



Modeling Reciprocity in Human Bilateral Negotiation

Ya'akov Gal and Avi Pfeffer
MIT CSAIL and
Harvard School of Engineering and Applied Sciences

1

Motivation

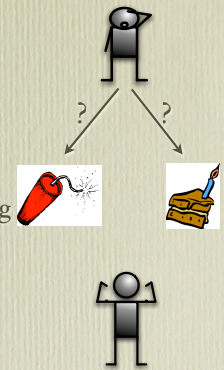
Reciprocity is a key determinant in human interaction over time.

Settings commonly

- include numerous of interaction rounds
- vary choices, payoffs at each round

Mixed human-computer settings are becoming prevalent.

Need a computational model of human reciprocity.



2

Research Focus

Modeling reciprocity requires considering two factors

- *Retrospective* reasoning : what is the extent to which others should be rewarded/punished for their past actions ?
- *Prospective* reasoning : what is the ramification (cost) of a potential action in the future ?

Example: diplomatic relations

3

Research Focus

Contribution: Learning reciprocal factors makes a difference for computer agents that bargain with people.

Approach:

- Probabilistic decision-making model.
- Represent reciprocal factors within agents' "psychological" utility functions.

4

Background: Prior Work

Equilibria in repeated games

- assume payoffs and strategies are constant at each stage game
- are difficult to compute
- humans deviate from equilibrium strategies (e.g. “turn the other cheek”)

Learning from human negotiation (Das et al. 02, Gal et al., 05)

- Prior worked learned social factors such as altruism, competitiveness in one-shot interaction
- but, did not address reciprocal factors

Repeated Negotiation Setting

Finite (unknown) number of rounds between proposer and responder agent.

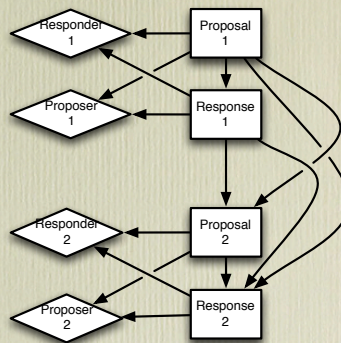
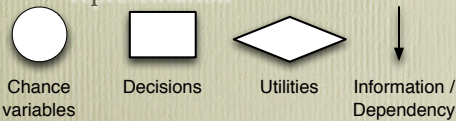
- Proposer can make multi-attribute exchange to responder.
- Responder agent can accept or reject exchange.
- Players incur payoffs after each round, based on result of negotiation.
- Each round varies
 - possible offers for proposer, players’ payoffs
 - proposer/responder roles

Probabilistic Model for Repeated Interaction

Generates a distribution over possible proposals and replies at each round.

Actions at each round depend on past history.

Issue: scalability of representation



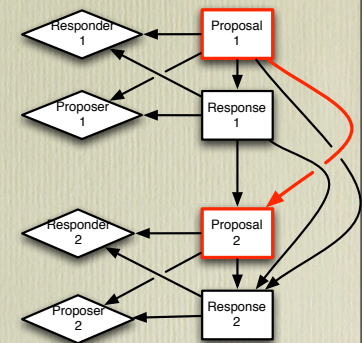
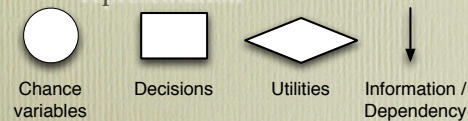
Multi-Agent Influence Diagram (MAID) (Milch and Koller, '01)

Probabilistic Model for Repeated Interaction

Generates a distribution over possible proposals and replies at each round.

Actions at each round depend on past history.

Issue: scalability of representation



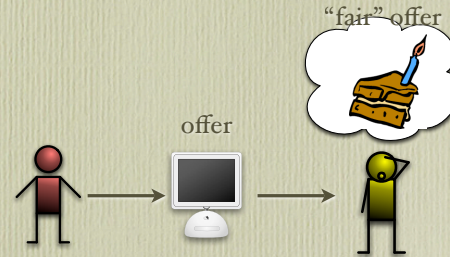
Multi-Agent Influence Diagram (Milch and Koller, '01)

I. Retrospective Reciprocity



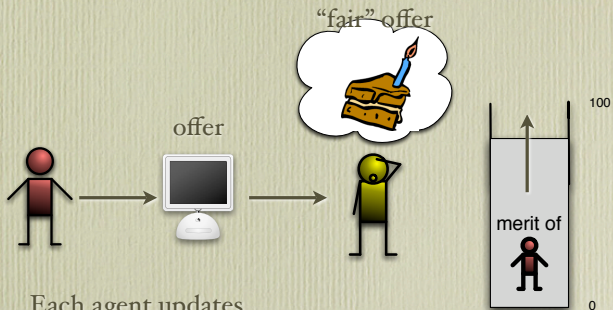
Each agent updates "merit" scalar that encapsulate other's past behavior.

I. Retrospective Reciprocity



Each agent updates "merit" scalar that encapsulate other's past behavior.

I. Retrospective Reciprocity

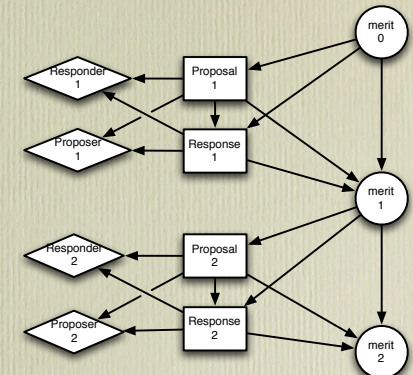


Each agent updates "merit" scalar that encapsulate other's past behavior.

merit is updated by difference in payoff between actual offer and "fair" offer

I. Retrospective Reciprocity

Agents update each other's merit values after each round.
 Merit vector is system "state".
 Makes the model compact.
 Actions at each round are conditionally independent from history given merit.



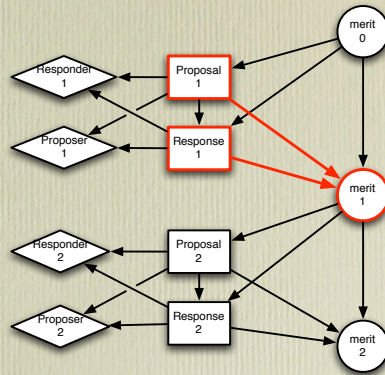
I. Retrospective Reciprocity

Agents update each other's merit values after each round.

Merit vector is system "state".

Makes the model compact.

Actions at each round are conditionally independent from history given merit.



13

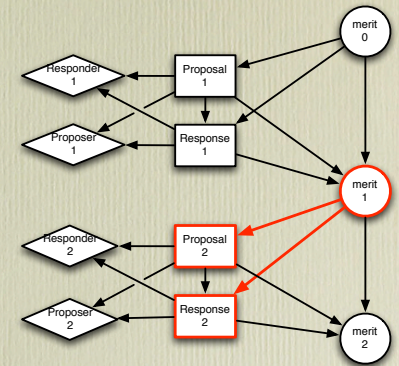
I. Retrospective Reciprocity

Agents update each other's merit values after each round.

Merit vector is system "state".

Makes the model compact.

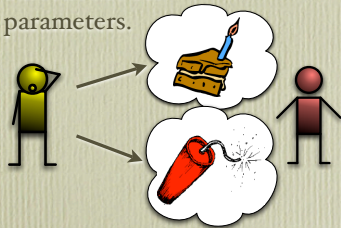
Actions at each round are conditionally independent from history given merit.



14

II. Prospective Reasoning

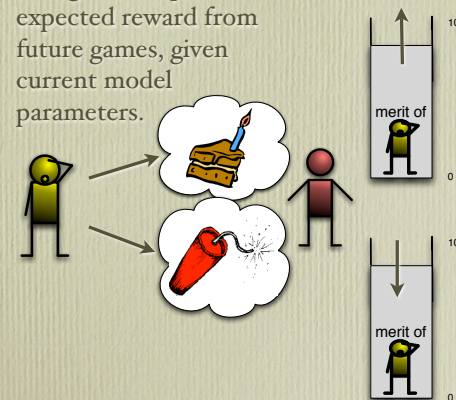
An agent computes the expected reward from future games, given current model parameters.



15

II. Prospective Reasoning

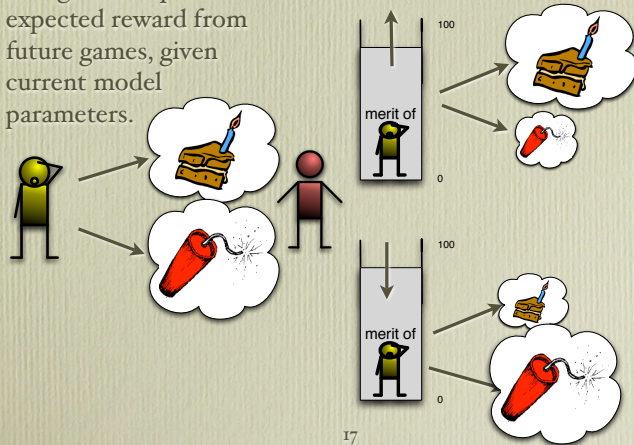
An agent computes the expected reward from future games, given current model parameters.



16

II. Prospective Reasoning

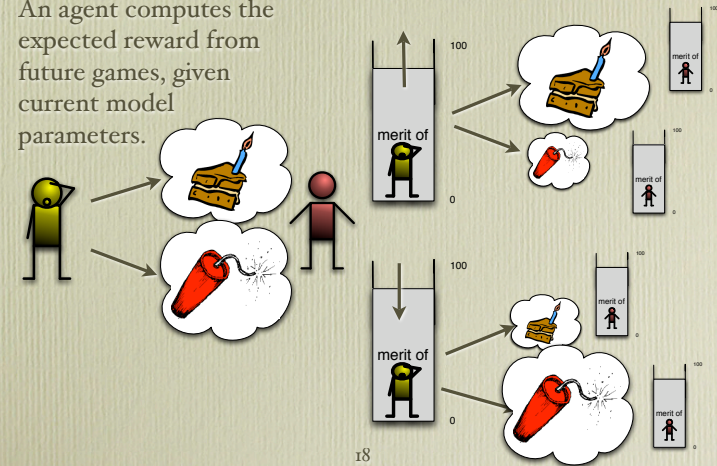
An agent computes the expected reward from future games, given current model parameters.



17

II. Prospective Reasoning

An agent computes the expected reward from future games, given current model parameters.



18

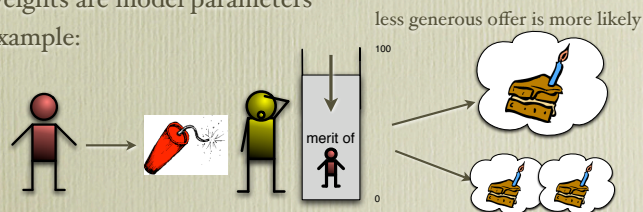
Making Decisions

Agents use “psychological” utility function, weighing

- payoff
- prospective utility: expected future outcome
- retrospective utility: product of other’s payoff with merit

Weights are model parameters

Example:



19

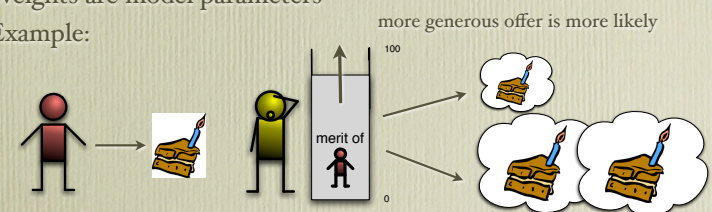
Making Decisions

Agents use “psychological” utility function, weighing

- payoff
- prospective reciprocity: expected future outcome
- retrospective reciprocity: product of other’s payoff with merit.

Weights are model parameters

Example:



20

Issues

Retrospective reasoning

- What is a “fair” exchange?
 - chose the pareto optimal exchange that maximized payoff product for both players. (Nash Bargaining)

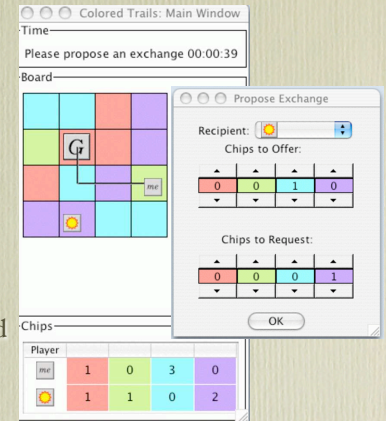
Prospective reasoning

- Computing future ramification is hard
 - future is unknown
 - need to consider the expected outcome of each action at each sampled game, given that the other agent is reasoning about reciprocity.
 - Used approximate methods; sampling; dynamic programming.

Experimental Set-up using Colored Trails (Grosz and Kraus '04)

Subjects played unknown, finite number of one-shot rounds

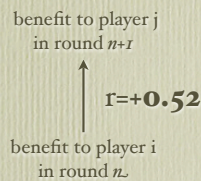
- Each round is played on a different board.
- players allocated chips, starting position and a single goal.
- proposer must make an offer to responder.
- individual score depends on reaching goal, chips retained and path taken



People Engage in Reciprocal Behavior

Collected 70 games; each includes between 1-13 Colored Trails rounds.

Correlations in the data confirm reciprocal reasoning.



Empirical Methodology

Used hierarchical model of different types of people

- EM to learn distribution over types
- gradient descent to learn utility weight values for each type

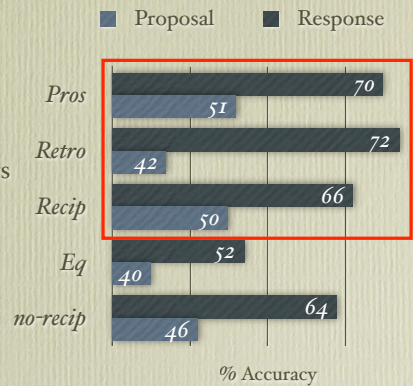
Compared between several approaches

- *recip*: learned both prospective, retrospective factors
- *pros, retro*: learned only one factor
- *equilibria strategies*: Nash bargaining, fairness equilibrium
- *no recip*: one-shot learning (Gal et al., 2004)

Results reported on held-out data set (using ten-fold cross validation)

Results: Prediction

Learning either reciprocity factor improves predictive power.
Learning both factors does not help.
Difficult to predict proposals. Up to 256 unique exchanges.



25

Future Work

Create “social” computer player to

- learn how people play in general.
- adapt to individual people over time

Evaluating player by

- playing new people in new types of situations.
- cultural distinctions

26

Colored Trails

Public release available for download

www.eecs.harvard.edu/ai/ct3

Can be used for

- investigating human-computer decision-making.
- facilitating multi-agent systems research.
- pedagogic purposes

