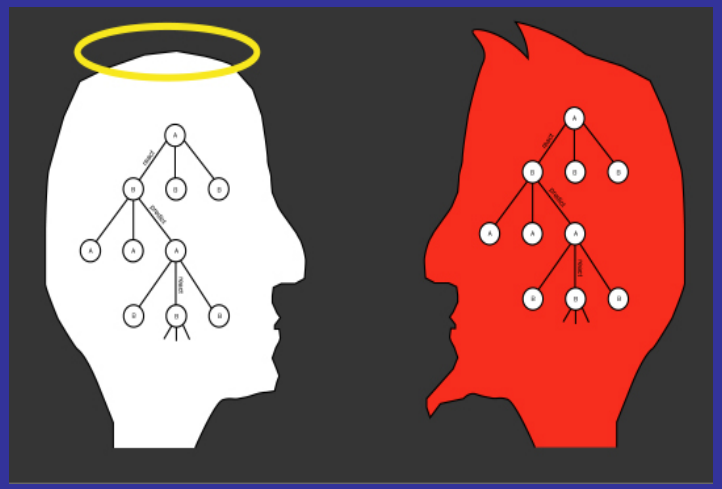
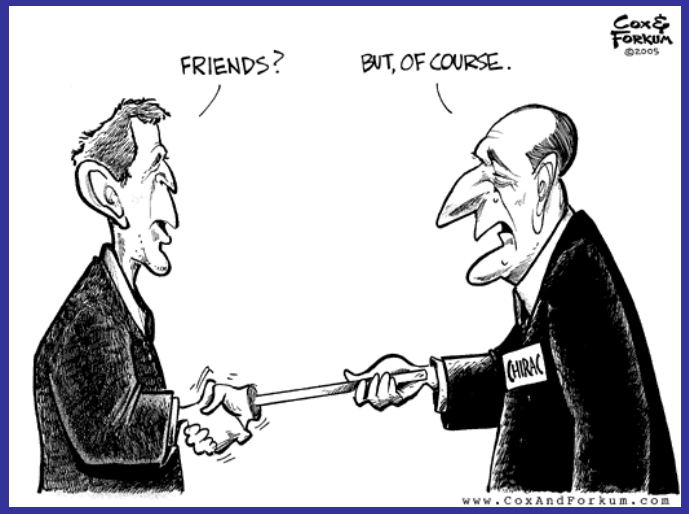
I believe it when I see it - Reputation versus Experience

Eva van den Broek
ex-CAI 1997-2003

game theory - cooperation – fish simulation - experiment – discussion

Outline





game theory - cooperation – fish simulation - experiment – discussion

Beauty contest

1. **Everybody chooses secretly an integer number between 1 and 100 (inclusive).**
2. **The average of all chosen numbers is calculated.**
3. **This average is multiplied by 0.7 and then rounded to the closest integer.**
4. **The winner is the number that is closest to the number calculated in 3.**

Game theory

- Strategies incorporate what ‘the other’ does
- Predicts equilibria, where no-one can do better by changing their strategy
- Does, however, NOT predict behavior well

Problem of Cooperation

How can cooperation evolve in a world of rational agents?

- Everyone is better of by defecting:
  - Biologists: fitness of a defector is higher
  - Economists: ‘utility’ of a defector is higher
- Any conjectures?

⇩

Conjectures

- Longer time horizon
- Kin selection (rB – C > 0)
- Direct reciprocity (Tit for Tat)
- Indirect reciprocity
  - group selection?
  - punishing?
  - reputation

Reputation mechanism

- Image scoring (Nowak and Sigmund, 1998)

- 100 players interact randomly
- cost < benefit, score = 0, strategy <-5,5>
- if coop last round, score +1
- next generation according to payoff
- If strategy = 0, “discriminating”

Evidence for image scoring

- People use image scoring (Wedekind and Milinski, 2000)
- People make strategic choices

  (Engelmann and Fischbacher, 2004)
- keep their image score above a threshold (Seinen and Schram, 2001)

Problem solved?

no.

Possibly wrong assumptions needed for image scoring:

1. All exchanges are alike
2. People react to ‘image’ like ‘experience’

Hypothesis

Reputation depends on homogeneity of population

→ Introducing heterogeneity in value of interactions leads to less cooperation

Evolutionary model (cleaner fish market)

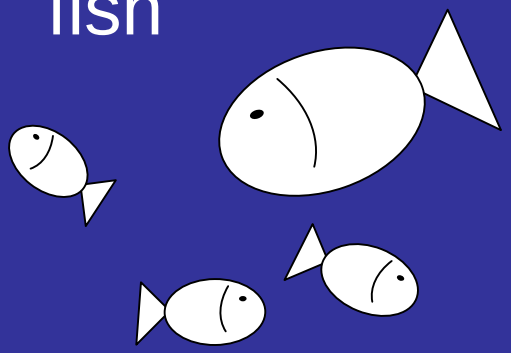
game theory - cooperation – fish simulation - experiment – discussion

Undersea market

- Repeated interaction
- Cleaner feeds on parasites; mutual benefit
- Cleaner fish can cheat on living tissue
- Client fish recognize and choose cleaners
- Client fish can end interaction

Agents

fish



Type: $\Theta_L \Theta_H$

Discriminate/not

cleaners

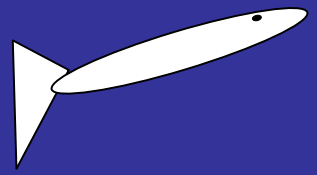


Image score: 0 or 1

Strategy

Lifetime

All fish approach one cleaner and see his image score.

If cleaner cooperated last round, image score = 1; otherwise image score = 0.

If cleaner's image score < fish's threshold level, the fish **flees**.

All fish reproduce and die

X 10

Evolution

Reproduction according to payoff

Client fish inherit threshold;

NOT type

Cleaner fish inherit strategy;

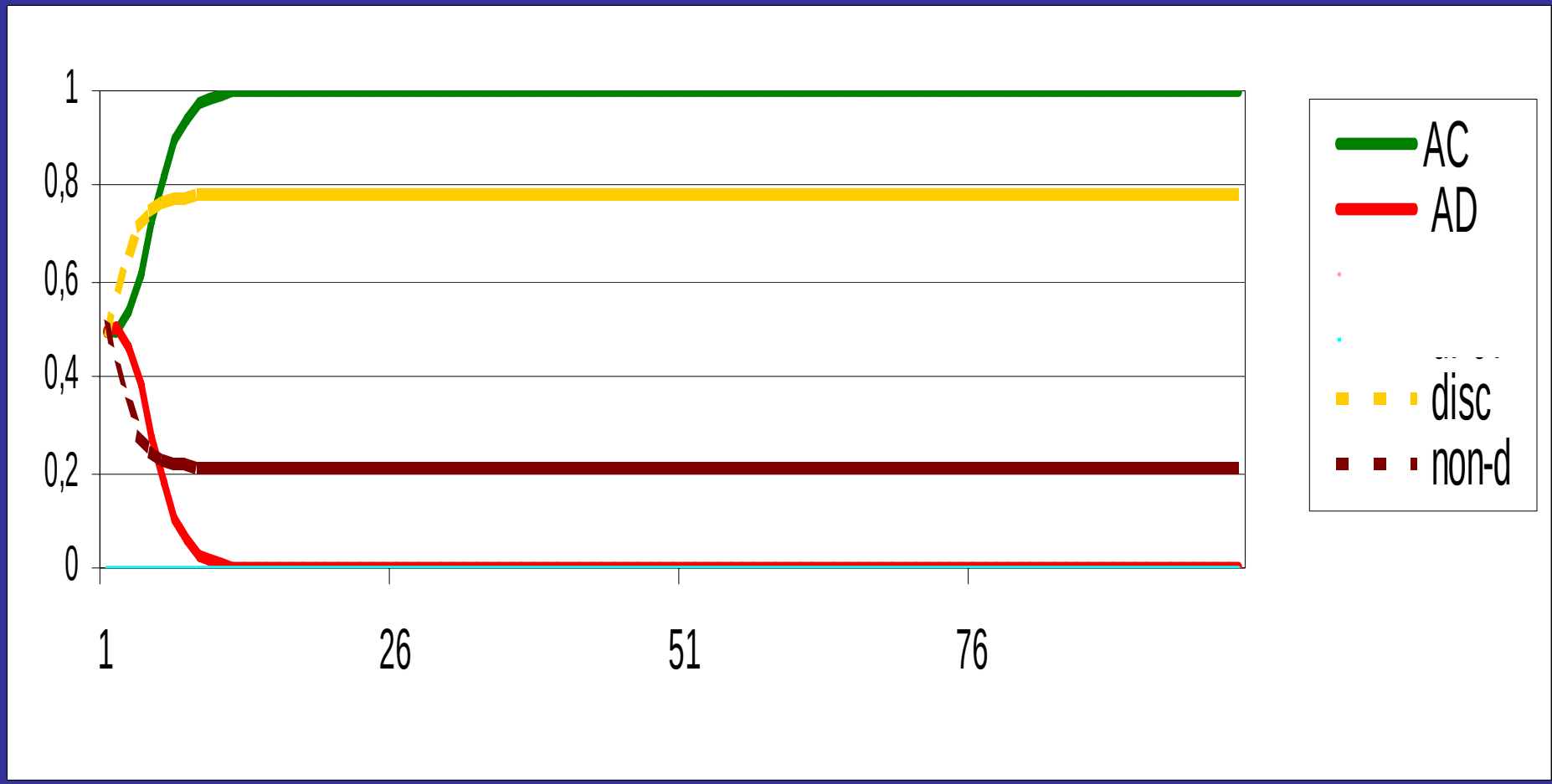
NOT image score

Homogeneous case

	fish	
cleaners	Θ_H	out
Cooperate	3,2	0,1
Defect	6,0	0,1

Results homogeneous

Strategy frequencies



Heterogeneous case

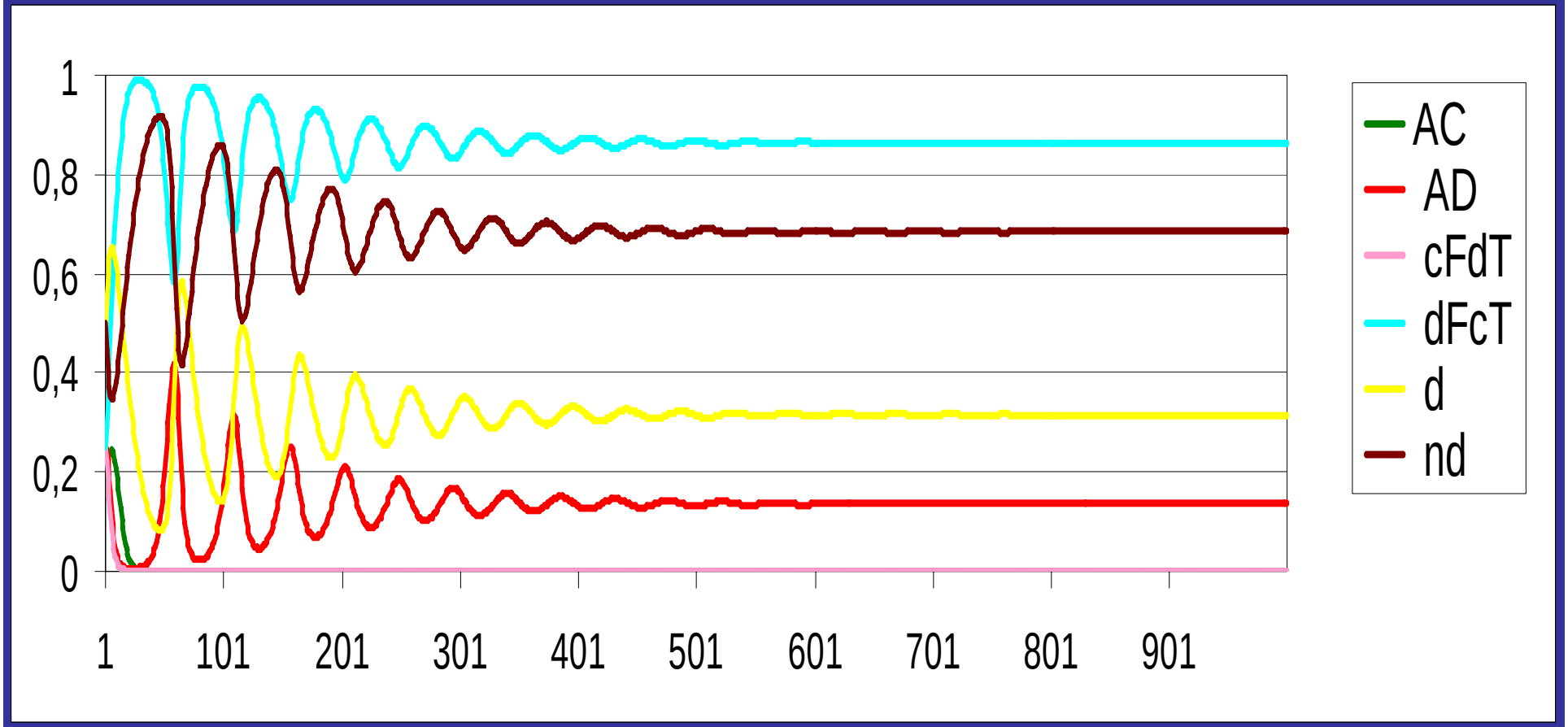
	fish		
cleaners	Θ_H	Θ_L	out
Cooperate	3, 2	1, 2	0, 1
Defect	6, 0	2, 0	0, 1

Strategies

	fish	
cleaners	Θ_H	Θ_L
Always Coop	C	C
Always Def	D	D
Cheater	D	C
Rev.Cheater	C	D

Results heterogeneous

Strategy frequencies



Generations

Conclusion

- After introducing heterogeneity, image scoring breaks down
- Cheater strategy thrives for this parameter constellation

Image scoring depends on the assumption of homogeneity

I believe it when I see it - Reputation versus Experience

motivation

Reputation (info about previous behavior)
is a valuable institution to foster trust

- ebay
- Job market
- I E N S
  INDEPENDENT INDEX
  onafhankelijke en betrouwbare restaurantgids

How do people interpret ‘reputation’?

motivation

Usual assumption:

reliable information from 3rd party (reputation)

=

information from own experience

For same history and experience, is this true?

background

Bohnet and Huck (AER, 2004):

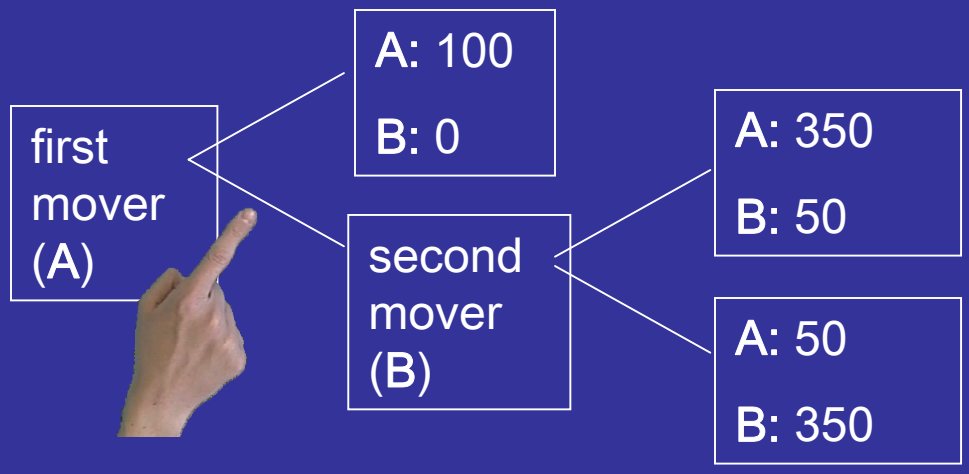
Partner treatment fosters trust in long run, more than reputation does. Why?

1. Reputation = noisy signal
2. Experience leads to emotional response

Reputation = noisy signal

But: Cox and Deck (working paper, 2002):

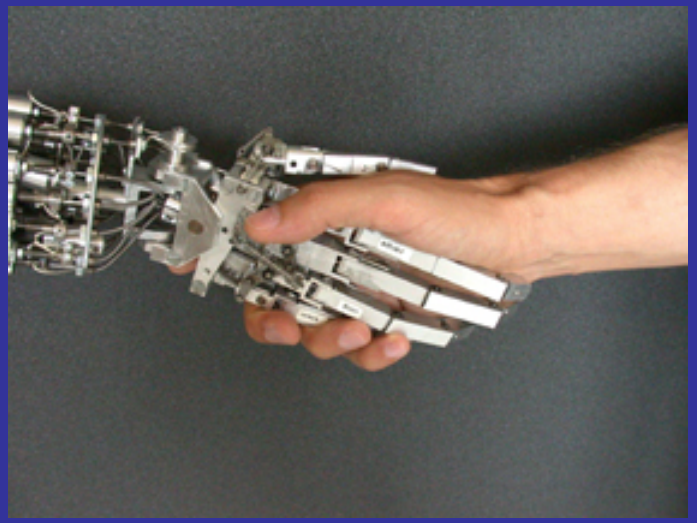
trembling trust game: positive signal is rewarded by 2nd movers



Experience leads to emotional response

Yes!

Camerer (2006) fMRI-scan shows ‘affection’



hypotheses

Differences in emotional response

Experience	*Reputation*
– personal	– anonymous
– “hot”	– “cold”
– repeated	– one shot

1. “Hot” experience => overreaction to positive / negative experience

2. Sense of repetition => more trust in partner, more forgiveness

hypotheses

"Hot" experience => overreaction to experience:

Partner + negative experience <

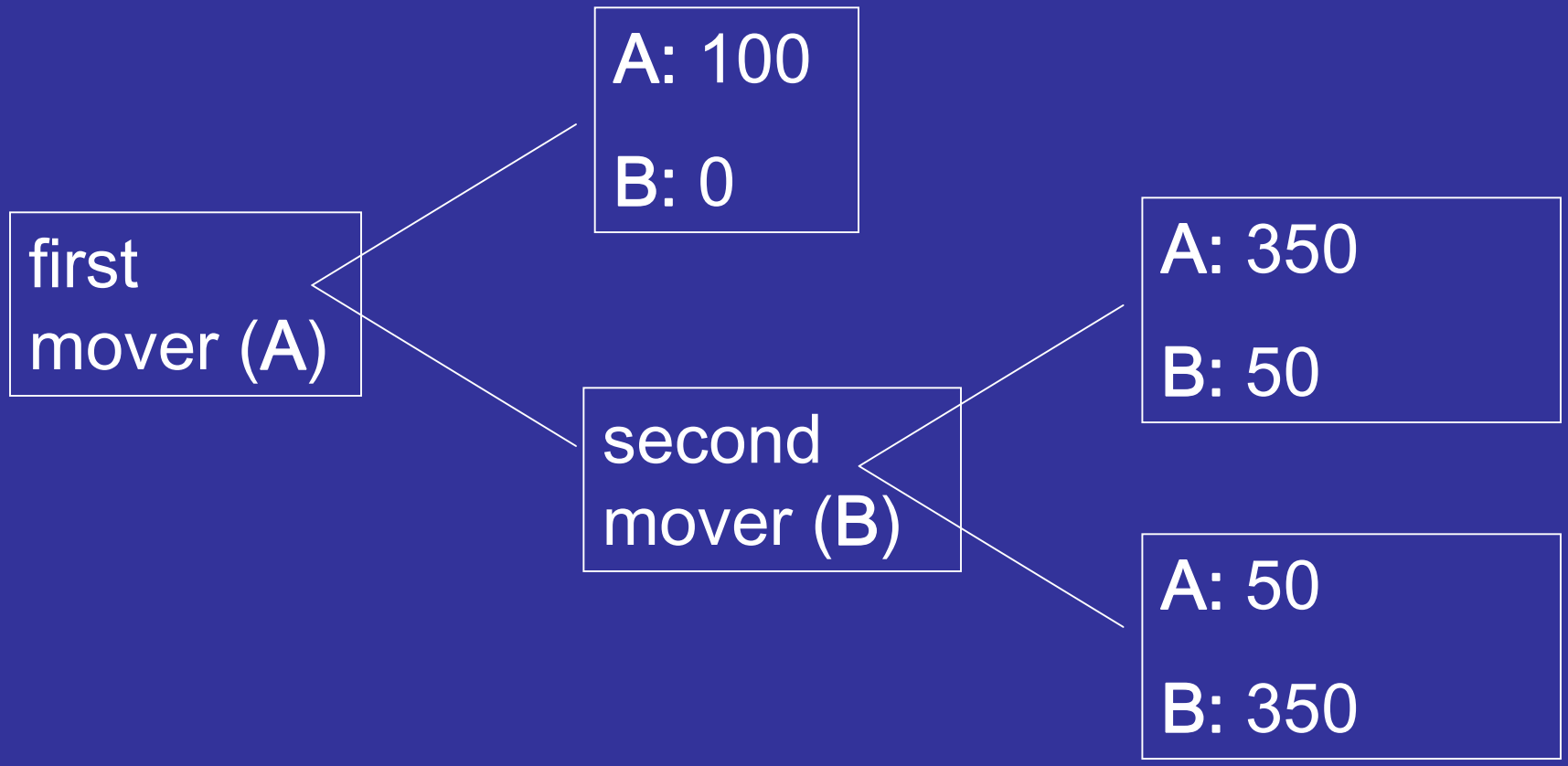
Reputation + negative experience + negative reputation

Sense of repetition => more trust in *Partner* :

Partner + negative experience >

Reputation + negative experience + negative reputation

experiment



experiment

- Binary trust game, played twice [part 1; part 2]
- Repetition unknown to subjects
- Fixed roles (player A and player B)
- 3 treatments- differ in part 2:

  -*Baseline* (same partner as part 1)

  -*Stranger* (stranger, but full information about behavior in part 1)

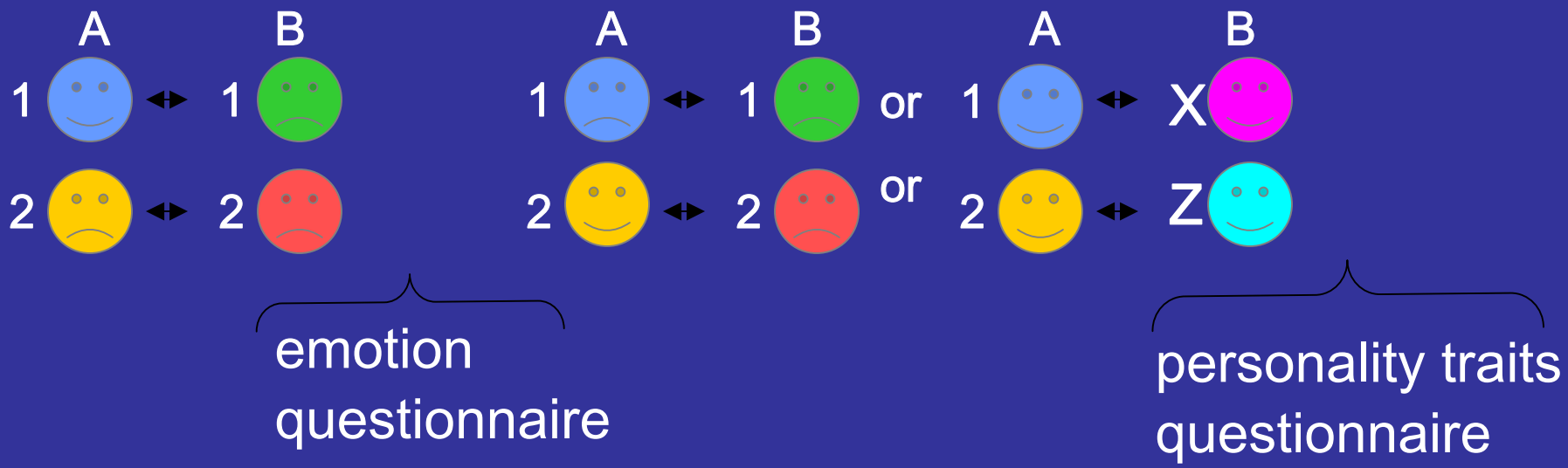
  -*Choice* (A chooses Partner or Stranger)

- Questionnaires: personality traits and emotions

experiment

part 1:

part 2:



subjects

CREED lab, summer 2006, 9 sessions

140 students (33 % female, 40 % economics)

Average age: 22

Average earnings: 10 euro (7,50-22,50)

30 minutes

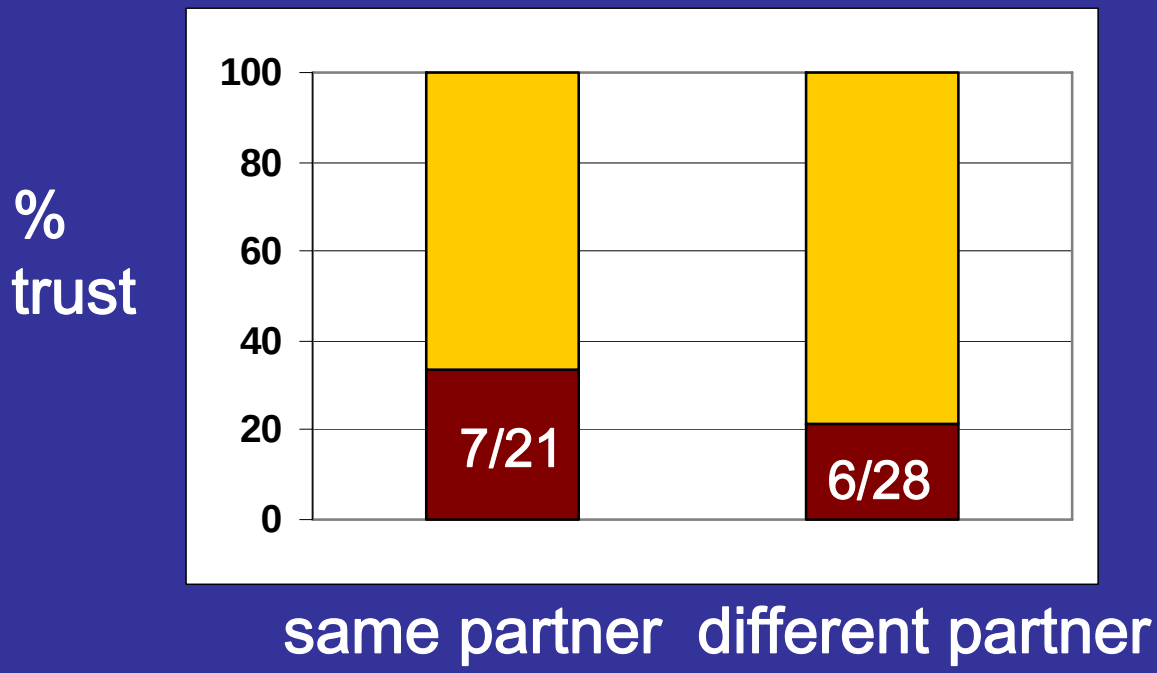
behavior: 1st round

- Total 70 observations
  - (42 same partner; 28 different partner)

- Players A (1st round):
  - trust: 41
  - no trust: 29
- Players B (1st round):
  - reward: 16
  - no reward: 54

behavior: 2nd round

- Decision to trust: A players (2nd round)
  - A players with no reward history



behavior: 2nd round

After “no reward”:

People prefer partner over stranger

- If playing with partner, people trust more
- Most subjects (7/11) chose to play with previous partner
    - when given the choice to play with same partner or with a stranger with same history

conclusions

- Reputation is valued (emotionally) differently from Own Experience
- Repetition hypothesis seems more likely than Overreaction
- The differences in trusting behavior are mediated by emotions

summary

A. Evolution did not work out

B. People are animals

C. AI is the mother of all sciences

discussion

Is it valid question to answer by means of simulation/ experiment?

Can ‘clean’ lab environment say anything about external, context-rich world?

Should we abstract from real world, like theory, or provide more details, like real world? (In casu: provide cues for recognition that stimulate ‘attachment’ and likeability)