- 在穩定的情況下，鷹派和鴿派的平均收益是相同
- $\frac{5}{12}$ 鴿派 和 $\frac{7}{12}$ 鷹派
- 平均收益為 $\frac{7(-25)}{12} + \frac{5(50)}{12} = \frac{75}{12} = 6\frac{1}{4}$

87

Best strategy depend frequency of other strategies

- Hawk :
 - better than dove when many doves
 - worse than dove when many hawks
- Dove :
 - better than hawk when many Hawks
 - worse than hawk when many doves

90

每一個群體採取鷹派和鴿派的混合策略

- 每一個群體在鷹鴿博弈沒有進化穩定的策略，但有進化穩定的混合策略，一部分 x 採取鷹派策略，另一部分 $(1-x)$ 採取鴿派策略。

91

五個策略的分析

94

每一個群體採取鷹派和鴿派的混合策略

- 每一個有政治群體如國家、政黨等，一部分 x 採取鷹派策略，另一部分 $(1-x)$ 採取鴿派策略。

92

Dove



95

Evolutionarily Stable Mixed Strategies in the Hawk-Dove Game.

- 由於任何進化穩定的混合策略都必須對應於博弈的混合納許均衡，這為我們提供了一種尋找可能的進化穩定策略的方法

93

鷹派 vs 鴿派 Hawk-Dove

- 鷹派兇猛，只有在遇到困難時才會撤退。
- 鴿派會威脅但不會傷害。
- 如果鷹派與鴿派打鬥，鴿派很快就會逃跑，不會受傷。

96

Hawk



97

報復者 (The retaliator)

- 報復者是條件的戰略家 (conditional strategist)，它的行為取決於在對手。
- 它一開始就像鴿派，但是，如果受到攻擊，會報復並表現得像鷹派。
- 如果報復者遇到另一個報復者，就會像鴿派。

100

鷹派 vs 鴿派 Hawk-Dove

- 如果鷹派遇到鷹派，它們就會一直戰鬥到其中之一受重傷或死亡為止。
- 如果鴿派遇到鴿派，它們會互相對視做樣子 (posture)，直到其中之一厭倦或後退。

98

The Retaliator



101

鷹派 vs 鴿派 Hawk-Dove

- 如果鷹派遇到鷹派，它們就會一直戰鬥到其中之一受重傷或死亡為止。
- 如果鴿派遇到鴿派，它們會互相對視做樣子 (posture)，直到其中之一厭倦或後退。

99

惡霸 (the bully)

- 他的行為就像鷹派，但在受到攻擊時會逃跑

102

The Bully



103

五個策略的分析

- 鴿子族群不穩定，因為它可能會受到鷹派和惡霸的入侵。
- 鷹派的族群不穩定因為它將會受到鴿派和霸凌者的入侵。

106

探測者-報復者 (prober-retaliator)

- 1. 如果對手不反擊，他就會繼續像鷹派一樣。
- 2 如果對手反擊，又恢復打鴿派。
- 3 如果受到攻擊，它會像報復者一樣報復。

104

五個策略的分析

- 惡霸族群是不穩定，因為它可能會被鷹派入侵。
- 探測者-報復者族群幾乎穩定，但只有報復者做得更好，只是輕微地。
- 報復者族群，沒有其他策略可以入侵。

107

the "Prober-Retaliator"



105

報復者在演化上是穩定

- 如果所有五個策略都一起模擬，只有報復者在演化上是穩定的

108

海盜(pirate) vs. 漁民 (fisher)

109

海盜(pirate) vs. 漁民 (fisher)

- 一個 Sea Bird (海鳥) 具有兩種生存策略，那麼它們的進化如何“決定”呢？

112

海盜(pirate)



110

海盜(pirate) vs. 漁民 (fisher)

- 海鳥可以用來養活自己的兩種策略：漁民 和 海盜。
- 鳥可以進化為漁民從海洋中捕魚，而鳥也可以進化為海盜將魚偷走。

113

漁民 (fisher)



111

海盜(pirate) vs. 漁民 (fisher)

- 兩者都是可行的策略，為什麼有些鳥類最終成為漁民，而另一些卻最終成為海盜？

114

海盜(pirate) vs. 漁夫(fisher)

- 對於物種而言，那種策略最穩定？
- “這取決於海鳥數目的狀況。”

115

海盜(pirate) vs. 漁夫(fisher)

- 當一個物種面臨如何最佳生存的選擇時，保持兩者之間的健康比例是確保其物種的最穩定策略。
- 這種健康比例的最著名例子，也許是人類的50/50性別比。

118

海盜(pirate) vs. 漁夫(fisher)

- “如果全部海鳥都在捕魚，那麼很可能一隻海鳥就可以突變成為海盜，因為很容易偷魚。”

116

性別比例 Sex Ratio

- 基因可以決定後代的性別，這賦予最高的適應性
- 沒有確定性別的最佳“純粹策略”。
- 性別決定的適應性是“頻率依賴的”

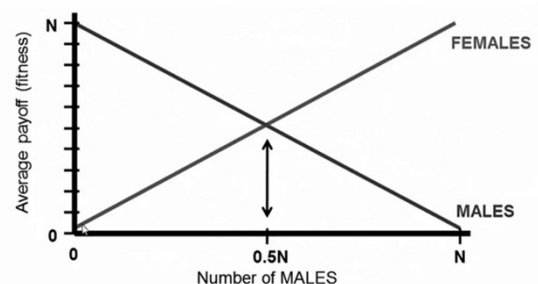
119

海盜(pirate) vs. 漁夫(fisher)

- 因為海盜比漁夫有優勢，所以海盜的基因會傳播開來，在幾代海鳥之後，每隻海鳥都是海盜。
- 但是在這種情況下，沒有更多的漁夫捕的魚可以偷走。
- 漁夫不是一種穩定的策略，因為它很容易被一隻具有海盜遺傳的突變鳥類偷魚。

117

性別比例 Sex Ratio



120

複製方程式是進化論 的基本理論 the replicator equation

121

複製方程式是進化論的基本理論

- The basic form of the replicator equation is:

$$\bullet \frac{dx_i}{dt} = x_i (f_i - \bar{f})$$

- where:
- x_i is the proportion of individuals using strategy i ,
- f_i is the fitness of individuals using strategy i ,
- \bar{f} is the average fitness of the population.

124

複製方程式是進化論的基本理論

- 複製方程式確實是演化的基礎數學模型之一。它描述了人口中不同類型的個體（或「策略」）的比例如何隨著時間的推移而變化，基於它們的相對適應性或繁殖成功率。

122

複製方程式是進化論的基本理論

- If a strategy's fitness f_i is higher than the average fitness \bar{f} , its frequency will increase.
- If it's lower than the average \bar{f} , its frequency will decrease.

125

複製方程式是進化論的基本理論

- 簡而言之，它模擬了「成功」策略（那些導致更高繁殖成功率的策略）如何在人群中變得更加普遍，而不太成功的策略則如何減少。

123

複製方程式是進化論的基本理論

- 這個概念有助於解釋有利的特徵或行為如何透過自然選擇傳播，從而導致演化。
- 複製方程式廣泛應用於生物學、生態學甚至經濟學等領域，用於模擬族群內的競爭、合作與適應。

126

複製方程式、遺傳演算法 和強化學習

- 複製方程式、遺傳演算法和強化學習都包含了適應、選擇和最佳化機制。

127

簡單族群模型

Simple Population Model

- 複製基因的 族群大小 (Population) $a(t)$
- $a(t)$ (Population) 可以根據適應性參數 f_A 隨時間 改變

$$\frac{da}{dt} = af_A$$

$$a(t) = a_0 e^{f_A t}$$

130

複製方程式、遺傳演算法 和強化學習

- 它們的主要區別在於具體的機制和應用，但共享進化啟發的學習和改進的共同基礎。
- 每種方法都以自己的方式模擬從回饋中“學習”，以隨著時間的推移達到更好的狀態或解決方案。

128

從長遠來看 $a(t)$

- $f_A > 0$, 爆炸 Explosion
- $f_A = 0$, 常數 Constant
- $f_A < 0$, 滅絕 Extinct

131

複製方程式 Replicator Equation

129

複製動力學

Replicator Dynamics

- $N(t)$ 個複製基因的群體。
- 每個複製基因的被分配策略 A (A 型基因) 或 B (B 型基因) 進行兩玩家博弈
- 在每個時期，每個複製基因都會與另一個複製基因隨機匹配進行兩玩家博弈

132

複製動力學

Replicator Dynamics

- 協調博弈 (coordination game) 的收益表 (the payoff table)

	A(x)	B (1-x)
A (x)	a , a	b, c
B (1-x)	c,b	d,d

- $a > c$ 且 $b < d$, 起始頻率 x : A型基因, 起始頻率 $1 - x$: B型基因.

133

複製動力學

Replicator Dynamics

- $N(t) = a(t) + b(t)$
- A型基因族群大小為
 - $x_A = \frac{a(t)}{N(t)}$
- B型基因族群大小為
 - $x_B = \frac{b(t)}{N(t)}$

136

收益表 (the payoff table)

- 對手A型基因, 則A型基因的收益是 $\pi_A(A, A) = a$ 。
- 對手B型基因, 則A型基因的收益 $\pi_A(A, B) = b$
- 對手A型基因, 則B型基因的收益是 $\pi_B(B, A) = c$ 。
- 對手B型基因, 則B型基因的收益 $\pi_B(B, B) = d$

134

每種基因類型的頻率
Frequency of Each Type

- A型基因的起始頻率是

$$x = \frac{a(t)}{N(t)} = \frac{a(t)}{a(t)+b(t)}$$
- B型基因的起始頻率是

$$1-x = \frac{b(t)}{N(t)} = \frac{b(t)}{a(t)+b(t)}$$
- $N(t) = a(t) + b(t)$,

137

複製動力學

Replicator Dynamics

- 複製基因的適應度 f 由收益 (如收益表所示)
- $x_A \geq 0, x_B \geq 0, x_A + x_B = 1$,
- $x_B = 1 - x_A = 1 - x$

135

演化模型

Evolutionary Model

- 演化在關於特徵頻率 (trait frequencies) 變化
- 基因有兩種類型: A 或 B
- $\frac{d(a)}{dt} = a(t)f_A$
 - $a(t) = a_0 e^{f_A t}$
- $\frac{d(b)}{dt} = b(t)f_B$,
 - $b(t) = b_0 e^{f_B t}$

138

A型基因的適應度

- $f_A(x, 1-x)$
 $= (\text{與A型基因互動的機率} * \pi_A(A, A) + (\text{與B型基因互動的機率}) * \pi_A(A, B))$
 $= x*a + (1-x)*b$

139

B型基因的適應性

- $f_B(x_A, x_B) = x*c + (1-x)*d$

142

A型基因的適應性

- $f_A(x_A, x_B)$
- $= x*a + (1-x)*b$

140

群體的平均適應性

- 適應性被解釋為繁殖率
- 群體的平均適應性 \bar{f} , 是兩個基因的適應性的加權平均值。

143

B型基因的適應性

- $f_B(x, 1-x)$
- $= ((\text{與A型基因互動的機率}) * \pi_B(B, A) + (\text{與B型基因互動的機率}) * \pi_B(B, B))$
- $= x*c + (1-x)*d$

141

群體的平均適應性 \bar{f}

- $\bar{f}(x_A, x_B)$
- $= x*f_A(x_A, x_B) + (1-x)*f_B(x_A, x_B)$

144

複製方程式的推導

Derivation of Replicator Equation

- $x = \frac{a(t)}{N} = \frac{a(t)}{a(t)+b(t)},$
- $N(t) = a(t)+b(t),$
- $1-x = \frac{b(t)}{N} = \frac{b(t)}{a(t)+b(t)},$
 $N(t) = a(t)+b(t),$

145

複製方程式的推導

Derivation of Replicator Equation

- $\frac{dx}{dt} = (a+b)^{-2} \left(\frac{da}{dt} (a+b) - a \left(\frac{da}{dt} + \frac{db}{dt} \right) \right)$
 $= (a+b)^{-2} \left(\frac{da}{dt} b - a \frac{db}{dt} \right)$
 $= (a+b)^{-2} (a f_A b - a b f_B)$
 $= x(1-x)(f_A - f_B)$

148

Quotient Rule

- derivative of f/g
 $= (f' g - g' f) / g^2$

146

複製方程式的推導

Derivation of Replicator Equation

- $\frac{dx}{dt} = x(1-x)(f_A - f_B)$
- $= x(f_A - f_B - x f_A + x f_B)$
- $\bar{f} = x f_A + (1-x) f_B$
- $= x f_A - x f_B + f_B$
- $\frac{dx}{dt} = x(f_A - \bar{f})$

149

複製方程式的推導

$$x = \frac{a(t)}{a(t)+b(t)}$$

$$\frac{dx}{dt} = \frac{d\left(\frac{a}{a+b}\right)}{dt}$$

$$= \frac{1}{(a+b)^2} \left(\frac{da}{dt} (a+b) - a \left(\frac{da}{dt} + \frac{db}{dt} \right) \right)$$

147

複製方程式的推導

Replicator Dynamics

- $\frac{dx}{dt} = x(f_A - \bar{f})$
- $\frac{dx}{dt} = x * (f_A(x_A, x_B) - \bar{f}(x_A, x_B))$
- $\frac{dx}{dt}$ 是 A 型基因的繁殖率

150

複製方程式的推導
Replicator Dynamics

- 按照同樣的邏輯

$$\frac{dN}{dt} = N * \bar{f}(a, b)$$

151

平衡條件

- 平衡條件： $\frac{dx}{dt} = x(1-x)(f_A - f_B) = 0$
 $\Rightarrow x=0, x=1, f_A = f_B$
- (1) 如果 $x=0$, $a(t)=0$, $b(t)=1$, 僅B型基因存在。
- (2) 如果 $x=1$, $a(t)=1$, $b(t)=0$, 僅A型基因存在
- (3) 如果 $f_A = f_B$ 然而, f_A 和 f_B 是任意常數 (arbitrary constant), 則不太可能發生

154

複製方程式的推導
Replicator Dynamics

- A型基因以 f_A 的速度繁殖，

$$\frac{da}{dt} = (f_A(x_A, x_B) - \bar{f}(x_A, x_B))$$

$$= a * f_A(a, b)$$

152

長期行為

- 如果 $f_A > f_B$, 最終只有A型基因存在。
- 如果 $f_A < f_B$, 最終只有B型基因存在。

155

複製方程式的推導
Replicator Dynamics

- 如果： $x > 0$ 且 $f_A(x_A, x_B) > \bar{f}(x_A, x_B)$ 則 $\frac{dx}{dt} > 0$
- A型基因的數量將會增加

153

兩種類型

- Type A' payoff: $f_A = ax + b(1-x)$
- Type B's payoff: $f_B = cx + d(1-x)$
- $f_A = ax + b(1-x) = f_B = cx + d(1-x)$
- $ax + b - bx = cx + d - dx$
- $x^* = \frac{(d-b)}{(a-b-c+d)}$

156

兩種類型

- 求 $x^* = \frac{(d-b)}{(a-b-c+d)}$
- 僅當 $0 < x^* < 1$ 時才有意義。

157

Average Payoff in Population

- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ 1-x \end{bmatrix}$
- $\mathbf{x}^T A \mathbf{x} = \begin{bmatrix} x & 1-x \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ 1-x \end{bmatrix}$
 $= \begin{bmatrix} x & 1-x \end{bmatrix} \begin{bmatrix} ax + b(1-x) \\ cx + d(1-x) \end{bmatrix}$
 $= x(ax + b(1-x)) + (1-x)(cx + d(1-x))$
 $= x f_A + (1-x) f_B$

160

Payoff in Matrix

- Type a' payoff : $f_A = ax + b(1-x)$
- Type b's payoff : $f_B = cx + d(1-x)$

$$Ax = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ 1-x \end{bmatrix} = \begin{bmatrix} ax + b(1-x) \\ cx + d(1-x) \end{bmatrix}$$

$$f_A = (Ax)_1 = ax + b(1-x)$$

$$f_B = (Ax)_2 = cx + d(1-x)$$

158

Two Types

- Average payoff in Population

$$= axx + bx(1-x) + cx(1-x) + d(1-x)(1-x)$$

161

Payoffs for type A and type B

- Type A' payoff : $f_A = ax + b(1-x)$
- Type B's payoff : $f_B = cx + d(1-x)$
- $A \mathbf{x} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ 1-x \end{bmatrix} = \begin{bmatrix} ax + b(1-x) \\ cx + d(1-x) \end{bmatrix}$
- $(A \mathbf{x})_1 = ax + b(1-x)$
- $(A \mathbf{x})_2 = cx + d(1-x)$
- $f_A = (A \mathbf{x})_1 = ax + b(1-x)$
- $f_B = (A \mathbf{x})_2 = cx + d(1-x)$

159

Payoff in Matrix Form

- Two Population , 2 types, $n=2$
- x_i frequency of type i
- $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in S$, State of population
- $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$
- $(A \mathbf{x})_i = a_{i1}x_1 + a_{i2}x_2$, mean payoff of type i
- $\mathbf{x}^T A \mathbf{x}$ mean payoff in population (all tyoes)

162

Replicator equation of n Types

$$\bullet \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\dot{x}_i = x_i((A \mathbf{x})_i - \mathbf{x}^T A \mathbf{x})$$

163

Replicator equation of n Types

$$\bullet \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\dot{x}_i = x_i((A \mathbf{x})_i - \mathbf{x}^T A \mathbf{x})$$

166

n types

- Large Population of n types
- x_i frequency of type i

$$\bullet \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} \in S, \text{ State of population}$$

164

Two types

$$\begin{aligned} \bullet x &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad x^T = [x_1 \quad x_2] \\ \bullet ((A \mathbf{x})_i) &= \left(\begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix} \right)_i = a_{11}x_1 + a_{12}x_2 \\ \bullet \mathbf{x}^T A \mathbf{x} &= [a_{11}x_1 + a_{21}x_2 \quad a_{12}x_1 + a_{22}x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= (a_{11}x_1 + a_{21}x_2)x_1 + (a_{12}x_1 + a_{22}x_2)x_2 \\ &= a_{11}x_1^2 + (a_{21} + a_{12})x_1x_2 + a_{22}x_2^2 \end{aligned}$$

167

Population Setting

- $(A \mathbf{x})_i = a_{i1}x_1 + a_{i2}x_2 \dots + a_{in}x_n$: average payoff of type i
- $\mathbf{x}^T A \mathbf{x}$: mean payoff in population (all types)
- $\frac{\dot{x}_i}{x_i}$: per capita rate of growth

165

Two types

$$\begin{aligned} \bullet \dot{x}_1 &= x_1((A \mathbf{x})_1 - \mathbf{x}^T A \mathbf{x}) \\ &= x_1 ((a_{21}x_1 + a_{22}x_2) - (a_{11}x_1^2 + (a_{21} + a_{12})x_1x_2 + a_{22}x_2^2)) \end{aligned}$$

168

Two types

- $\dot{x}_2 = x_2((A \mathbf{x})_2 - \mathbf{x}^T A \mathbf{x})$
- $= x_1 ((a_{11}x_1 + a_{12}x_2) - (a_{11}x_1^2 + (a_{21} + a_{12})x_1x_2 + a_{22}x_1^2))$

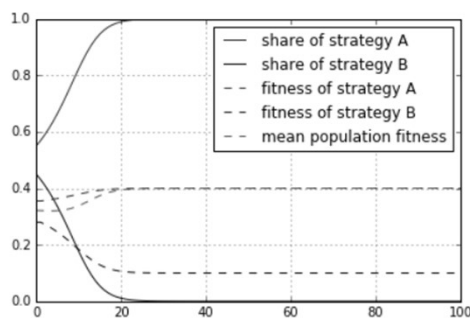
169

Three types

- Average fitness \bar{f} of the population is:
- $\bar{f} = x_R * f_R + x_P * f_P + x_S * f_S$
- The Replicator equations are
- $\frac{dx_R}{dt} = x_R * (f_R - \bar{f})$
- $\frac{dx_P}{dt} = x_P * (f_P - \bar{f})$
- $\frac{dx_S}{dt} = x_S * (f_S - \bar{f})$

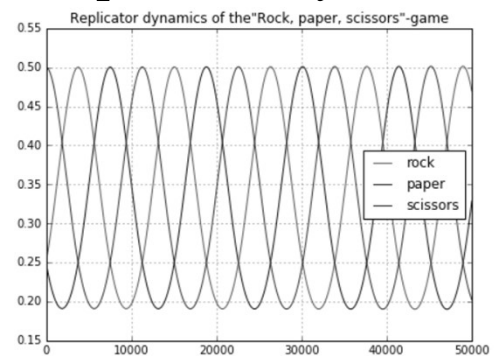
172

Replicator Dynamics



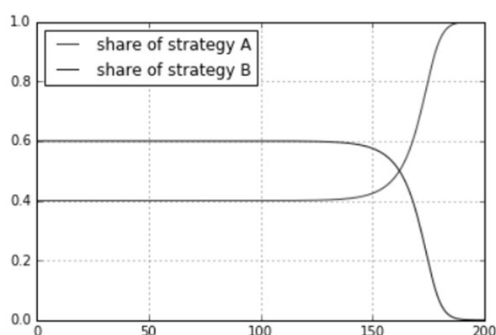
170

Replicator Dynamics



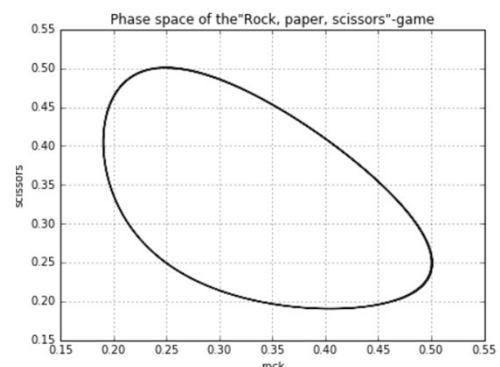
173

Replicator Dynamics



171

Replicator Dynamics



174

Replicator Dynamics

- Three strategies - Rock, paper, scissors
- x_R is the proportion of players playing R , resp. x_P , x_S
- $x_R + x_P + x_S = 1$

175

五彩側斑蜥蜴的博弈

- 在五彩側斑蜥蜴的進化遊戲中，領土、潛行和模仿的石頭剪刀布動態說明了戰略多樣性。

178

Payoff of Player 1 Rock- Paper – Scissors game

	R	P	S
R	0	-1	1
P	1	0	-1
S	-1	1	0

176

五彩側斑蜥蜴的博弈

multicoloured side-blotched lizards

- 多姿多彩的側斑蜥蜴是博弈論原理如何在演化生物學中發揮作用的有趣例子。
- 這些蜥蜴發現於北美洲西部，表現出三種不同的顏色變形：橙色喉雄性、藍喉雄性和黃喉雄性。
- 每個變形體都採用不同的交配策略，這就是賽局理論發揮作用的地方。
- 每個變體的交配策略可以透過著名的「石頭剪刀布」類比來理解

179

Rock- Paper – Scissors game

- Orange Lizard (Rock)
- Yellow Lizard (Paper)
- Blue Lizard (Scissors)

177

orange-throated male side-blotched lizard (Rock)



180

a blue-throated male side-blotched lizard(Scissors)



181

Replicator Dynamics

- Average fitness \bar{f} of the population is:
- $\bar{f} = x_R * f_R + x_P * f_P + x_S * f_S$
- The Replicator equations are
- $\frac{dx_R}{dt} = x_R * (f_R - \bar{f})$
- $\frac{dx_P}{dt} = x_P * (f_P - \bar{f})$
- $\frac{dx_S}{dt} = x_S * (f_S - \bar{f})$

184

a yellow-throated male side-blotched lizard (Paper)



182

Replicator Dynamics

- # simulate
- import matplotlib.pyplot as plt
- %matplotlib inline
- fig = plt.figure(figsize=(15,5))
- fig.add_subplot(1,2,1)
- xR = [0.25]
- xP = [0.25]
- xS = [0.5]
- dt = 0.001

185

Replicator Dynamics

- Starting frequencies are: $x_R=0.25$, $x_P=0.25$, $x_S=0.5$
- Fitness for player playing R is $f_R=0.25*0+0.25*-1+0.5*1 = 0.5$
- Fitness for player playing P is $f_P=0.25*1+0.25*0+0.5*-1 = -0.25$
- Fitness for player playing S is $f_S=0.25*-1+0.25*1+0.5*0 = 0.5$

183

Replicator Dynamics

- for t in range(50000):
- # oscillations
- fR = xR[t] * 0 + xP[t] * -1 + xS[t] * 1
- fP = xR[t] * 1 + xP[t] * 0 + xS[t] * -1
- fS = xR[t] * -1 + xP[t] * 1 + xS[t] * 0
- # asymptotically stable (no cycling)
- #fR = xR[t] * 0 + xP[t] * -1 + xS[t] * 2
- #fP = xR[t] * 2 + xP[t] * 0 + xS[t] * -1
- #fS = xR[t] * -1 + xP[t] * 2 + xS[t] * 0
- f = xR[t] * fR + xP[t] * fP + xS[t] * fS
- xR.append(xR[t] + (xR[t] * (fR - f)) * dt)
- xP.append(xP[t] + (xP[t] * (fP - f)) * dt)
- xS.append(xS[t] + (xS[t] * (fS - f)) * dt)

186

Replicator Dynamics

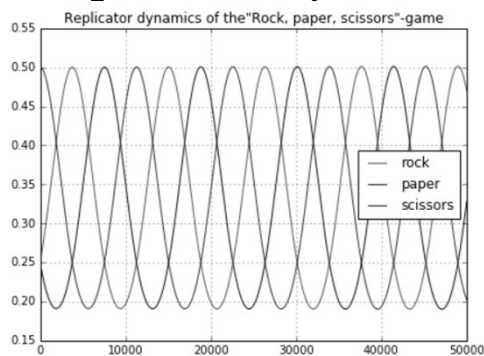
- `plt.plot(xR, 'g', label = 'rock')`
- `plt.plot(xP, 'b', label = 'paper')`
- `plt.plot(xS, 'r', label = 'scissors')`
- `plt.title('Replicator dynamics of the"Rock, paper, scissors"-game')`
- `plt.legend(loc='best')`
- `plt.grid()`
- `fig.add_subplot(1,2,2)`
- `plt.plot(xR, xS)`
- `plt.title('Phase space of the"Rock, paper, scissors"-game')`
- `plt.xlabel('rock')`
- `plt.ylabel('scissors')`
- `plt.grid()`

187

複製動力學應用於 生物學和演化動力學

190

Replicator Dynamics



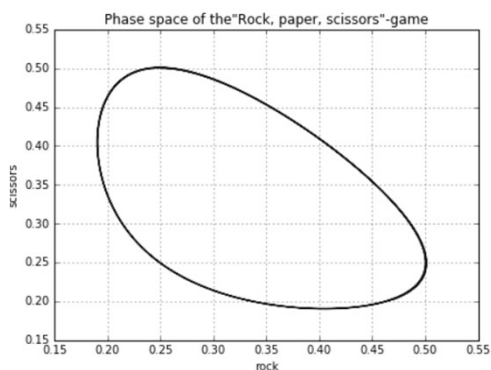
188

生物學和演化動力學

- 物種演化：賦予適應性優勢（生存或繁殖成功）的基因隨著時間的推移頻率會增加，而較不成功的基因則會減少。
- 動物行為：例如，在捕食者-被捕食者模型中，合作、攻擊或迴避等策略可以根據群體的生存成功而演化。

191

Replicator Dynamics



189

複製動力學應用於 人工智慧和機器學習

192

人工智慧和機器學習

- 遺傳演算法：解決方案或“策略”被編碼為“基因”，那些在解決問題方面表現良好的解決方案或“策略”被選擇並在下一代中複製，模仿生物進化。

193

迷因 meme

- 迷因是在線上共享的概念、笑話或媒體片段（如圖像、影片或文字），通常會在社交媒體和網路上快速傳播。
- 迷因通常是幽默的、相關的或發人深省的，並且它們通常依賴共同的文化或社會理解。人們使用迷因來評論時事、取笑相關情況，或以簡單、視覺的方式表達想法和感受。
- 隨著人們添加新的標題或進行更改，模因也會不斷演變，賦予它們新的含義並保持它們在不同環境中的相關性

196

複製動力學應用於 社會學與文化演化

194

兩種雌性策略

197

社會學與文化演化

- 思想傳播：提供社會優勢（例如合作、信任）或被視為對社會有益的想法或行為被更廣泛地採用，而不太成功的想法則逐漸消失。
- 文化實踐：隨著時間的推移，某些文化規範或實踐可能佔據主導地位，特別是如果它們被視為提高社會中的群體凝聚力或個人成功。
- 社群網路：迷因、趨勢或意識形態如何透過社群媒體或社群傳播的動態可以使用複製動態進行建模，成功的想法會激增，而其他想法則會消失。

195

兩種雌性策略

- 假設兩種雌性策略：羞怯和快速 coy and fast
- 以及在這種情況下，害羞的雌性不會與雄性交配，直到他經歷了漫長而昂貴的求愛期。
- 快速的雌性會立即與任何雄性交配。

198

兩種雄性策略

- 兩種雄性策略：忠誠和花心faithful and philanderer。
- 忠誠的雄性準備好長時間求愛，交配後，留在雌性身邊並幫助她撫養幼崽。
- 如果雌性不立即與花心的雄性交配，花心的雄性很快就會失去耐心，並且在交配後，它們也不會留下來。

199

回報的計算

	忠誠 Faithful q	花心 Philanderer (1-q)
靦腆 Coy (p)	(15-10-3, 15-10-3) (2, 2)	(0, 0)
快速 Fast (1-p)	(15-10, 15-10) (5, 5)	(15-20, 15) (-5, 15)

202

回報的計算

- 假設每個養育的幼崽都是+15分。
- 養育一個幼崽需要20分，但如果共同的話，則只需10分。浪費時間為 -3。
- 在雌性靦腆而雄性忠誠的情況下，回報將是 $15 - 10 - 3 = +2$ 。
- 在快速雌性的情況下，回報將為 $15 - 10 = +5$ 。

200

混合策略的計算

- $2q = 5q + (-5)(1-q)$
- $8q = 5$
- $q = 5/8$
- $2p + 5(1-p) = 15(1-p)$
- $12p = 10$
- $p = 5/6$
- 這是一種無止盡的振盪；
- 系統將趨於穩定狀態，即 $5/6$ 的女性靦腆， $5/8$ 的男性忠誠。

203

回報的計算

- 如果一隻快速的雌性開始佔據主導地位，那麼花心的雄性就會在種群中出現，並且比它忠誠的競爭對手做得更好。
- 快速的雌性的回報 = $15 - 20 = -5$ ，而花心的雄性的回報 +15。

201

Genetic Algorithm

204

Steps of a Genetic Algorithm

- **Initialization:** Generate an initial population randomly.
- **Fitness Evaluation:** Compute the fitness of each individual.
- **Selection:** Choose individuals based on their fitness.
- **Crossover:** Produce offspring by recombining pairs of selected parents.
- **Mutation:** Apply random changes to offspring genes.
- **Replacement:** Form the next generation from the offspring and/or parents.
- **Repeat:** Steps 2–6 until a termination condition is met.

205

3. Selection (Roulette Wheel Selection)

- Since all individuals have the same fitness, the probability of selecting any individual is equal. We randomly choose two pairs of individuals for crossover.
- Let's say we select X_1 and X_2 as one pair, and X_3 and X_4 as the second pair.

208

1. Initialization

- We start with a randomly generated population of 4 individuals (binary strings of length 6):
- $P^{(0)} = \{X_1^{(0)} = 010011, X_2^{(0)} = 101010, X_3^{(0)} = 111000, X_4^{(0)} = 000111\}$

206

4. Crossover (Single-Point)

- For the first pair $X_1 = 010011$ and $X_2 = 101010$,
- assume we choose a crossover point at position 3:
- Parent 1: 010011
- Parent 2: 101010
- Performing single-point crossover at position 3:
- Offspring 1: 010010
- Offspring 2: 101011
- For the second pair $X_3 = 111000$ and $X_4 = 000111$,
- assume we choose a crossover point at position 2:
- Parent 1: 111000
- Parent 2: 000111
- Performing single-point crossover at position 2:
- Offspring 1: 110111
- Offspring 2: 001000
- Now the new population (offspring) after crossover is
- $O^{(0)} = \{O_1 = 010010, O_2 = 101011, O_3 = 110111, O_4 = 001000\}$

209

2. Fitness Function

- The fitness function $f(X)$ counts the number of 1s in the binary string.
 $f(010011) = 3, f(101010) = 3, f(111000) = 3, f(000111) = 3$
- Since all individuals have the same fitness, selection is uniform in the first generation.

207

5. Mutation

- Let's assume a mutation probability of 0.1. We randomly mutate one bit of one offspring.
- Suppose we mutate the 4th bit of $O_4 = 001000$, flipping it from 0 to 1.
- The new string becomes: $O_4 = 001100$
- The new population after mutation is:
- $O^{(0)} = \{O_1 = 010010, O_2 = 101011, O_3 = 110111, O_4 = 001100\}$

210

6. Fitness Evaluation

- Now, we evaluate the fitness of the new population.
 $f(010010)=3, f(101011)=4, f(110111)=5, f(001100)=2$
- The best solution is ,
 $O_3=110111$ with a fitness of 5.

211

Summary of Evolutionary Process

- **Initial Population:**
 - $P^{(0)}=\{X_1^{(0)}=010011, X_2^{(0)}=101010, X_3^{(0)}=111000, X_4^{(0)}=000111\}$
- **First Generation (After Crossover and Mutation)**
 - $P^{(1)}=\{010010, 101011, 110111, 001100\}$
- **Next Generations:** The process continues until we reach the optimal solution.
- This illustrates how a Genetic Algorithm evolves a population of candidate solutions over generations to find the optimal solution. Each generation brings us closer to the best solution by leveraging crossover, mutation, and selection mechanisms.

214

7. Replacement

- We replace the old population with the new one.
- The population for the next generation is
- $P^{(1)}=\{010010, 101011, 110111, 001100\}$

212

Example Applications

- Optimization problems (e.g., scheduling, routing)
- Machine learning (e.g., neural network optimization)
- Engineering design (e.g., evolving structural designs)
- Game strategy development

215

8. Termination Condition

- We continue this process for a fixed number of generations or until we find the string with all 1s (fitness = 6).
- Suppose after a few more generations, we evolve the solution 111111, which is the optimal solution.

213

Reinforcement Learning (RL)

- Reinforcement Learning (RL) is a type of machine learning where an agent learns to make decisions by performing actions in an environment to achieve some goal. It is based on a reward system, where the agent receives feedback in the form of rewards or penalties based on its actions, guiding it to learn the best strategies over time.

216

Reinforcement Learning (RL)

- 1.Agent: This is the learner or decision-maker, like a robot or an AI algorithm.
- 2.Environment: The world in which the agent operates. For example, a video game, a robot's surroundings, or a simulated business model.
- 3. Actions: Choices the agent can make. For example, moving left or right, picking up an object, etc.

217

a simple example of a robot learning to navigate a maze to reach a goal using reinforcement learning (RL).

220

Reinforcement Learning (RL)

- 4.States: The situations the agent finds itself in at any point in time within the environment.
- 5.Reward: Feedback the agent gets after taking an action in a particular state. Positive rewards reinforce good actions, while negative rewards (or penalties) discourage bad actions.
- 6. Policy: The strategy or set of rules the agent follows to decide what action to take based on its current state.

218

情境

- 假設一個機器人在一個網格迷宮中學習如何找到通往目標的路徑。機器人（代理）透過與環境的互動，學習一套能最大化累計獎勵的策略。以下是它的工作原理

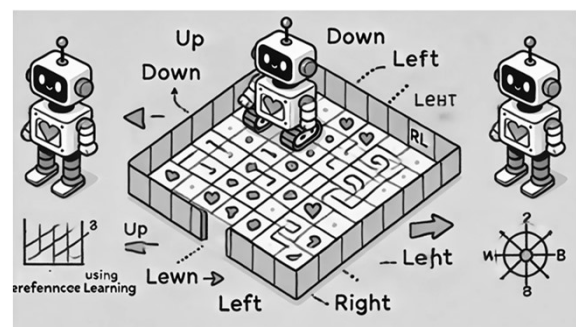
221

Reinforcement Learning (RL)

- The agent's goal in reinforcement learning is to learn a policy that maximizes the total cumulative reward over time, allowing it to perform tasks more effectively and efficiently.
- A common analogy is training a pet: the pet (agent) learns to perform tricks (actions) by receiving treats (rewards) when it performs the correct action. Over time, it learns the best tricks to get the most treats.

219

a robot navigating a maze using reinforcement learning



222

步驟分析

- 1. 定義環境
 - 狀態：迷宮中的每個網格單元都是一個狀態。機器人的當前位置即為它目前的狀態。
 - 動作：機器人在每一步可以選擇四個動作：上移、下移、左移或右移。
 - 獎勵：到達目標點時獲得+10分。每移動一步扣除-1分，以鼓勵找到最短路徑。若撞到障礙物則扣除-10分。

223

步驟分析

- 4. 選擇動作（探索與利用）
 - 在每一步，機器人決定是：
 - 探索（隨機選擇一個動作，以便更多了解迷宮）。
 - 利用（根據當前經驗，選擇當前狀態Q值最高的動作）。探索和利用之間的平衡由一個參數（如 ϵ ）控制，並隨著機器人的經驗逐漸減少。

226

步驟分析

- 2. 初始化Q表（Q學習）
 - Q表是一個矩陣，其中行表示狀態（網格單元），列表示動作（上、下、左、右）。每個元素 $Q(s, a)$ 表示在狀態 s 下採取動作 a 的估計價值。
 - 初始時，所有Q值都設為零，因為機器人還沒有經驗

224

步驟分析

- 執行動作並觀察結果
 - 機器人執行所選動作（例如向左移動），並進入一個新狀態（在迷宮中的新位置）。它會根據結果獲得一個獎勵（例如未到達目標則扣-1分，撞到障礙物則扣-10分，到達目標則加+10分）

227

步驟分析

- 3. 交互循環（回合）
 - 機器人從迷宮中的隨機位置開始，並重複此過程直到到達目標。此循環會運行多次，以便機器人持續學習。

225

步驟分析

- 6. 更新Q值（學習）
 - 根據獲得的獎勵和最佳未來動作的估計值，更新狀態-動作對 (s, a) 的Q值，遵循Q學習公式：

$$Q(s, a) = Q(s, a) + \alpha \cdot (\text{獎勵} + \gamma \cdot \max_a Q(s', a) - Q(s, a))$$
 其中：
 - α 是學習率（機器人更新知識的速度），
 - γ 是折扣因子（未來獎勵的權重），
 - s' 是執行動作 a 後的新狀態。

228

步驟分析

- 7. 重複步驟4 - 6，直到到達目標或超過最大步數。機器人繼續執行動作、更新Q值並從結果中學習。一旦到達目標或超過該回合的最大步數，回合結束，開始新的一輪。

229

Further Reading

232

步驟分析

- 8. 多輪學習（回合）
— 隨著每個回合的進行，機器人不斷改進Q值，逐步偏向能夠以最少步數達到目標的動作。經過多輪回合後，Q表會收斂到一套指引機器人高效率到達目標的策略。

230

複製動力學應用於選舉

233

步驟分析

- 9. 提取最終策略
— 經過充分的訓練，Q表的數值反映了最佳策略。機器人現在可以在每個狀態下選擇Q值最高的動作，引導它沿著最短路徑到達目標，並避開障礙。

231

選舉

- 在兩黨制度的背景下，複製動態可以解釋選民偏好如何在兩個主導政黨之間轉變。假設甲方由於一個受歡迎的政綱platform而贏得了幾次選舉。
- 隨著時間的推移，更多選民可能會支持甲方，從而增強其主導地位。如果乙方採用類似的政綱platform進行競爭，就會產生一種動態，雙方根據對方的成功來制定自己的策略。

234

複製動力學應用於 商業策略和創新擴散

235

經濟與市場競爭

- 企業競爭：擁有卓越策略（例如更好的定價模式、更高的產品品質）的成功企業會增加市場份額，而實力較弱的企業或退出市場。
- 消費者選擇：消費者傾向於複製成功的購買決策，因此在市場上表現良好的產品會隨著時間的推移而越來越受歡迎
- 創新：優於其他技術或商業模式的新技術或商業模式將逐漸取代舊方法，不適應的企業將被淘汰。

238

商業策略和創新擴散

- 創新擴散：市場上成功的新技術、產品或商業模式往往會隨著其他公司的採用而複製。例如，電動車或數位支付系統的傳播可以使用複製動力學進行建模
- 企業進化：隨著時間的推移，能夠很好地適應不斷變化的環境（例如新法規、市場條件、消費者偏好）的公司會複製他們的成功，而那些沒有進化的公司則面臨衰退。

236

複製動力學應用於 語言與語言學

239

複製動力學應用於 經濟與市場競爭

237

語言與語言學

- 語言競爭：為使用者提供社會或經濟利益的語言或方言往往會傳播，而其他語言或方言可能會下降。這可以解釋為什麼英語等全球主導語言能夠迅速傳播。
- 語言特徵：如果某些文法結構或語音特徵更容易學習或提供溝通優勢，它們可能會在語言社群中變得更加常見。

240

複製動力學應用於 政治學與意識形態傳播

241

金融與投資策略

- 投資行為：投資人可能會模仿成功的策略，隨著這些策略產生正回報，更多的投資人會複製它們，從而推動市場行為的轉變。
- 金融市場：對沖基金和演算法交易者經常調整和複製成功的交易策略，導致某些主導市場行為的出現。

244

政治學與意識形態傳播

- 政治策略：與民眾產生共鳴或成功實現政策目標的政黨或運動將獲得更多追隨者，而較不成功的運動則會衰退。
- 社會運動：複製動態可以解釋某些社會運動（例如民權或環保主義）如何隨著時間的推移而傳播，取得成功並吸引更多支持者。

242

複製動力學應用於 生態與生態系穩定性

245

複製動力學應用於 金融與投資策略

243

生態與生態系穩定性

- 物種競爭：在生態系統中，更適合其環境（例如，更有效地利用資源、更好地適應氣候）的物種將蓬勃發展和傳播，而其他物種則會衰落。
- 資源分配：在資源有限的環境中，制定策略以更好地獲取或利用資源的生物體將在競爭中勝出，從而導致進化轉變。

246

複製動力學應用於 行為生態學與合作

247

Network Game Theory

- https://www.youtube.com/watch?v=QY_4LXI4ExY
- Evolutionary Game Theory
- <https://www.youtube.com/watch?v=HxgVYhhArSk>
- *Cooperative Structures*
- <https://www.youtube.com/watch?v=53z-KK26yd8>

250

行為生態學與合作

- 公共物品：在個人為共享資源做出貢獻的情況下，即使面臨搭便車問題，如果能夠提供整體優勢，有利於群體的合作策略也可以傳播。
- 利他主義：在演化生物學中，如果群體或親屬的利益大於個人的成本，像利他主義這樣的行為就會演化，隨著時間的推移，這些行為可以在群體中複製。

248

- How multicoloured side-blotched lizards put game theory into evolutionary action
- <https://aeon.co/videos/how-multicoloured-side-blotched-lizards-put-game-theory-into-evolutionary-action>

251

UBC ISCI 344 - Evolutionary Game Theory

- <https://www.youtube.com/watch?v=NIFhapYWdxU>

249