Decision Making
in Robots and Autonomous Agents

Game Theory: How should robots reason about *interactive* decisions?

Subramanian Ramamoorthy
School of Informatics

15 February, 2019

Robots Often Face *Strategic* Adversaries

Key issue we seek to model: Misaligned/conflicting interest

On Self-Interest

What does it mean to say that agents are self-interested?

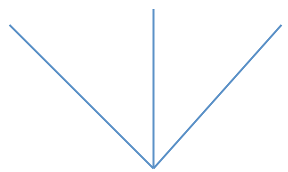
- It does not necessarily mean that they want to cause harm to each other, or even that they care only about themselves.

- Instead, it means that each agent has his *own* description of which states of the world he likes—which can include good things happening to other agents

—and that he *acts in an attempt to bring about these states of the world* (better term: *inter-dependent* decision making)

Basic Constructs of Game (Extensive Form)

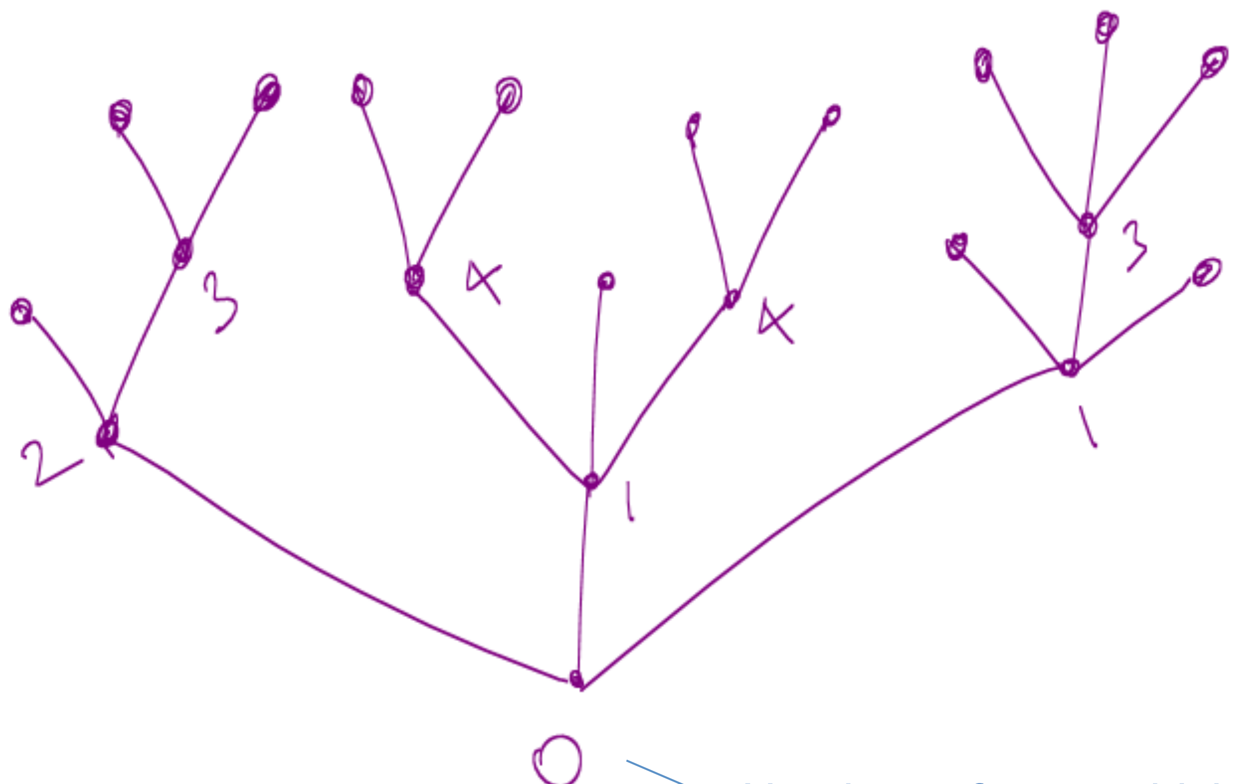
- **Move**: A point of decision for the player, defined by a set of actions that could be chosen (e.g., I am holding King-10-2). In an abstract form:



The move looks identical in a different game involving, say, passing – calling- betting

- **Choice**: The particular action that has actually been chosen (play King)
- **Play**: A sequence of choices, one following another, until the game terminates (King -> 10 -> 2)

Abstract Relation Between Moves



Number refers to which player has to make a choice

Game Trees

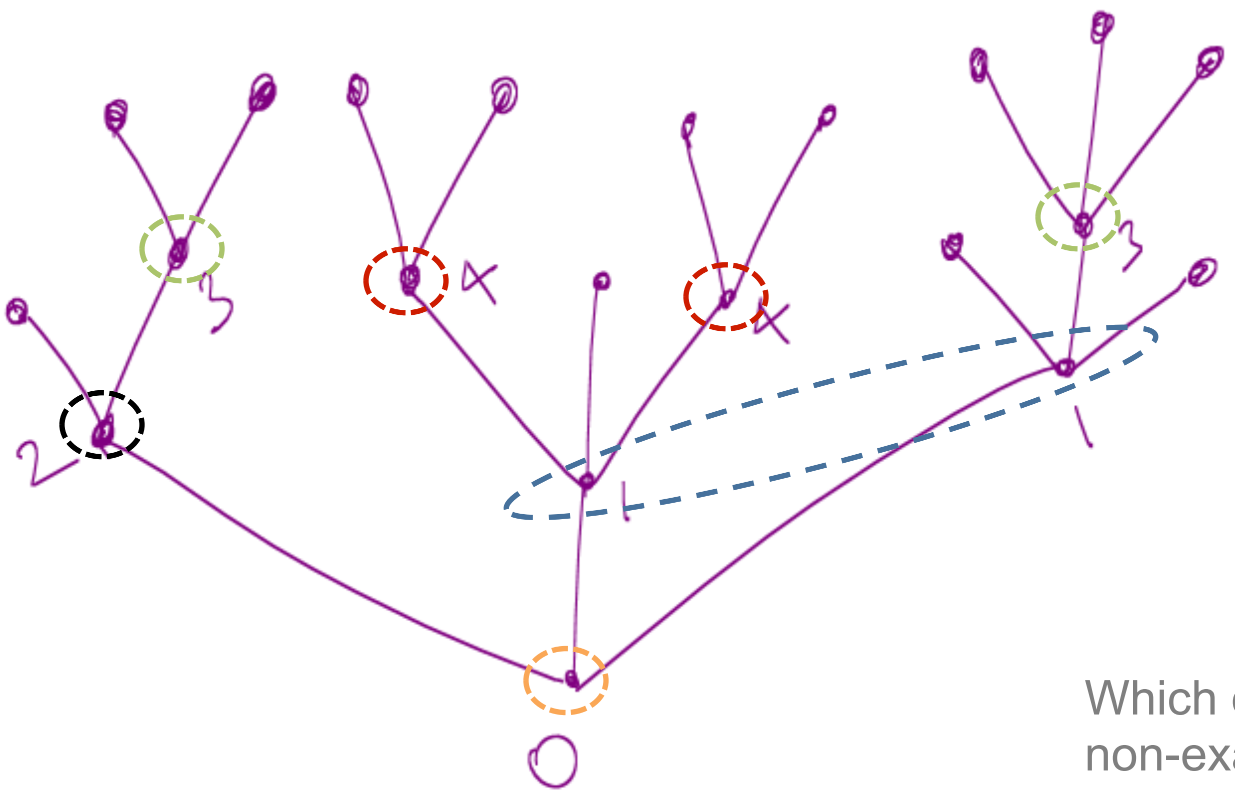
- The previous picture is often referred to as a game tree, a listing of moves and consequences
- It is a tree in the mathematical sense

- Strictly speaking, many games should have a game graph (not just a tree)
  - Why?
- However, the convention is to treat the play (history of moves) as unique even if it revisits a specific state

Information Sets

- What can each player know when he makes a choice at a move?
- We are not asking about their way of playing – just, what is the most they could possibly know without violating the rules of the game?
  - e.g., think of card games with chance moves, or where a player picks a card and places it face down on the table
- Rules of the game specify which moves are indistinguishable
- Two requirements for information set:
  - Moves must be assigned to same player
  - Moves must have same number of alternatives

Information Sets - Pictorially



Which one is a non-example?

Specification of an Extensive Form Game

- A finite tree with a distinguished node
- A partition of the nodes of the tree into $n+1$ sets (specifying who takes the move)
- A probability distribution over the branches of chance moves
- A refinement of the player partition into information sets (characterizes, for each player, the ambiguity of location of the game tree of each of his moves)
- An identification of corresponding branches for each of the moves in each information set
- A set of outcomes and an assignment of outcomes to each endpoint of the tree

How Do Players Actually Choose?

- Each player has a linear utility function M_i over outcomes
- Each player is fully aware of the rules, and will maximize expected utility

- **Pure strategy**: prescription of decision for each situation
- For any fixed strategy, given the rules of the game, the game tree can be evaluated directly to yield a value
- If there are chance moves, selection of strategies of players defines a distribution over plays and the payoff is expected value w.r.t. this distribution

A Different Simple Model of a Game

- Two decision makers
  - Robot (has an action space: *A*)
  - Adversary (has an action space: θ)
- *Cost* or *payoff* (to use the term common in game theory) depends on actions of both decision makers:

  $R(a, \theta)$ – denote as a matrix corresponding to product space

$$
\begin{array}{cc}
 & \Theta \\
A & \begin{array}{|c|c|c|} \hline 1 & -1 & 0 \\ \hline -1 & 2 & -2 \\ \hline 2 & -1 & 1 \\ \hline \end{array}
\end{array}
$$

This is the **normal form** – simultaneous choice over moves

Representing Payoffs

In a general, bi-matrix, normal form game $(n, \mathcal{A}_{1\ldots n}, R_{1\ldots n})$

Action sets of players

Payoff function: $\mathcal{A} \rightarrow \Re$

$$R_1 = \begin{array}{cc} & a_2 \\ a_1 & \begin{pmatrix} & \vdots & \\ \cdots & R_1(a) & \cdots \\ & \vdots & \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{cc} & a_2 \\ a_1 & \begin{pmatrix} & \vdots & \\ \cdots & R_2(a) & \cdots \\ & \vdots & \end{pmatrix} \end{array}$$

a.k.a. utility $u_2(a)$

The combined actions $(a_1, a_2, \ldots, a_n)$ form an

action profile a ∈ A

Example: Rock-Paper-Scissors

- Famous children's game
- Two players; Each player simultaneously picks an action which is evaluated as follows,
  - Rock beats Scissors
  - Scissors beats Paper
  - Paper beats Rock

$$R_1 = \begin{array}{c c} & \begin{array}{ccc} \mathsf{R} & \mathsf{P} & \mathsf{S} \end{array} \\ \begin{array}{c} \mathsf{R} \\ \mathsf{P} \\ \mathsf{S} \end{array} & \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{c c} & \begin{array}{ccc} \mathsf{R} & \mathsf{P} & \mathsf{S} \end{array} \\ \begin{array}{c} \mathsf{R} \\ \mathsf{P} \\ \mathsf{S} \end{array} & \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \end{array}$$

TCP Game

- Imagine there are only two internet users: you and me
- Internet traffic is governed by TCP protocol, one feature of which is the *backoff* mechanism: when network is congested then backoff and reduce transmission rates for a while
- Imagine that there are two implementations: C (correct, does what is intended) and D (defective)
- If you both adopt C, packet delay is 1 ms; if you both adopt D, packet delay is 3 ms
- If one adopts C but other adopts D then D user gets no delay and C user suffers 4 ms delay

TCP Game in Normal Form

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Note that this is another way of writing a bi-matrix game: First number represents payoff of row player and second number is payoff for column player

Some Famous Matrix Examples - What are they Capturing?

- Prisoner's Dilemma: Cooperate or Defect (same as TCP game)

$$R_1 = \begin{array}{c c} & \begin{array}{cc} \mathsf{C} & \mathsf{D} \end{array} \\ \begin{array}{c} \mathsf{C} \\ \mathsf{D} \end{array} & \begin{pmatrix} 3 & 0 \\ 4 & 1 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{c c} & \begin{array}{cc} \mathsf{C} & \mathsf{D} \end{array} \\ \begin{array}{c} \mathsf{C} \\ \mathsf{D} \end{array} & \begin{pmatrix} 3 & 4 \\ 0 & 1 \end{pmatrix} \end{array}$$

- Bach or Stravinsky (von Neumann called it Battle of the Sexes)

$$R_1 = \begin{array}{c c} & \begin{array}{cc} \mathsf{B} & \mathsf{S} \end{array} \\ \begin{array}{c} \mathsf{B} \\ \mathsf{S} \end{array} & \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{c c} & \begin{array}{cc} \mathsf{B} & \mathsf{S} \end{array} \\ \begin{array}{c} \mathsf{B} \\ \mathsf{S} \end{array} & \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \end{array}$$

- Matching Pennies: Try to get the same outcome, Heads/Tails

$$R_1 = \begin{array}{c c} & \begin{array}{cc} \mathsf{H} & \mathsf{T} \end{array} \\ \begin{array}{c} \mathsf{H} \\ \mathsf{T} \end{array} & \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{c c} & \begin{array}{cc} \mathsf{H} & \mathsf{T} \end{array} \\ \begin{array}{c} \mathsf{H} \\ \mathsf{T} \end{array} & \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \end{array}$$

Different Categorization: Common Payoff

A common-payoff game is a game in which for all action profiles $a \in A_1 \times \cdots \times A_n$ and any pair of agents i, j, it is the case that $u_i(a) = u_j(a)$

	Left	Right
Left	1, 1	0, 0
Right	0, 0	1, 1

Pure coordination:
e.g., driving on a side of the road

Different Categorization: Constant Sum

A two-player normal-form game is constant-sum if there exists a constant c such that for each strategy profile $a \in A_1 \times A_2$ *it is the case that* $u_1(a) + u_2(a) = c$

	Heads	Tails
Heads	$1, -1$	$-1, 1$
Tails	$-1, 1$	$1, -1$

Pure competition:
One player wants to coordinate
Other player does not!

What Can Players Do?

What can players do?

- Pure strategies (a_i): select an action.
- Mixed strategies (σ_i): select an action according to some probability distribution.

Strategies

Notation.

- σ is a joint strategy for all players.

$$R_i(\sigma) = \sum_{a \in \mathcal{A}} \sigma(a) R_i(a)$$

Expected utility

- σ_{-i} is a joint strategy for all players except i.
- $\langle \sigma_i, \sigma_{-i} \rangle$ is the joint strategy where i uses strategy σ_i and everyone else σ_{-i}.

Solution Concepts

Many ways of describing what one ought to do:

- Dominance
- Minimax
- Pareto Efficiency
- Nash Equilibria
- Correlated Equilibria

Remember that in the end game theory aspires to predict behaviour given specification of the game.

Normatively, a solution concept is a *rationale* for behaviour

Concept: Dominance

- An action is strictly dominated if another action is always better, i.e,

$$\exists a_i' \in \mathcal{A}_i \; \forall a_{-i} \in \mathcal{A}_{-i} \qquad R_i(\langle a_i', a_{-i} \rangle) > R_i(\langle a_i, a_{-i} \rangle).$$

- Consider prisoner's dilemma.

$$R_1 = \begin{array}{c c} & \begin{array}{cc} \mathsf{C} & \mathsf{D} \end{array} \\ \begin{array}{c} \mathsf{C} \\ \mathsf{D} \end{array} & \begin{pmatrix} 3 & 0 \\ 4 & 1 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{c c} & \begin{array}{cc} \mathsf{C} & \mathsf{D} \end{array} \\ \begin{array}{c} \mathsf{C} \\ \mathsf{D} \end{array} & \begin{pmatrix} 3 & 4 \\ 0 & 1 \end{pmatrix} \end{array}$$

  - For both players, D dominates C.

Concept: Iterated Dominance

- Actions may be dominated by mixed strategies.

$$R_1 = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{cc} D & E \end{array} \\ \begin{pmatrix} 1 & 1 \\ 4 & 0 \\ 0 & 4 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{cc} D & E \end{array} \\ \begin{pmatrix} 4 & 0 \\ 1 & 2 \\ 0 & 1 \end{pmatrix} \end{array}$$

- If strictly dominated actions should not be played. . .

$$R_1 = \begin{array}{c} \\ \cancel{A} \\ \cancel{B} \\ C \end{array} \begin{array}{c} \begin{array}{cc} \cancel{D} & E \end{array} \\ \begin{pmatrix} \cancel{1} & \cancel{1} \\ \cancel{4} & \cancel{0} \\ \cancel{0} & 4 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{c} \\ \cancel{A} \\ \cancel{B} \\ C \end{array} \begin{array}{c} \begin{array}{cc} \cancel{D} & E \end{array} \\ \begin{pmatrix} \cancel{4} & \cancel{0} \\ \cancel{1} & \cancel{2} \\ \cancel{0} & 1 \end{pmatrix} \end{array}$$

- This game is said to be dominance solvable.

Concept: Minimax

- Consider matching pennies.

$$R_1 = \begin{array}{cc} & \begin{array}{cc} \mathsf{H} & \mathsf{T} \end{array} \\ \begin{array}{c} \mathsf{H} \\ \mathsf{T} \end{array} & \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{cc} & \begin{array}{cc} \mathsf{H} & \mathsf{T} \end{array} \\ \begin{array}{c} \mathsf{H} \\ \mathsf{T} \end{array} & \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \end{array}$$

- Q: What do we do when the world is out to get us?
  A: Make sure it can't.

- Play strategy with the best worst-case outcome.

$$\operatorname*{argmax}_{\sigma_i \in \Delta(\mathcal{A}_i)} \min_{a_{-i} \in \mathcal{A}_{-i}} R_i(\langle \sigma_i, \sigma_{-i} \rangle)$$

- Minimax optimal strategy.

Minimax

- Back to matching pennies.

$$R_1 = \begin{array}{c c} & \begin{array}{cc} \text{H} & \text{T} \end{array} \\ \begin{array}{c} \text{H} \\ \text{T} \end{array} & \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{array} \qquad \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = \sigma_1^*$$

- Consider Bach or Stravinsky.

$$R_1 = \begin{array}{c c} & \begin{array}{cc} \text{B} & \text{S} \end{array} \\ \begin{array}{c} \text{B} \\ \text{S} \end{array} & \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \end{array} \qquad \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix} = \sigma_1^*$$

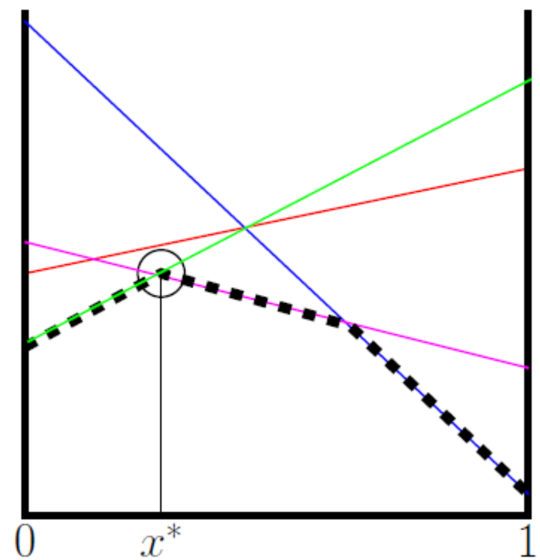
- Minimax optimal guarantees the saftey value.

- Minimax optimal never plays dominated strategies.

Computing Minimax

- Minimax optimal strategies via linear programming.

$$\operatorname*{argmax}_{\sigma_i \in \Delta(\mathcal{A}_i)} \min_{a_{-i} \in \mathcal{A}_{-i}} R_i(\langle \sigma_i, \sigma_{-i} \rangle)$$



Pick-a-Hand

- There are two players: chooser (player I) & hider (player II)
- The hider has two gold coins in his back pocket. At the beginning of a turn, he puts his hands behind his back and either takes out one coin and holds it in his left hand, or takes out both and holds them in his right hand.
- The chooser picks a hand and wins any coins the hider has hidden there.
- She may get nothing (if the hand is empty), or she might win one coin, or two.

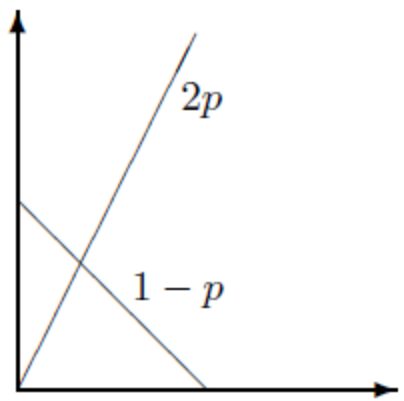
Pick-a-Hand, Normal Form:

chooser \ hider	$L1$	$R2$
L	1	0
R	0	2

- Hider could minimize losses by placing 1 coin in left hand, most he can lose is 1
- If chooser can figure out hider's plan, he will surely lose that 1
- If hider thinks chooser might strategise, he has incentive to play R2, …
- All hider can guarantee is max loss of 1 coin
- Similarly, chooser might try to maximise gain, picking R
- However, if hider strategizes, chooser ends up with zero
- So, chooser can't actually guarantee winning anything

Pick-a-Hand, with Mixed Strategies

- Suppose that chooser decides to choose R with probability p and L with probability 1 – p
- If hider were to play pure strategy R2 his expected loss would be 2p
- If he were to play L1, expected loss is 1 – p
- Chooser maximizes her gains by choosing p so as to **maximize min{2p, 1 – p}**



- Thus, by choosing R with probability 1/3 and L with probability 2/3, chooser assures expected payoff of 2/3, **regardless of whether hider knows her strategy**

Mixed Strategy for the Hider

- Hider will play R2 with some probability q and L1 with probability 1−q
- The payoff for chooser is 2q if she picks R, and 1 − q if she picks L
- If she knows q, she will choose the strategy corresponding to the maximum of the two values.
- If hider knows chooser's plan, he will choose q = 1/3 to minimize this maximum, guaranteeing that his expected payout is 2/3 (because 2/3 = 2q = 1 − q)
- Chooser can assure expected gain of 2/3, hider can assure an expected loss of no more than 2/3, regardless of what either knows of the other's strategy.

Safety Value as Incentive

- Clearly, without some extra incentive, it is not in hider's interest to play *Pick-a-hand* because he can only lose by playing.
- Thus, we can imagine that chooser pays hider to entice him into joining the game.
- 2/3 is the maximum amount that chooser should pay him in order to gain his participation.

Another Game

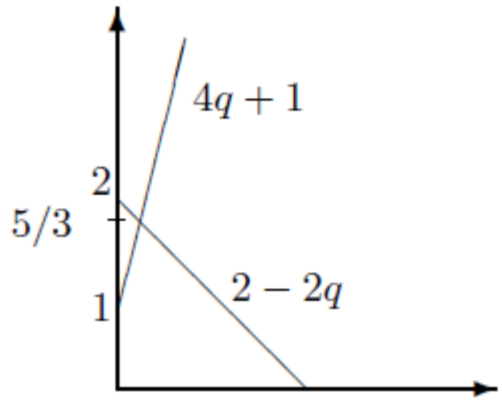
player I		player II L	player II R
	T	0	2
	B	5	1

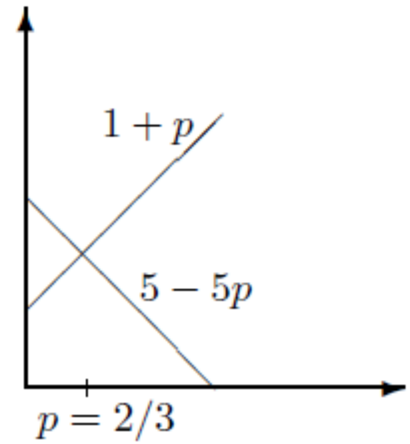
Mixed strategies:

- Suppose player I plays T with probability p and B with probability 1−p
- Player II plays L with probability q and R with probability 1 – q

- For player I, expected payoff is 2(1 – q) for playing pure strategy T; 4q + 1 for playing pure strategy B.
- If she knows q, she'll pick the strategy corresponding to max{2(1 – q), 4q + 1}
- Player II can choose q = 1/6 so as to minimize this maximum, and expected amount player II will pay player I is 5/3.

The Game Analysed Graphically





If pl. I knows q, she'll pick strategy based on max{2(1 − q), 4q + 1}. Player II can choose q = 1/6 so as to minimize this maximum. Expected amount player II will pay player I is **5/3**.

For pl. II, expected loss is 5(1 − p) if he plays pure strategy L and 1+p if he plays pure strategy R; he will aim to minimize this expected payout. In order to maximize this minimum, player I will choose p = 2/3, yielding expected gain **5/3**.

Concept: Nash Equilibrium

- What action should we play if there are no dominated actions?
- Optimal action depends on actions of other players.
- A best response set is the set of all strategies that are optimal given the strategies of the other players.

$$\mathrm{BR}_i(\sigma_{-i}) = \{\sigma_i \quad | \quad \forall \sigma'_i \quad R_i(\langle \sigma_i, \sigma_{-i} \rangle) \geq R_i(\langle \sigma'_i, \sigma_{-i} \rangle)\}$$

- A Nash equilibrium is a joint strategy, where all players are playing best responses to each other.

$$\forall i \in \{1 \ldots n\} \qquad \sigma_i \in \mathrm{BR}_i(\sigma_{-i})$$

Nash Equilibrium

- A Nash equilibrium is a joint strategy, where all players are playing best responses to each other.

$$\forall i \in \{1 \ldots n\} \qquad \sigma_i \in \mathrm{BR}_i(\sigma_{-i})$$

- Since each player is playing a best response, no player can gain by unilaterally deviating.
- Dominance solvable games have obvious equilibria.
  - Strictly dominated actions are never best responses.
  - Prisoner's dilemma has a single Nash equilibrium.

Nash Equilibrium - Example

- Consider the coordination game.

$$R_1 = \begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} \boxed{2} & 0 \\ 0 & 1 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} \boxed{2} & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

- Consider Bach or Stravinsky.

$$R_1 = \begin{array}{cc} & \begin{array}{cc} B & S \end{array} \\ \begin{array}{c} B \\ S \end{array} & \begin{pmatrix} \boxed{2} & 0 \\ 0 & \boxed{1} \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{cc} & \begin{array}{cc} B & S \end{array} \\ \begin{array}{c} B \\ S \end{array} & \begin{pmatrix} \boxed{1} & 0 \\ 0 & \boxed{2} \end{pmatrix} \end{array}$$

Nash Equilibrium - Example

- Consider matching pennies.

$$R_1 = \begin{array}{c|c} & \begin{array}{cc} \mathsf{H} & \mathsf{T} \end{array} \\ \begin{array}{c} \mathsf{H} \\ \mathsf{T} \end{array} & \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{array} \qquad R_2 = \begin{array}{c|c} & \begin{array}{cc} \mathsf{H} & \mathsf{T} \end{array} \\ \begin{array}{c} \mathsf{H} \\ \mathsf{T} \end{array} & \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \end{array}$$

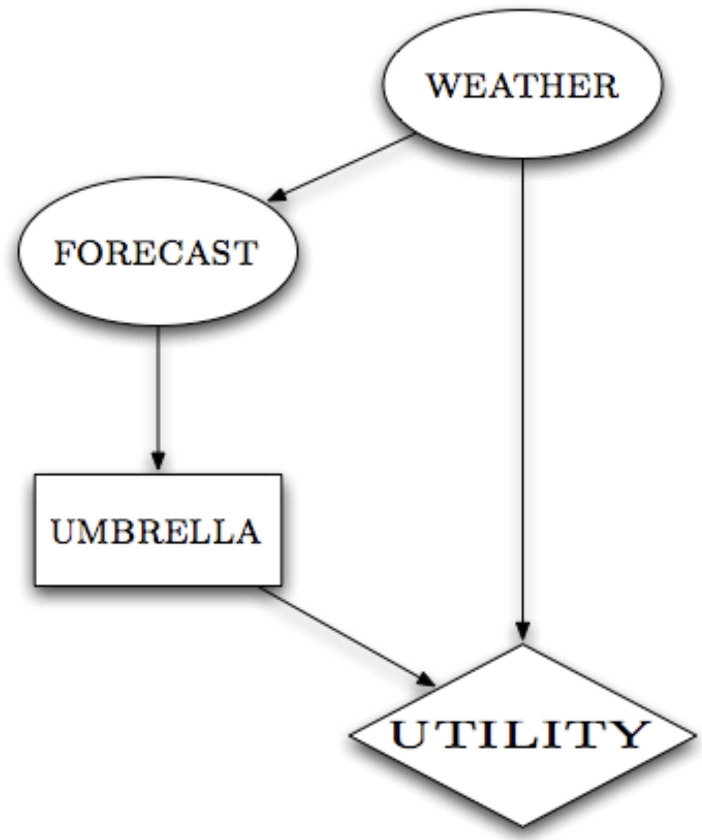
  - No pure strategy Nash equilibria. Mixed strategies?

$$\mathrm{BR}_1\Big(\langle 1/2, 1/2 \rangle\Big) = \{\sigma_1\}$$

  - Corresponds to the minimax strategy.

Influence Diagrams [Howard & Matheson '84]

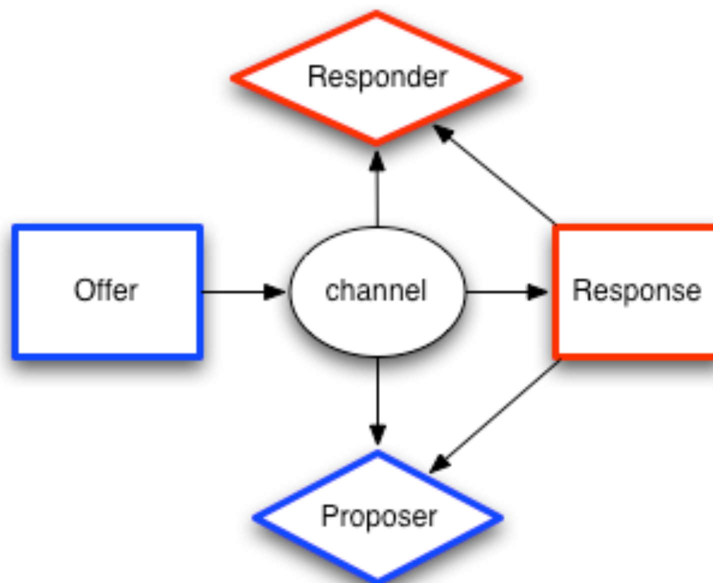
- Influence Diagrams (ID) extend Bayesian Networks for decision making.
- *Rectangles* are decisions; *ovals* are chance variables; *diamonds* are utility functions.
- Graph topology describes decision problem.
- Each node specifies a probability distribution (CPD) given each value of parents.



Multi-agent Influence Diagrams

[Milch and Koller '01]

- Extend Influence Diagrams to the multi-agent case.
- Rectangles and diamonds represent decisions and utilities associated with agents; ovals represent chance variables.
- A strategy for a decision is a mapping from the informational parents of the decision to a value in its domain.
- A strategy profile includes strategies for all decisions.



Reasoning Patterns through IDs

- Informally, a reasoning pattern is a form of argument that leads to and explains a decision
  - e.g.
    - modus ponens in logic
    - explaining away in Bayes nets

- **What reasoning patterns can agents use in *interactive* decision making contexts?**

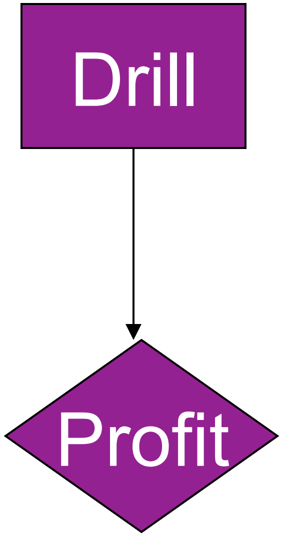
[A. Pfeffer & Y. Gal, On the reasoning patterns of agents in games, In Proc. AAAI 2007]

Characterization of Reasoning Patterns

- Four basic reasoning patterns, each characterized by paths in a multiple-agent version of influence diagrams

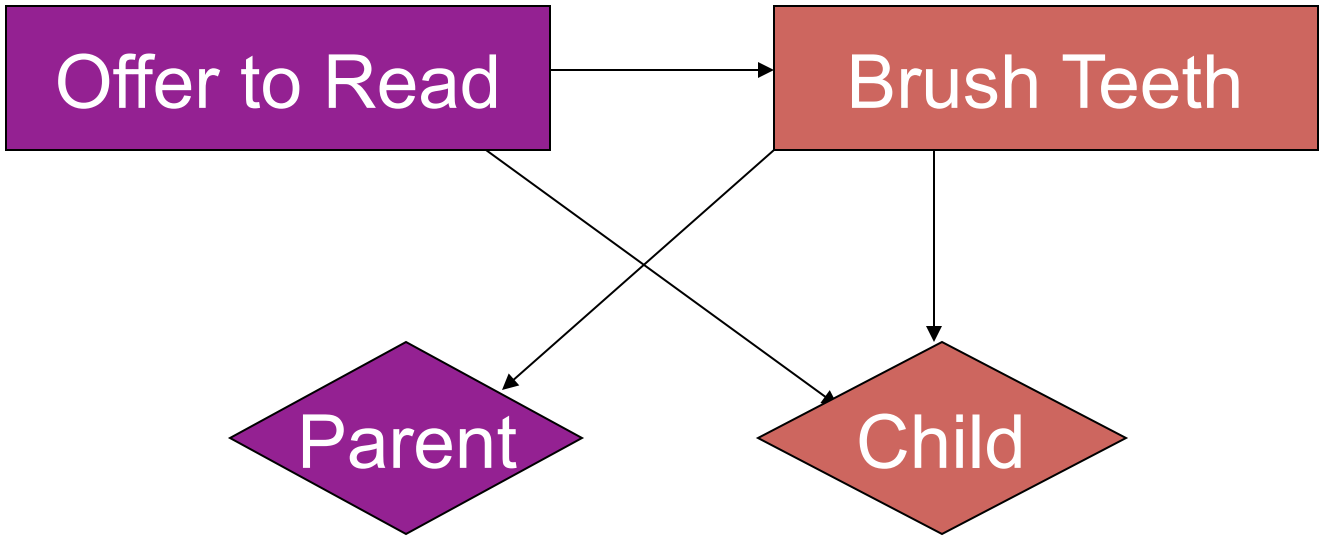
- Characterization based on graphical criteria only
  - could further refine characterization based on numerical parameters

Reasoning Pattern #1: Direct Effect



- An agent takes a decision because of its direct effect on its utility
  – without being mediated by other agents' actions

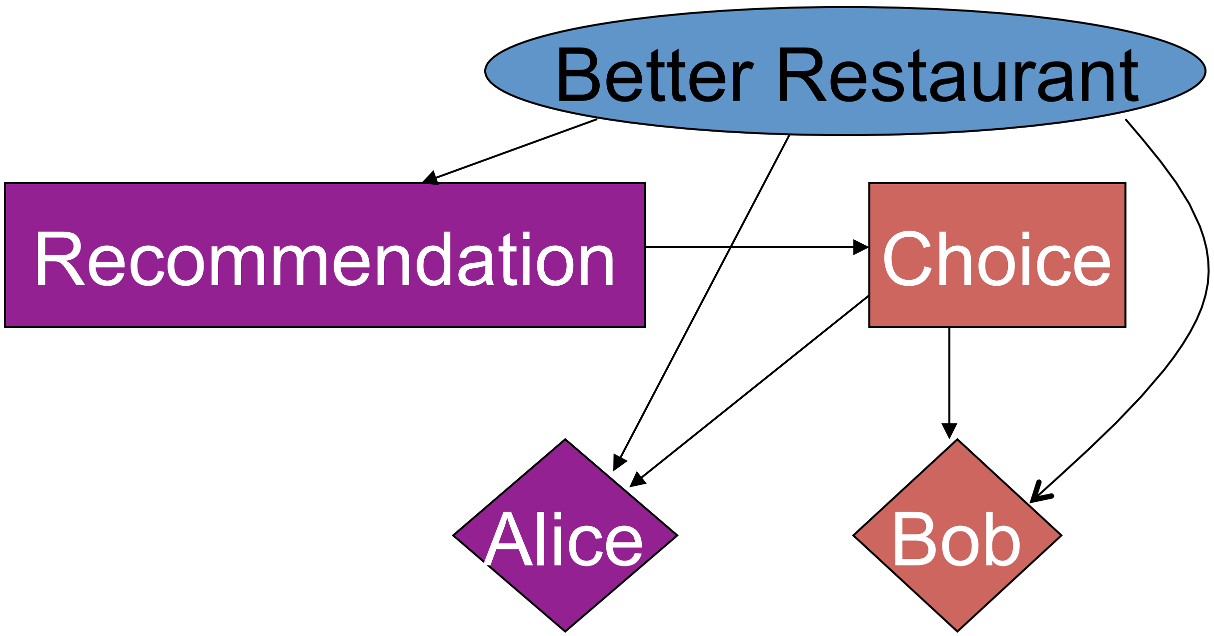
Reasoning Pattern #2: Manipulation



- Child knows about parent's action
- Parent does not care about reading, but wants child to brush teeth
- Child dislikes brushing teeth but likes being read to

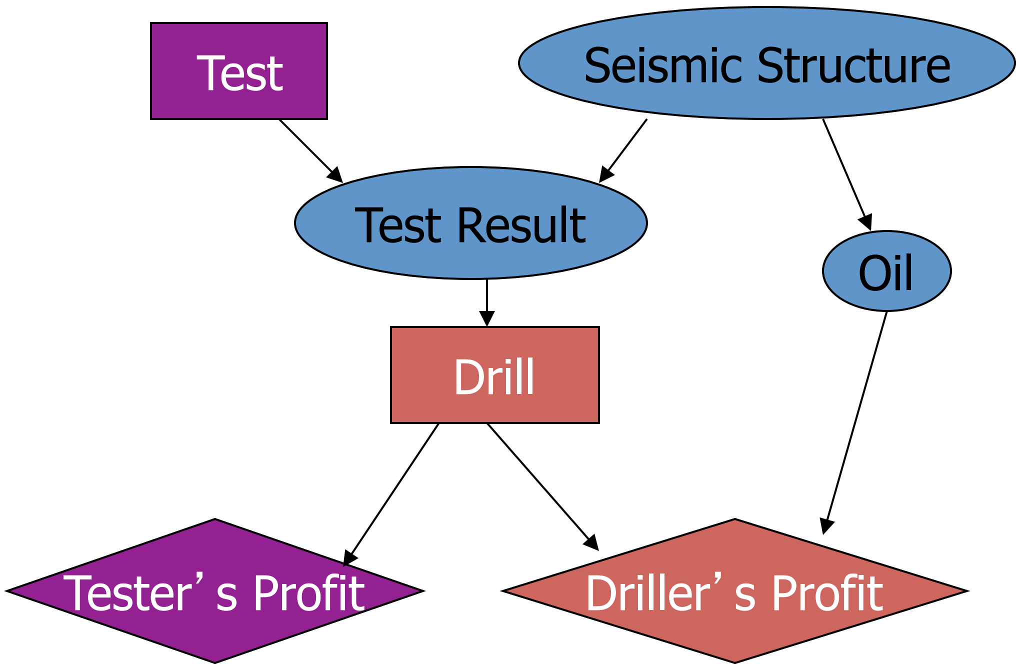
⇒ Parent can manipulate child

Reasoning Pattern #3: Signaling



- A communicates something that she knows to B, thus influencing B’s behavior

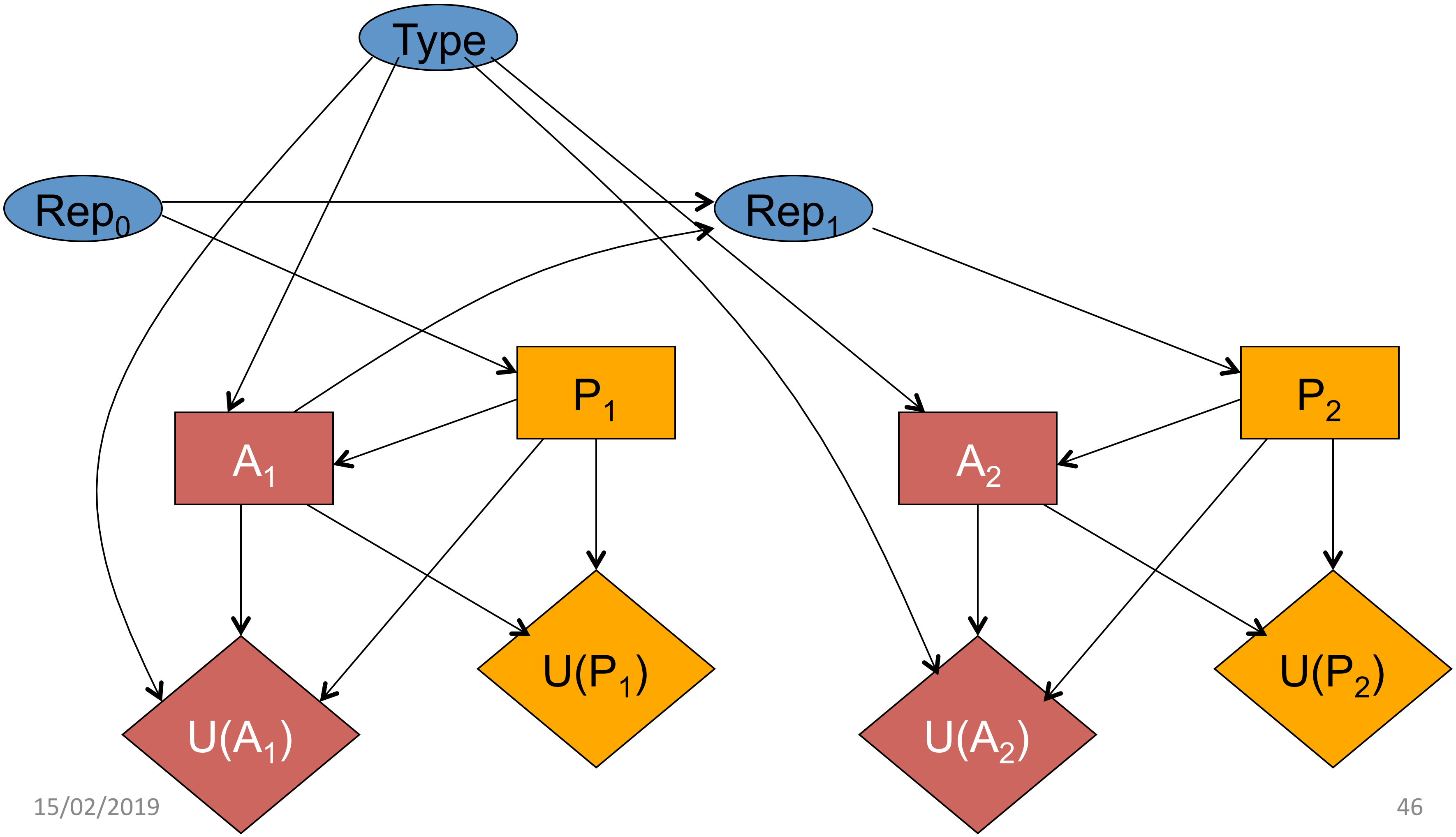
Reasoning Pattern #4: Revealing/Denying



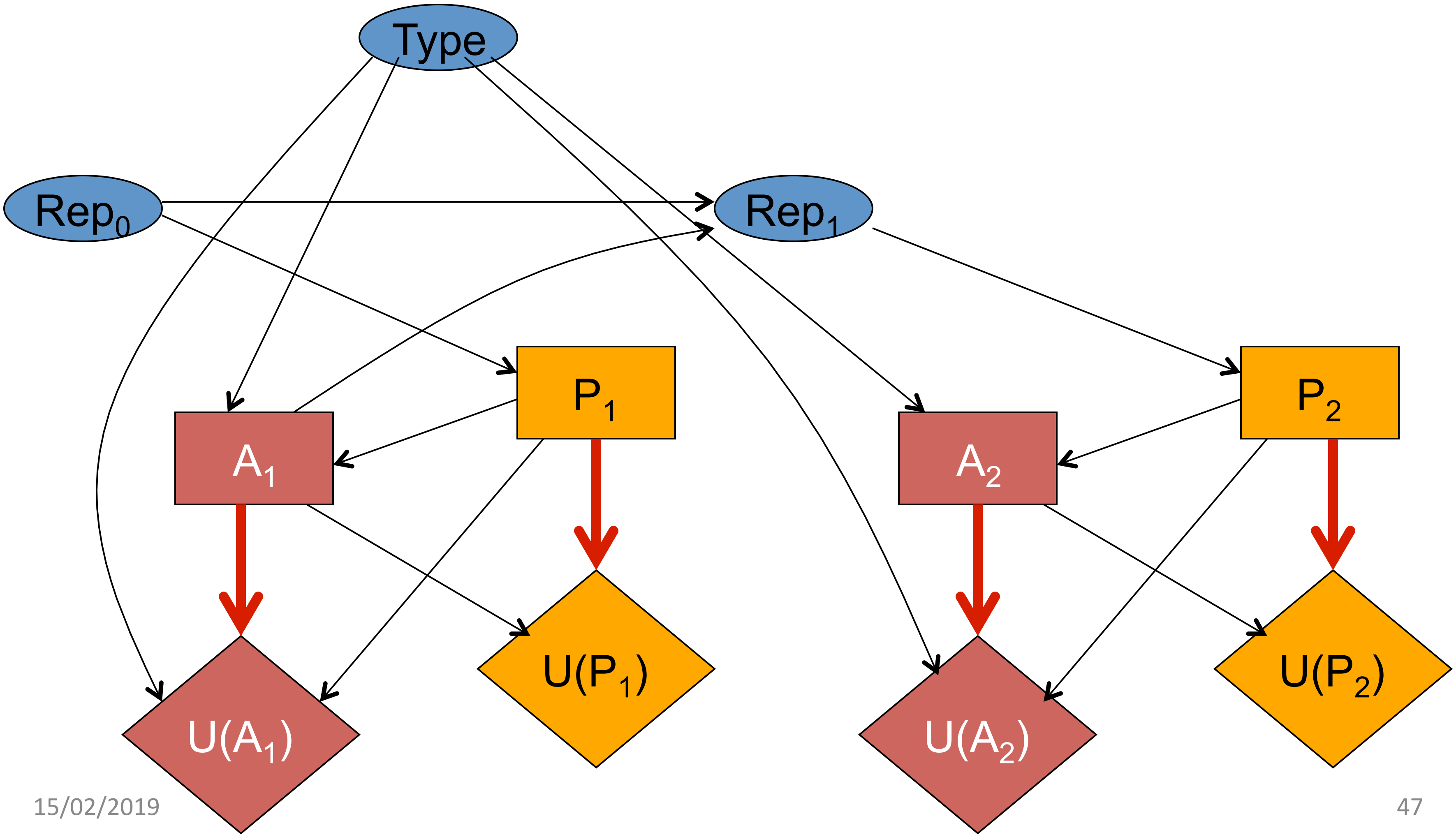
- Driller cares about oil
- Tester receives fee if driller drills
- Tester causes driller to find out (or not) about information tester herself does *not* know

Example: Two Stage Principal-Agent *Game*

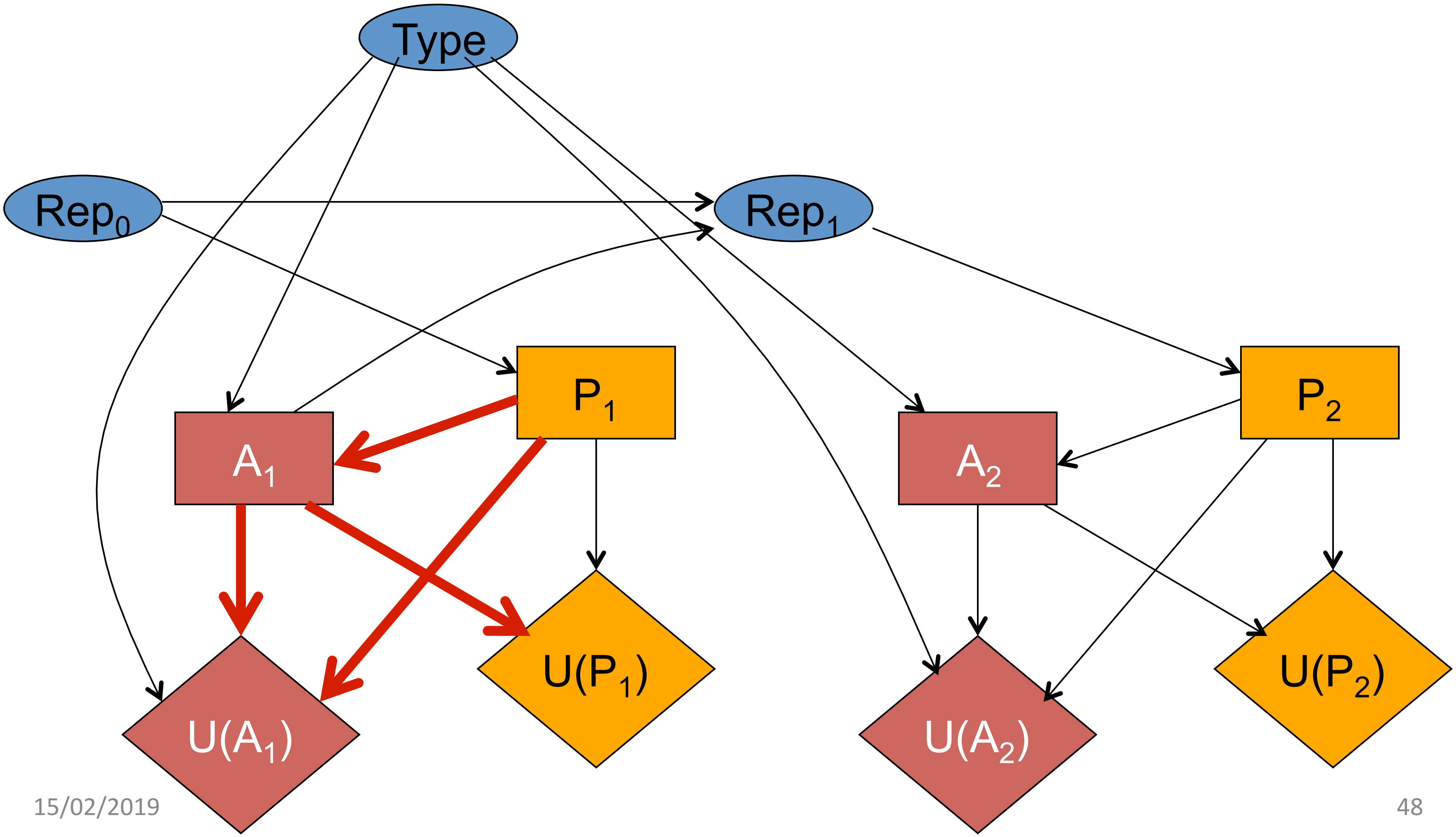
Type: described parameters specific to an agent
Rep: Quantification of "Reputation"



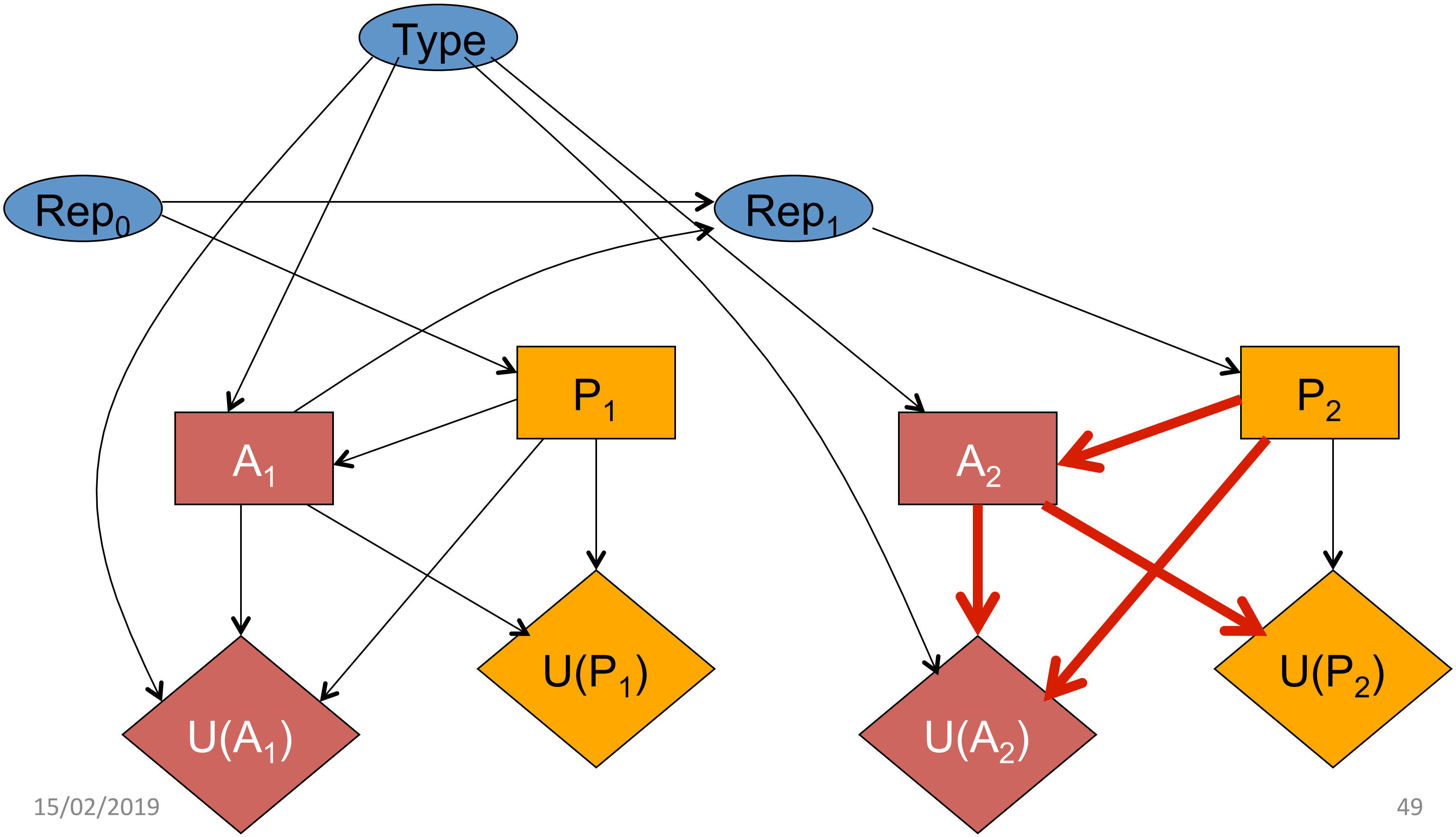
Direct Effect For All Four Decisions



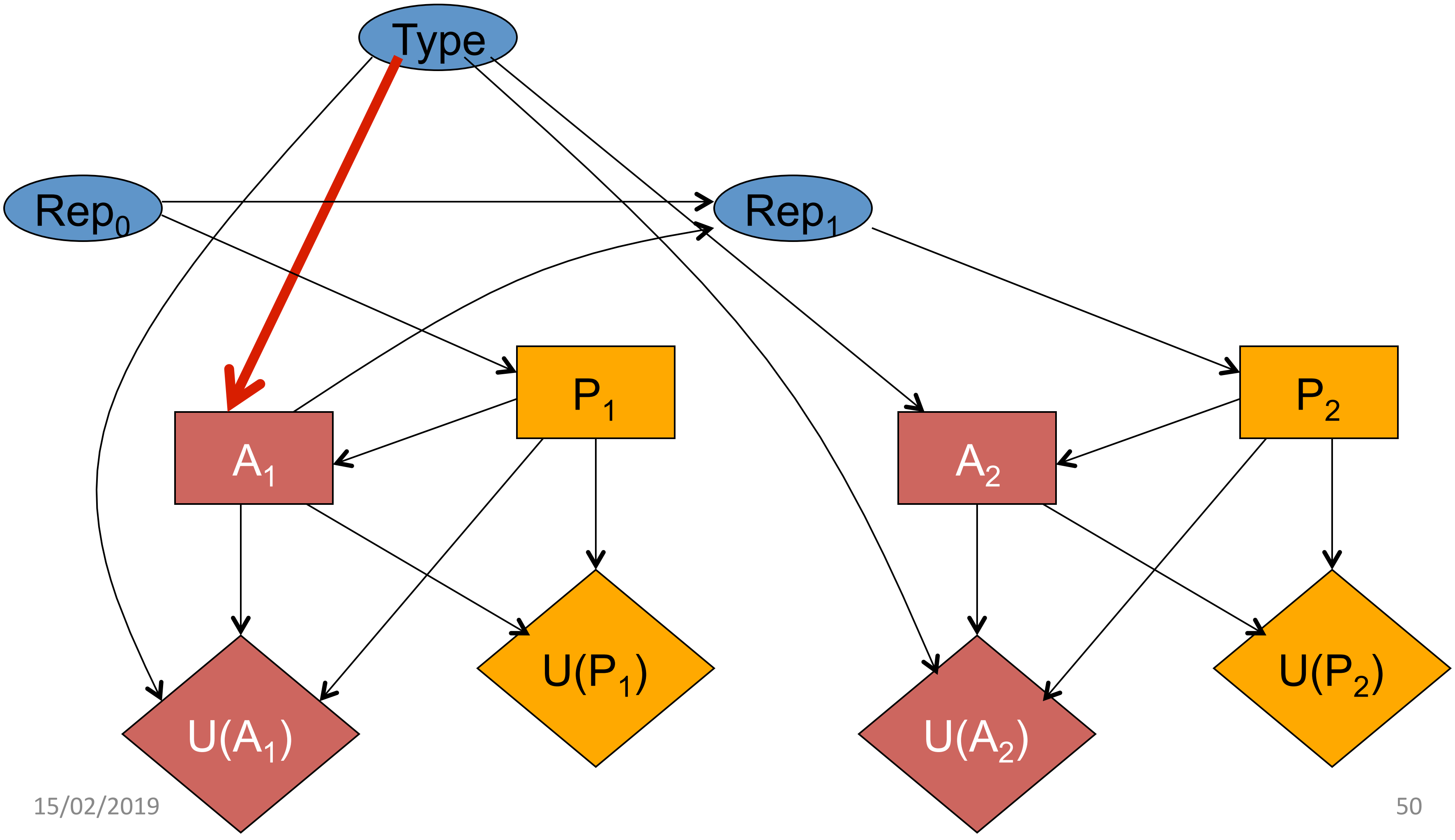
Manipulation ($P_1 \rightarrow A_1$)



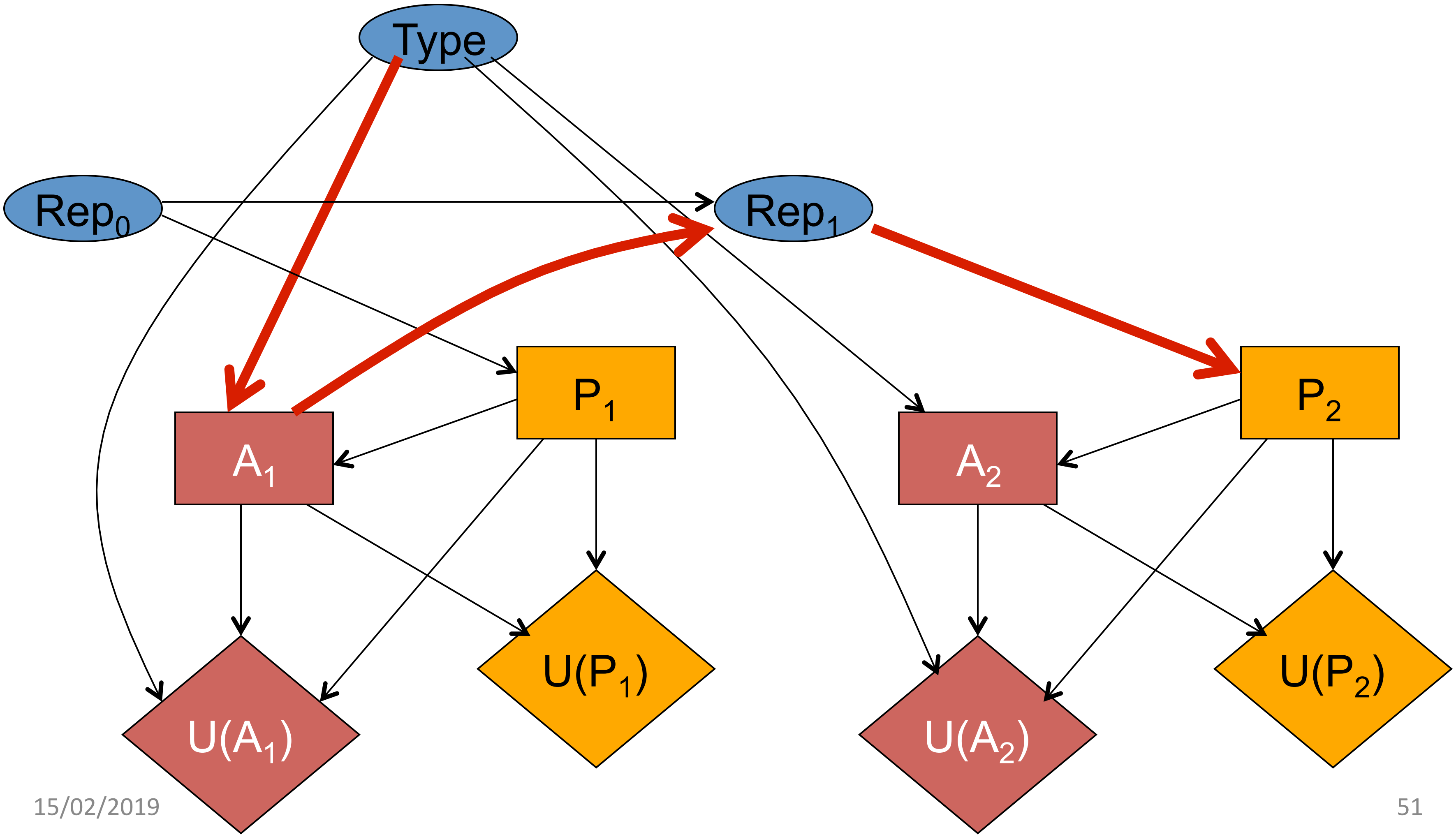
Manipulation ($P_2 \rightarrow A_2$)



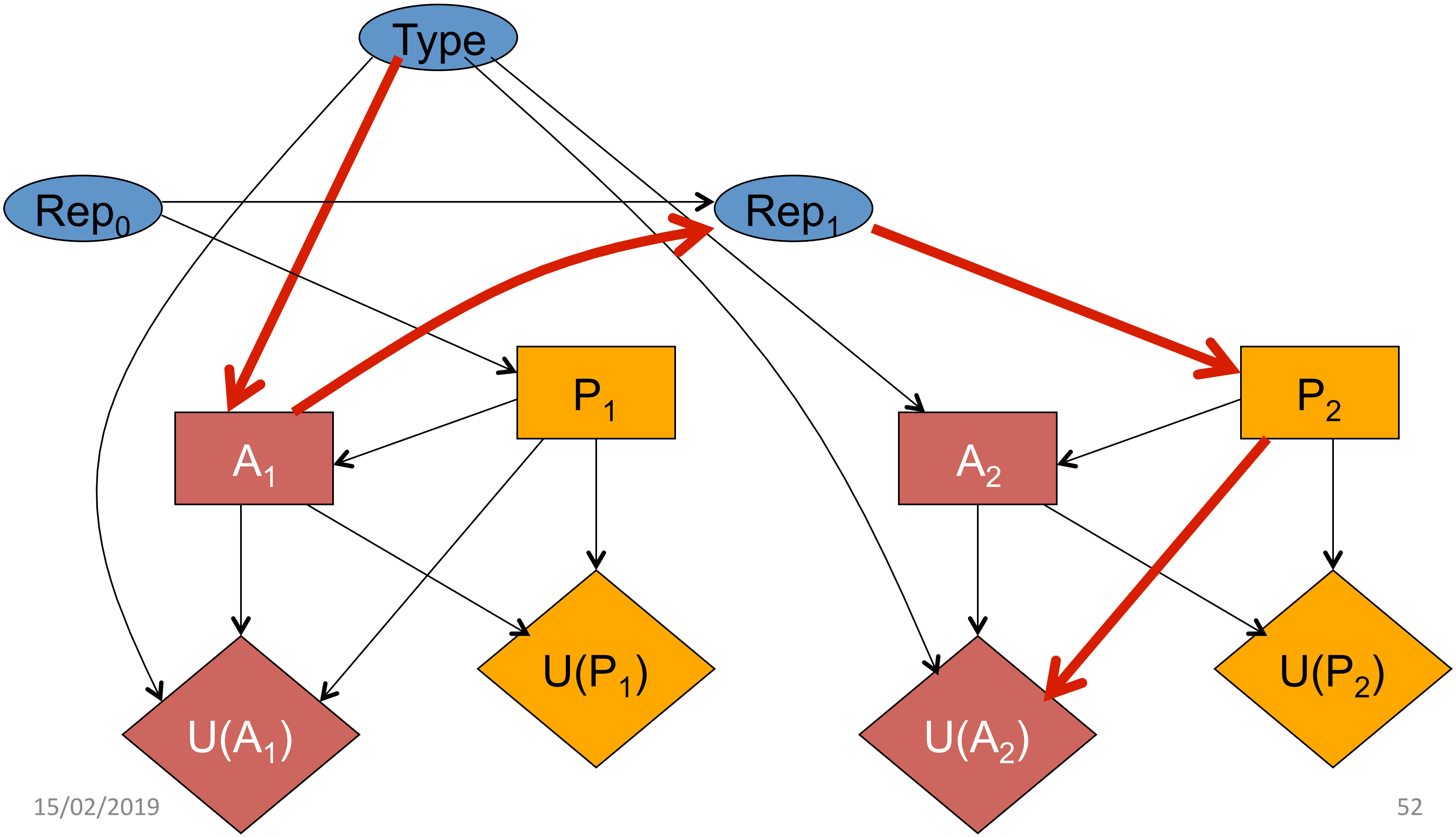
Signaling (A_1 signals Type to P_2)



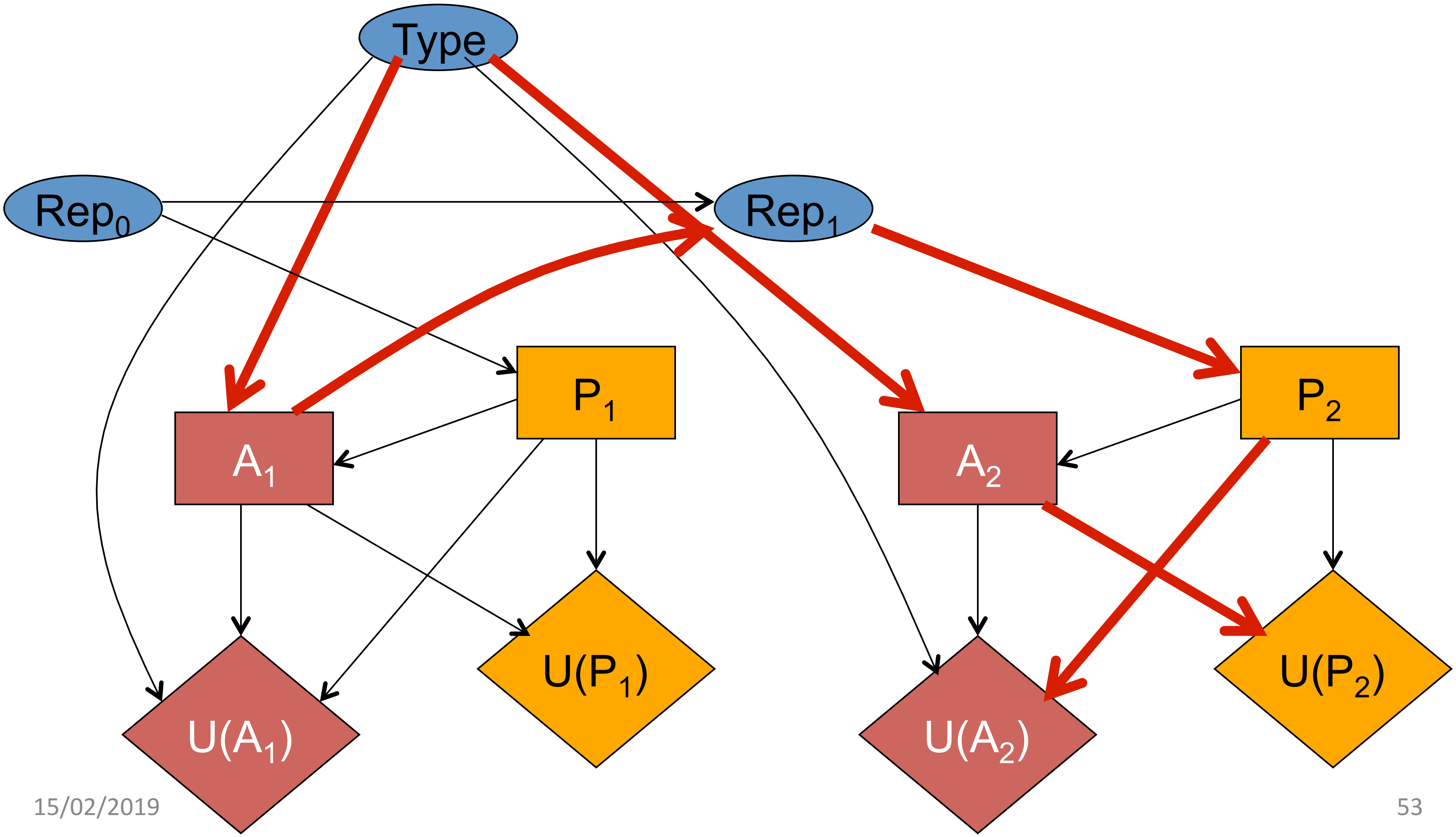
Signaling (A_1 signals Type to P_2)



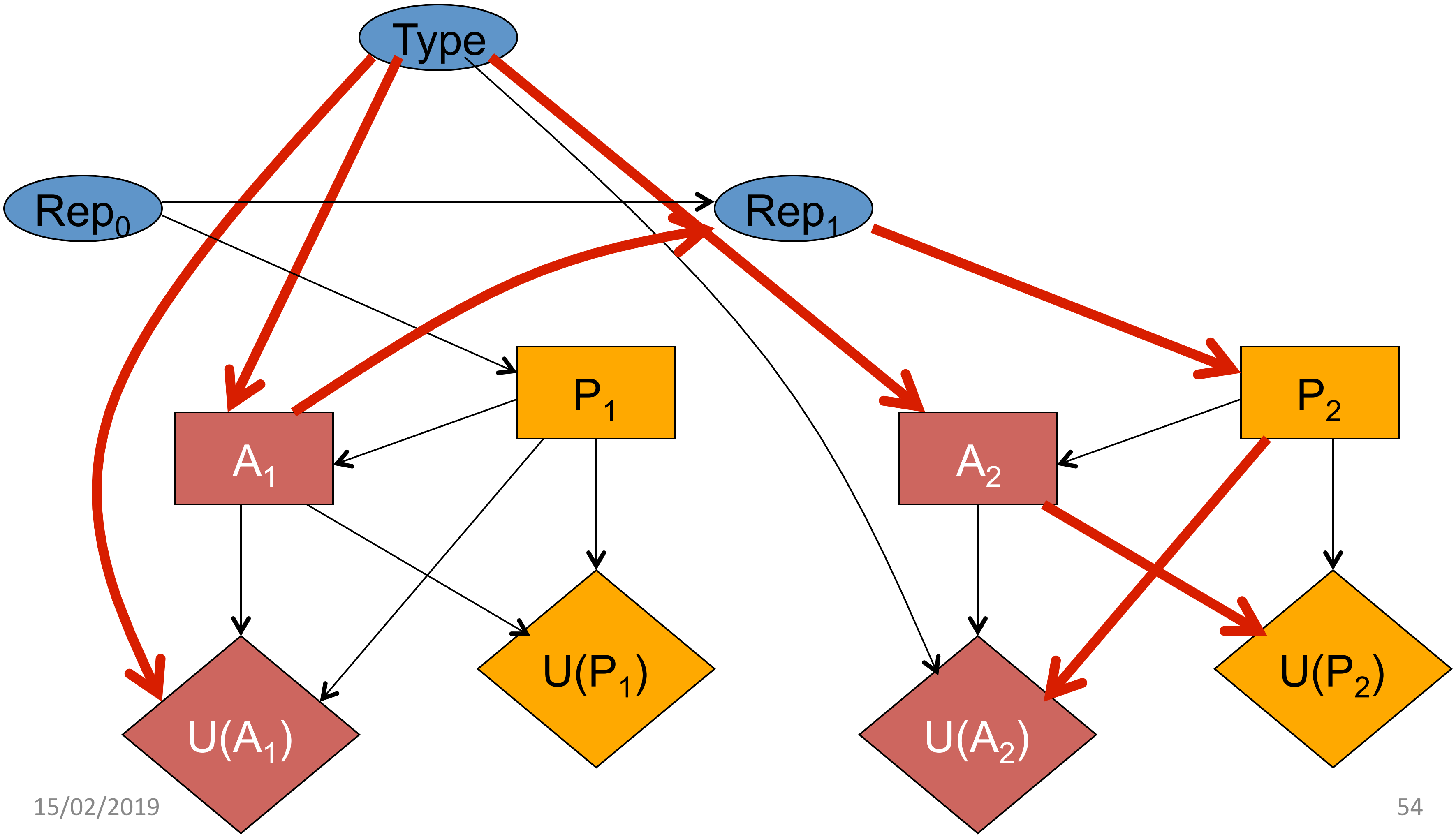
Signaling (A_1 signals Type to P_2)



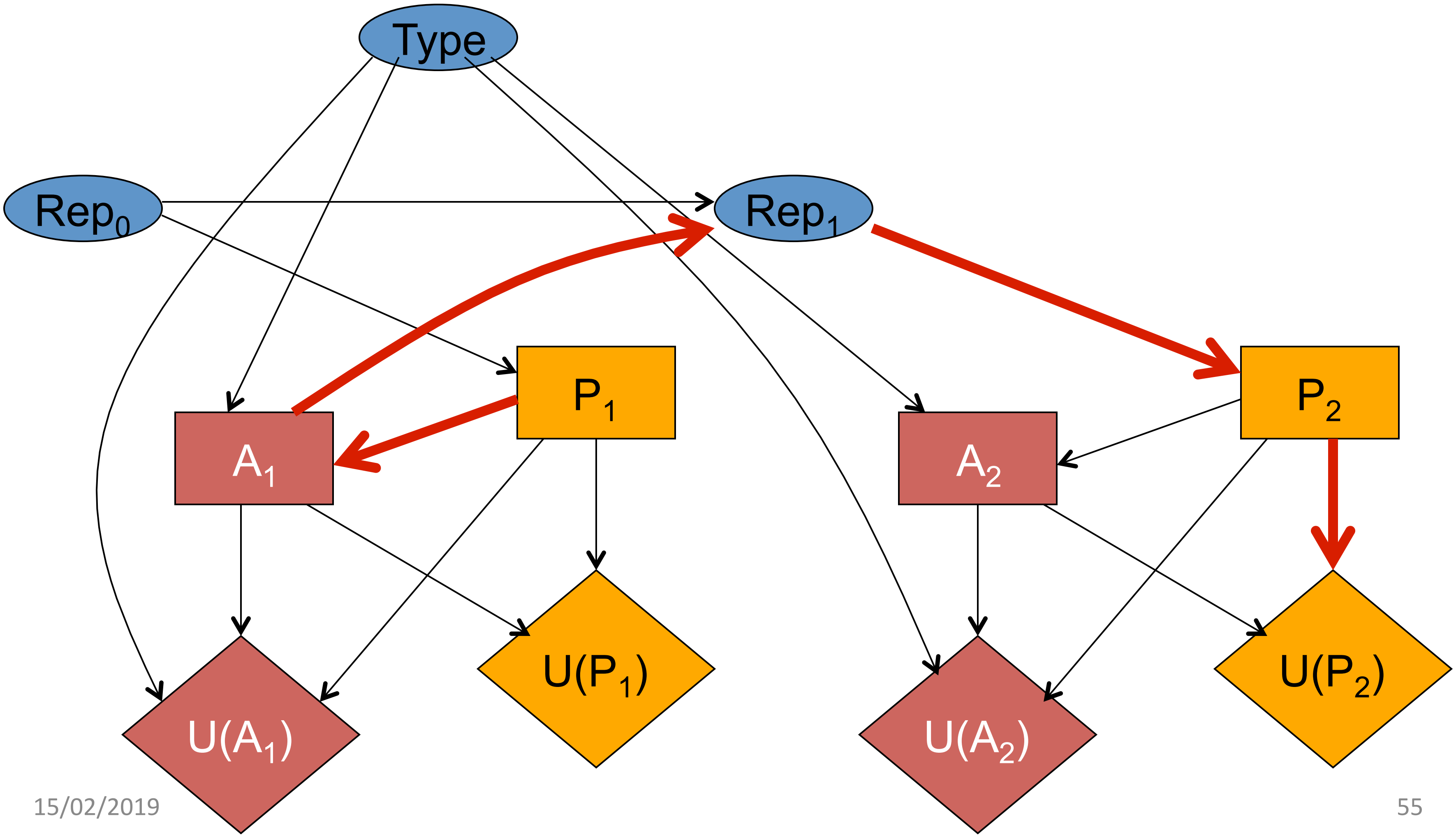
Signaling (A_1 signals Type to P_2)



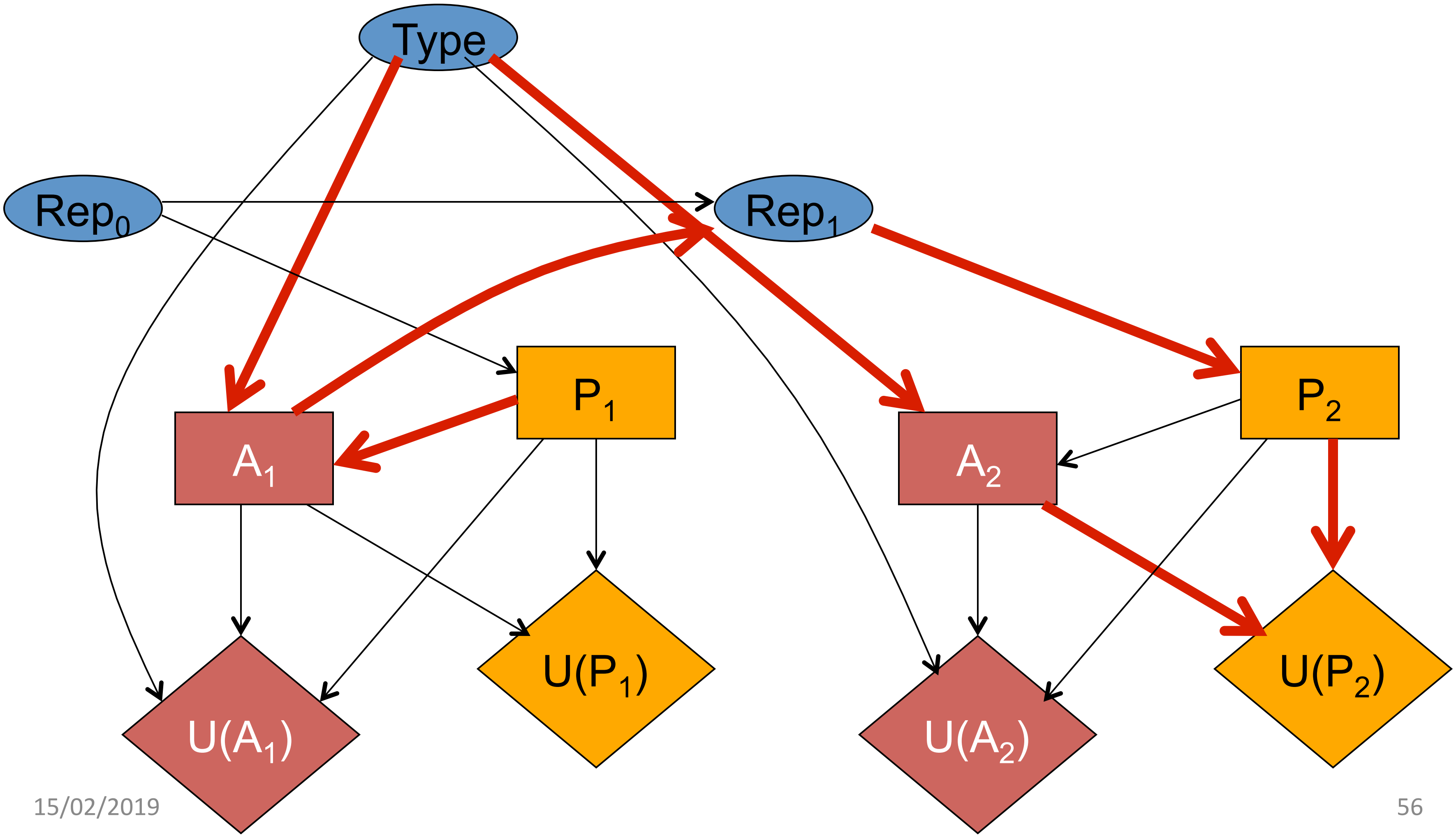
Signaling (A_1 signals Type to P_2)



Revealing/Denying (P_1 reveals Type to P_2)



Revealing/Denying (P_1 reveals Type to P_2)



References

Some of the examples and related content are from:

- R.D. Luce, H. Raiffa, Games and Decisions, Dover 1985.
- A. Karlin, Y. Peres, Game Theory, Alive, AMS Press 2017.
- Tutorial at IJCAI 2003 by Prof Peter Stone, University of Texas