Cooperation

Emergence in complex systems

Julien Lie-Panis[1,2]

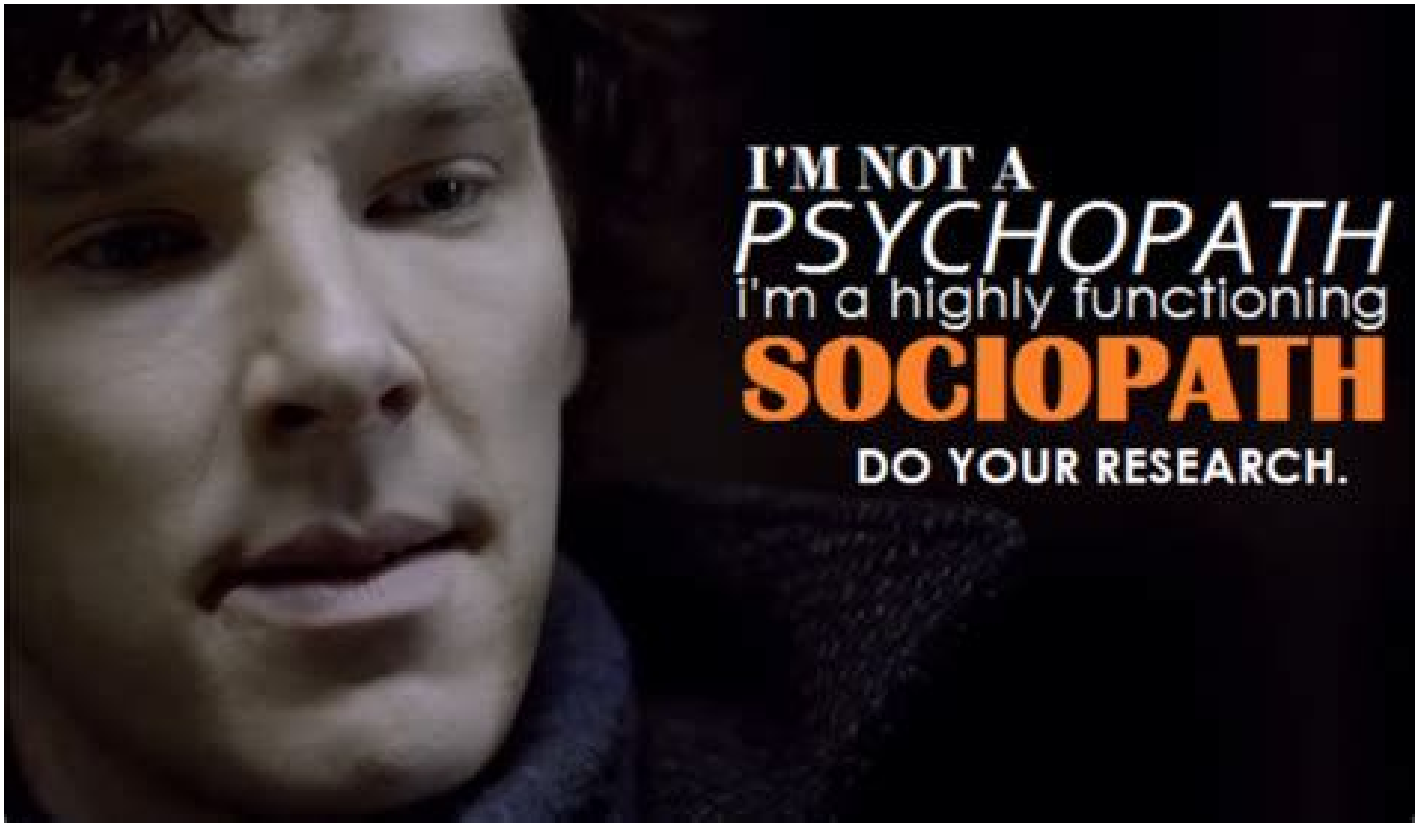
[1]Laboratoire Traitement et Communication de l'Information
Telecom Paris

[2]Institut Jean Nicod
Ecole Normale Superieure

November 17, 2021

DEC

Why altruism?



Economist: :(
(Game Theory // Nash equilibrium)

Biologist: :) (?)
(Evolutionary Game Theory // ESS, which resist *mutant* strategies)

Golden Balls

The Prisoner's dilemma

Prisoner B \ Prisoner A	C	D
C	$(-1,-1)$	$(-20,0)$
D	$(0,-20)$	$(-10,-10)$

Years in prison faced by both prisoner's depending on their actions (cooperate or defect on the other)

The Prisoner's dilemma

Prisoner B \ Prisoner A	C	D
C	$(-1,-1)$	$(-20,0)$
D	$(0,-20)$	$(-10,-10)$

Years in prison faced by both prisoner's depending on their actions (cooperate or defect on the other)

The Prisoner's dilemma

Prisoner B \ Prisoner A	C	D
C	$(-1,-1)$	$(-20,0)$
D	$(0,-20)$	$(-10,-10)$

Years in prison faced by both prisoner's depending on their actions (cooperate or defect on the other)

The Prisoner's dilemma

Prisoner B \ Prisoner A	C	D
C	$(-1,-1)$	$(-20,0)$
D	$(0,-20)$	$(-10,-10)$

Years in prison faced by both prisoner's depending on their actions (cooperate or defect on the other)

The Prisoner's dilemma

Prisoner B \ Prisoner A	C	D
C	$(-1,-1)$	$(-20,0)$
D	$(0,-20)$	$(-10,-10)$

Years in prison faced by both prisoner's depending on their actions (cooperate or defect on the other)

$\rightarrow$ *The only ESS for the prisoner's dilemma is* (D, D)

The Prisoner's dilemma

Benefit of help: b
Cost of helping someone else: c
Benefit of mutual cooperation: $b - c > 0$

		Chooser	
		A	R
Signaler	C	$(b-c, b-c)$	$(-c, b)$
	D	$(b, -c)$	$(0, 0)$

Payoffs for a typical cooperative dilemma

$\rightarrow$ *The only ESS for the prisoner's dilemma is* (D, D)

So why do we manage to cooperate ?

https://ncase.me/trust/

https://ncase.me/trust/

One game: the social optimum (C,C) is not reached. What if games are *repeated* with a certain probability?

https://ncase.me/trust/

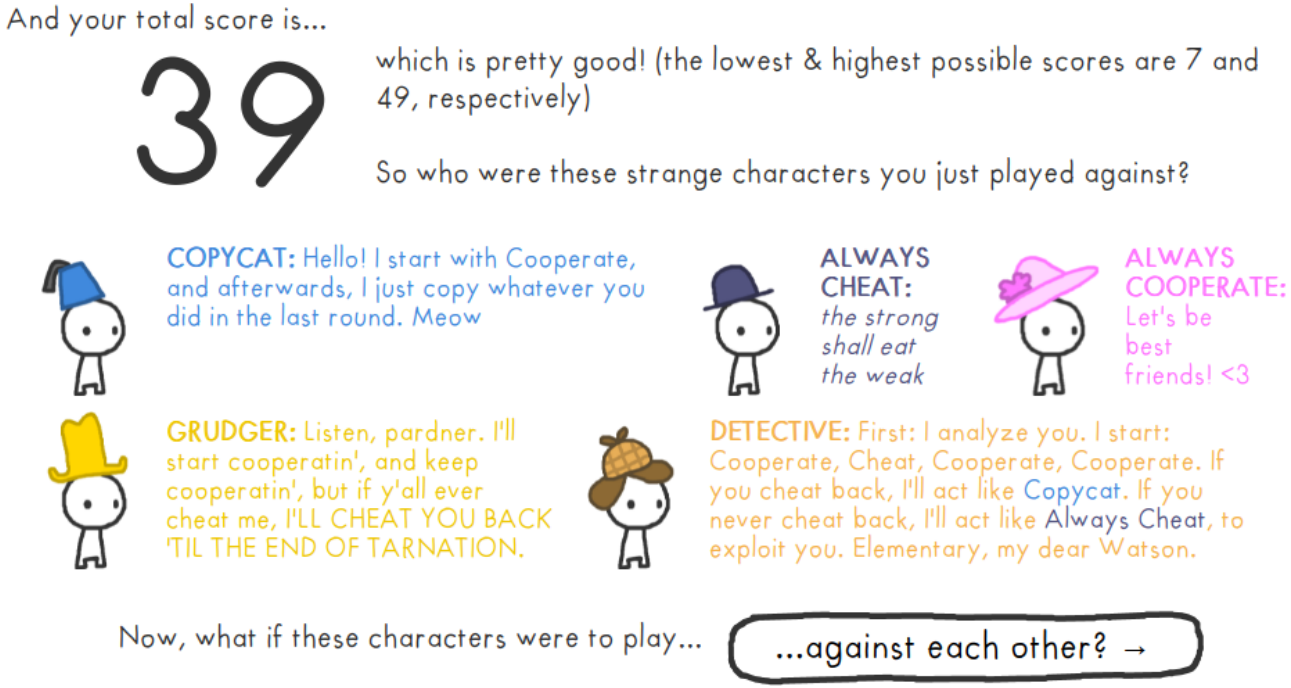
One game: the social optimum (C,C) is not reached. What if games are *repeated* with a certain probability?

Online: think of your strategy, and play the game against the computer. Note your score.

https://ncase.me/trust/

One game: the social optimum (C,C) is not reached. What if games are *repeated* with a certain probability?

Online: think of your strategy, and play the game against the computer. Note your score.



Tournament

What about when players enter in indefinitely repeated interactions (probability of ending μ)? How do you think will win the "tournament":

A) A pure defector, who always plays D — and thus exploits others' generosity

B) A pure cooperator, who always plays C — and thus benefits from repetition when encountering other cooperators (as they gain $b - c$ for several rounds)

C) "Tit-for-tat": a player who starts of by playing C and then just repeats the other players' last interaction.

D) Some other more complex strategy, which figures out others' player strategy in the first few rounds, using a neural network

Tournament

What about when players enter in indefinitely repeated interactions (probability of ending μ)? How do you think will win the "tournament":

A) A pure defector, who always plays D — and thus exploits others' generosity

B) A pure cooperator, who always plays C — and thus benefits from repetition when encountering other cooperators (as they gain $b - c$ for several rounds)

C) **"Tit-for-tat": a player who starts of by playing C and then just repeats the other players' last interaction.**

D) Some other more complex strategy, which figures out others' player strategy in the first few rounds, using a neural network

Reciprocal cooperation

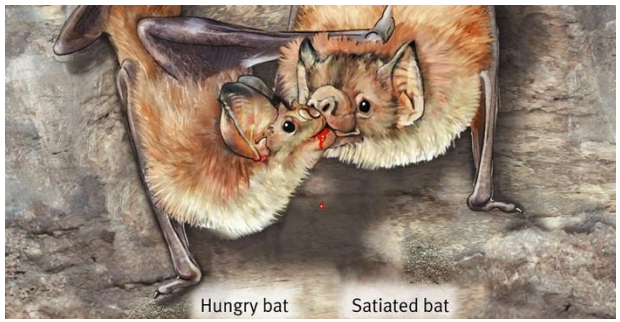
Indefinite repetition can allow reciprocal cooperation to emerge ("do unto others as you would have do unto you") — as shown by an actual tournament organized by R. Axelrod [Trivers, 1971, Axelrod and Hamilton, 1981].

Reciprocal cooperation

Indefinite repetition can allow reciprocal cooperation to emerge ("do unto others as you would have do unto you") — as shown by an actual tournament organized by R. Axelrod [Trivers, 1971, Axelrod and Hamilton, 1981].

Refinement: indirect reciprocity and reputation
[Nowak and Sigmund, 1998, Panchanathan and Boyd, 2003]

Non-human animal examples:



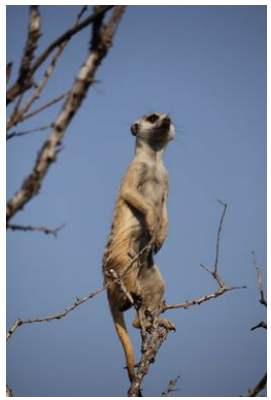
Explanations for helpful behavior

These examples are quite rare! In most instances, helpful acts can be explained by a simpler mechanism.

Explanations for helpful behavior

These examples are quite rare! In most instances, helpful acts can be explained by a simpler mechanism.

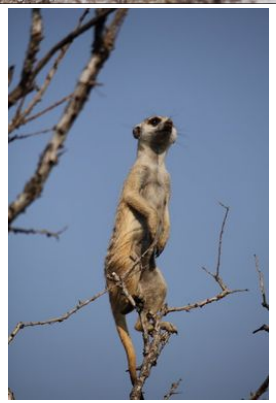
What explanation would you invoke for the following behavior?



Explanations for helpful behavior

Explanations seen up until now:

Kin altruism: $-C_i$, $+B_o$ ($rB_o > C_i$)



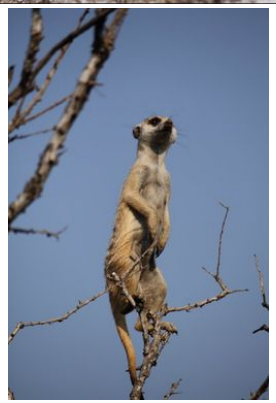
Mutualism or by-product cooperation: $+ B_i$, $+B_o$

Reciprocal cooperation: $+B_i$ because help is recip., $+B_o$.

Explanations for helpful behavior

Explanations seen up until now:

Kin altruism: $-C_i$, $+B_o$ ($rB_o > C_i$)



Mutualism or by-product cooperation: $+ B_i$, $+B_o$

Reciprocal ~~**altruism**~~ cooperation: $+B_i$ because help is recip., $+B_o$.

Refinements and limits

None of these explanations seem able to address the extent of human cooperation!

Refinements and limits

None of these explanations seem able to address the extent of human cooperation!

- We help unrelated individuals
- Including strangers we will likely never meet again
- We take into account mistakes and/or attenuating circumstances
- We don't value efficiency [Burum et al., 2020]
- We do not keep track of favours with certain partners [Hoffman et al., 2015]
- We sometimes obscure our good deeds [Hoffman et al., 2018]
- We value and trust others more when they appear uncalculating, most or uninterested in material gains
- We don't simply trust based on behavior in a similar game, or reciprocate help with help, but instead make a wide-series of cross-inferences, and hold a "general" reputation

Refinements and limits

None of these explanations seem able to address the extent of human cooperation!

- We help unrelated individuals
- Including strangers we will likely never meet again
- We take into account mistakes and/or attenuating circumstances
- We don't value efficiency [Burum et al., 2020]
- We do not keep track of favours with certain partners [Hoffman et al., 2015]
- We sometimes obscure our good deeds [Hoffman et al., 2018]
- We value and trust others more when they appear uncalculating, most or uninterested in material gains
- We don't simply trust based on behavior in a similar game, or reciprocate help with help, but instead make a wide-series of cross-inferences, and hold a "general" reputation

Next session: signaling.

Refinements and limits

None of these explanations seem able to address the extent of human cooperation!

- We help unrelated individuals
- Including strangers we will likely never meet again
- We take into account mistakes and/or attenuating circumstances
- We don't value efficiency [Burum et al., 2020]
- We do not keep track of favours with certain partners [Hoffman et al., 2015]
- We sometimes obscure our good deeds [Hoffman et al., 2018]
- We value and trust others more when they appear uncalculating, most or uninterested in material gains
- We don't simply trust based on behavior in a similar game, or reciprocate help with help, but instead make a wide-series of cross-inferences, and hold a "general" reputation

Next session: signaling. NB: Actually Tit for Tat is not an ESS!

References I


Axelrod, R. and Hamilton, W. D. (1981).
The Evolution of Cooperation.
page 11.


Burum, B., Nowak, M. A., and Hoffman, M. (2020).
An evolutionary explanation for ineffective altruism.
Nature Human Behaviour, pages 1–13.
Publisher: Nature Publishing Group.


Hoffman, M., Hilbe, C., and Nowak, M. A. (2018).
The signal-burying game can explain why we obscure positive traits and good deeds.
Nature Human Behaviour, 2(6):397–404.
Number: 6 Publisher: Nature Publishing Group.

References II


Hoffman, M., Yoeli, E., and Nowak, M. A. (2015).
Cooperate without looking: Why we care what people think and not just what they do.
Proceedings of the National Academy of Sciences, 112(6):1727–1732.


Nowak, M. A. and Sigmund, K. (1998).
Evolution of indirect reciprocity by image scoring.
Nature, 393(6685):573–577.
Number: 6685 Publisher: Nature Publishing Group.


Panchanathan, K. and Boyd, R. (2003).
A tale of two defectors: the importance of standing for evolution of indirect reciprocity.
Journal of Theoretical Biology, 224(1):115–126.

References III


Trivers, R. L. (1971).
The Evolution of Reciprocal Altruism.
The Quarterly Review of Biology, 46(1):35–57.