

Reciprocal Altruism

Lecture 11

Psych 250

9/27/07

Today

- Reciprocal Altruism
- Prisoner's Dilemma
- ESS for reciprocity
- Cognitive programs required
 - Reasoning about social exchange

Adaptive Problem



Cognitive Programs



Neurophysiology

Altruism: helping behavior directed toward non-relatives

What kinds of programs would have been needed to guide social exchange?

Are there brain regions that govern the ability to reason about social exchange?

Reciprocal Altruism

- Basic Idea:
 - I will help you now and later, when I need help, you will help me.
- But, there is a problem:

- Cheating/defection is a big problem for the evolution of reciprocal altruism.
- Prisoner's Dilemma

The Prisoner's Dilemma

| | | PARTNER | |
|-----|-----------|-------------------------------------------------------------------------|----------------------------------------------------------------------------|
| | | Cooperate | Defect |
| YOU | Cooperate | You: 6mths Partner: 6mths (reward for mutual cooperation) | You: 20 yrs Partner: 0 (sucker's payoff) |
| | Defect | You: 0 Partner: 20 yrs (temptation to defect) | You: 10 yrs Partner: 10yrs (punishment for mutual defection) |

Cooperate = stay silent
Defect = confess

Prisoner's Dilemma

- What should you do in a 1-shot Prisoner's Dilemma?
 - What is the best strategy? ESS?
- How does this apply to the real world?
- Iterated Prisoner's Dilemma
 - Interact with one person more than once
- What kind of strategy would work?

Evolutionary Stable Strategy

- An evolutionary stable strategy (ESS):
 - A strategy (way of making decisions) that, at any frequency in the population, cannot be invaded by any other strategy -- it is the one that works the best.
- Defection can be seen as an ESS
 - It beats out the cooperative strategy.
- Is there any way that cooperation can evolve given the stability of an “always defect” strategy?

Is it possible for niceness to evolve?

- The answer is, of course, yes.
 - When played only once, however, “always defect” is the most stable strategy
- If the Prisoner’s Dilemma is played more than once, and both players do not know how many times it will be played, “always defect” is no longer the most stable strategy.
- What strategy can out compete “always defect”?

Tit for Tat

- Features of tit for tat:
 - C = cooperate
 - D = Defect

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------------|---|---|---|---|---|---|---|---|
| TFT | | | | | | | | |
| Strategy X | D | C | D | D | C | C | D | C |

Properties of a Tit for Tat strategy

- Tit for Tat is a form of cooperation
 - Based on reciprocity!
- It out competed all the other types because of two basic features:
 - It was ***retaliatory***, which discouraged the opponent from continuing a defection strategy.
 - It was ***forgiving***. After any number of defections, as soon as the opponent cooperated, TFT forgave by cooperating on the next trial.

- TFT is an evolutionarily stable strategy so long as the probability that two individuals will meet again is sufficiently large and unknown.
- (example of paying bill in bar at airport)

What cognitive programs do you need for this ability?

To be able to retaliate against defectors and to rekindle reciprocal relations gone bad, you need to be able to:

- 1)
- 2)
- 3)

Specialized mechanisms dedicated to recognizing individuals.

- Strong evidence that there are neural circuits dedicated to recognizing faces.
- Prosopagnosia -- a neuropsychological disorder where individuals lose the ability to recognize faces.

Person-specific memory

- Important to have information about whether individuals are cooperators or cheaters.
- Evidence that we store person-specific memories.
 - When asked “Is your roommate nice?”
- Calls up:
 - A person summary of either yes or no
 - (semantic information)
 - instances when your roommate was NOT nice
 - (episodic memory)
- Bounded information

Memory for cheaters

- Also some evidence that people tend to remember individuals who are cheaters more easily than individuals who are not.
- “Linda experiment”

Need to be able to reason about cheating

- Let's take a closer look at what it means to reason about something.
- What are some of the things we reason about?
- We are able to do these things effortlessly.
 - Reasoning abilities make certain kinds of tasks easy to do.

What kind of reasoning mechanisms should we possess?

- General purpose reasoning abilities or specialized reasoning abilities?
- Should have reasoning mechanisms that solve adaptive problems that our ancestors faced over evolutionary history.
- The ability to detect cheaters was one of the adaptive problems we faced.
 - Therefore expect to see circuits that allow us to reason about cheating and other aspects of social exchange.

Properties that you would expect to see in a well designed reasoning instinct:

- Specialized for solving an adaptive problem
- Reliably develops in all human beings
- Develops without conscious effort
- Develops without any formal instruction
- Applied without awareness of underlying logic
- Distinct from more general abilities to process information or behave intelligently

Classical view of the mind

- General purpose reasoning abilities.
- There are very few reasoning processes.
- They are content independent.
 - This means that the same general reasoning processes are used to solve *all* reasoning problems.
 - Two such reasoning processes are:

Modus Ponens

Premise: **If P then Q**

**If the bathroom door is shut,
then someone is in there**

Given Premise: **P**

Can conclude: **Q**

Modus Tollens

Premise: **If P then Q**

**If you sit in the last row,
then you will fall asleep**

Given Premise: $\sim Q$

Can conclude: $\sim P$

Classical view of the mind

- Under the classical view of the mind, these reasoning processes were *general* and could be applied to *anything*.
- They seemed very powerful and it was nice to think that the laws of thinking were governed by the laws of logic.
- So, if our minds contained reasoning inferences that obeyed the laws of logic, we would be able to check the validity of statements using:
 - modus ponens
 - modus tollens

How to falsify a logical statement

- According to the rules of logic, what would you need to do to check the *validity* of a statement?
- Conversely, what would have to do to *falsify* a logical statement?
- How good are these content independent reasoning inferences at solving problems?

WASON SELECTION TASK

If P then Q:

If the card has a 7 on one side,
then there is an E on the other side.

Which card(s) do you need to turn over?

7

3

E

S

Rule:

If you are drinking beer, then you must be over 21.

Which people do you have to check?

Beer

Pepsi

23

18

Detecting logic violations

- According to formal logic, to determine whether the rule has been violated:

Question:

Why was this problem much easier compared to the first descriptive problem?

Answer:

We have specialized circuits that evolved to reason about social exchange.

Cheating

- The evolution of reciprocal altruism required that individuals be able to detect cheaters.
- What is cheating in a social exchange?
 - Taking the benefit without paying the associated cost or meeting the requirement.
 - Adaptive logic

Features of a cheater detection system

- Shouldn't need to be familiar with the content
- Should use the rules of adaptive logic not necessarily formal logic
- Should be sensitive to perspective
- Should be sensitive to costs and benefits
- Should be sensitive to intentional cheating versus innocent mistakes
- Should be able to detect cheaters not altruists