Structure and Synthesis of Robot Motion

Motion Synthesis with Strategic Considerations II

Subramanian Ramamoorthy
School of Informatics

27 February, 2012

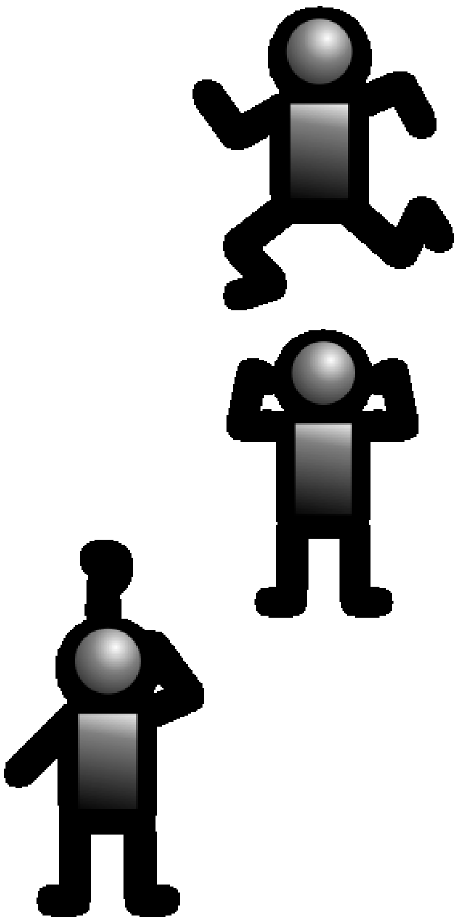
Acknowledgement

Parts of the content of this lecture are based on a tutorial at AAAI 2008 by Y. Gal and A. Pfeffer

For these parts, I have used their original slides, with a bit of editorial additions to fit within a lecture.

Multi-agent Decision-making

- Two sources of uncertainty
  - the environment
  - other agents
- Multi-agent decision problem, or game, includes
  - strategies
  - outcomes
  - utilities
- A solution concept for a game includes a strategy *profile* for all agents.



What is a Game ?

A game includes a set of agents $\mathbf{N} = \{1, \ldots, n\}$.
For each agent i, includes a set of strategies $\mathbf{S}_i$.

Joint strategy profile $(s_1, \ldots, s_n)$ determines outcome of game, where $s_i \in \mathbf{S}_i$.

Payoff function, $u_i : \mathbf{S}_1, \ldots, \mathbf{S}_n \rightarrow R$
represents utility for i given $(s_1, \ldots, s_n)$

Let $\mathbf{S}_{-i} = \mathbf{S}_1 \times, \ldots, \times \mathbf{S}_{i-1} \times \mathbf{S}_{i+1} \times, \ldots, \times \mathbf{S}_n$ be
the joint set of strategies for all players other than i.

Normal Form Representation: The Prisoners' Dilemma

Bob

Alice

	C	*D*
C	*(-1,-1)*	*(-5,0)*
D	*(0,-5)*	*(-3,-3)*

- Row player is Alice; column player is Bob. Values in cell *(C,D)* denotes payoff to Alice when playing *C* and to Bob when playing *D.*
- A **dominant** strategy is one which is best for an agent regardless of other agents' actions.
- The dominant strategy for both players in the prisoners' dilemma is to defect *(D,D)*.

Battle of the Sexes

- Row player is Alice; column player is Bob.
  - Alice prefers watching a football match (*FM*) over going to the ballet (*B*); conversely for Bob.
  - Both players do not like to mis-coordinate.
- Entry *(FM, B)* denotes payoff to (Alice, Bob) when Alice goes to *FM* and Bob goes to *B*.

Bob

Alice

	FM	*B*
FM	*(2,1)*	*(0,0)*
B	*(0,0)*	*(1,2)*

Battle of the Sexes

- Dominant strategies do not exist for either Alice or Bob.
  - But given Alice's strategy, Bob can choose a strategy to maximize his utility (and similarly for Alice)

Bob

Alice		*FM*	*B*
	FM	*(2,1)*	*(0,0)*
	B	*(0,0)*	*(1,2)*

Battle of the Sexes

- Dominant strategies do not exist for either Alice or Bob.
  - But given Alice's strategy, Bob can choose a strategy to maximize his utility (and similarly for Alice)

Bob

Alice

	FM	*B*
FM	***(2,1)***	*(0,0)*
B	*(0,0)*	***(1,2)***

Nash Equilibrium

- A strategy profile $\mathbf{s}^* = (s_1, \ldots, s_n)$ is a Nash equilibrium if no player has the incentive to deviate from its assigned strategy.

- Formally, for every player i and $s_i \in \mathbf{S}_i$,

$$u_i(s_i^*, \mathbf{s}_{-i}^*) \geq u_i(s_i, \mathbf{s}_{-i}^*)$$

Bob

Alice

	FM	*B*
FM	*(2,1)*	*(0,0)*
B	*(0,0)*	*(1,2)*

Matching Pennies

- Alice and Bob can each turn a penny to "heads" or "tails". The payoffs depend on whether Alice and Bob coordinate. The game is zero sum.

Bob

		H	*T*
Alice	*H*	*(1,-1)*	*(-1,1)*
	T	*(-1,1)*	*(1,-1)*

Mixed Strategies

For each player i, a mixed strategy profile defines a probability $\sigma_i(s_i)$ for each pure strategy s_i.
Let σ be a mixed strategy profile for all players.

The expected utility for i given σ is

$$u_i(\sigma) = \sum_{\mathbf{s} \in \mathbf{S}} \prod_j \sigma_j(s_j) \cdot u_i(\mathbf{s})$$

Mixed Strategy Equilibrium

σ is a Nash equilibrium if no player has the incentive to deviate from its assigned strategy.

Formally, for every player i and $\sigma_i \in \Delta \mathbf{S}_i$,

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*)$$

Theorem (Nash 50) Every finite game has a mixed strategy equilibrium.

Mixed strategy equilibrium for Matching Pennies: Alice and Bob choose heads and tails with probability 0.5.

Sequential Decisions

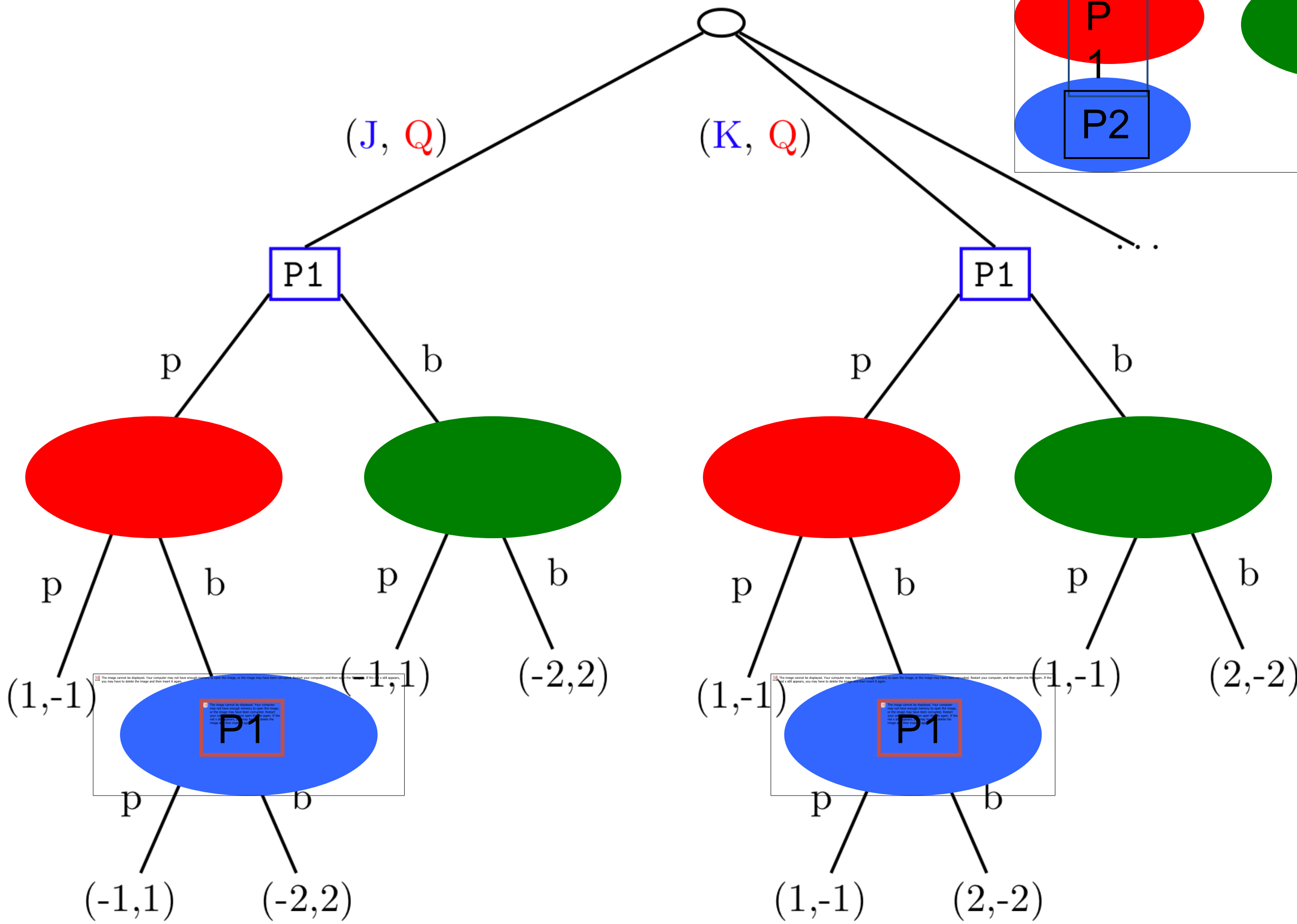
- Normal form represents situation where agents make simultaneous decisions
- What happens when players make decisions sequentially?
- We need to be able to represent situations in which different agents have different information

- **Extensive form games:** like single agent decision trees, plus information sets

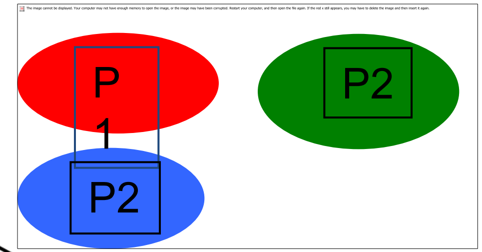
“Constrained” Poker [Kuhn 1950]

- Two players (P1,P2) each given £2
- Three cards in the deck: K, Q, J
- All players put £1 in the pot and pick a card, visible only to themselves.
- P1 bets £1 or passes;
- P2 bets £1 or passes;
- if P1 passes and P2 bets
- P1 can bet its £1 or pass.
- If both players bet (or pass), player with higher card wins £2 (or £1).
- If one player passes and the other bets, the betting player wins £1

Partial Extensive Form Game Tree for Kuhn Poker



Information Sets



Solution Algorithms: Normal Form Game

- Exact solutions:
- Two player zero-sum games
  - Can be solved by a linear program in polynomial time (in number of strategies)
- Two player general-sum game
  - Can be solved by a linear complimentary program (exponential worst-case complexity) [Lemke-Howson '64]
- Approximate solutions for multi-player games:
  - Continuation and triangulation methods [Govindan and Wilson '03, McKelvey & McLennan '96]
  - Search [Porter, Nudelman and Shoham '05]
- Off-the-shelf packages

  Gambit, Game tracer

Solution Algorithms: Extensive Form Games

- Two player perfect information zero-sum game
  - can be solved by minimax search (with alpha-beta pruning)

- Two player perfect information general-sum game
  - Can be solved using backward-induction

- Two player imperfect information general-sum game
  - Can be solved using sequence form algorithm

    [Koller, Megiddo, von Stengel '94].

An Ultimatum Game Example

- Two players: Proposer and responder player.
- Proposer can offer some split of 3 coins to Responder.

An Ultimatum Game Example

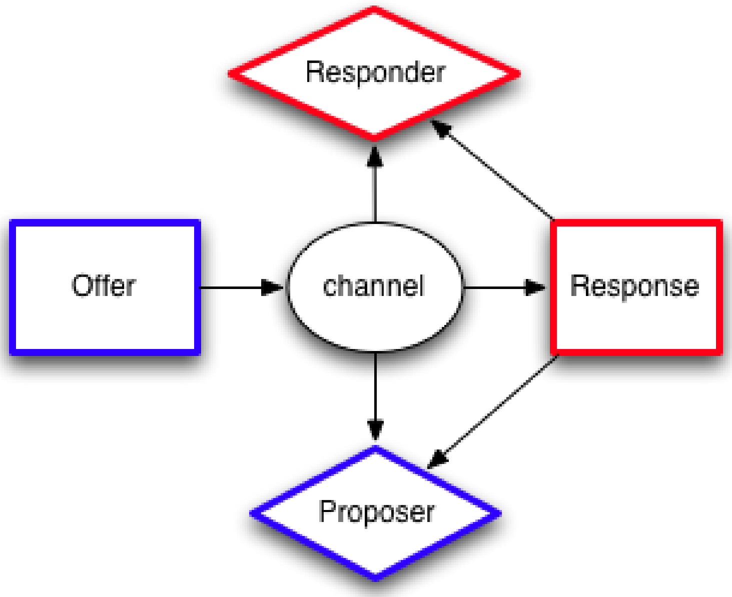
- Proposer can offer some split of 3 coins to Responder. If Responder accepts, offer is enforced; if Responder rejects, both receive nothing.

An Ultimatum Game Example

- Proposer can offer some split of 3 coins to Responder. If Responder accepts, offer is enforced; if Responder rejects, both receive nothing. Offer may be corrupted and set to (1,2) split (proposer/responder) by noisy channel with 0.1 probability.

Multi-agent Influence Diagrams [Milch and Koller '01]

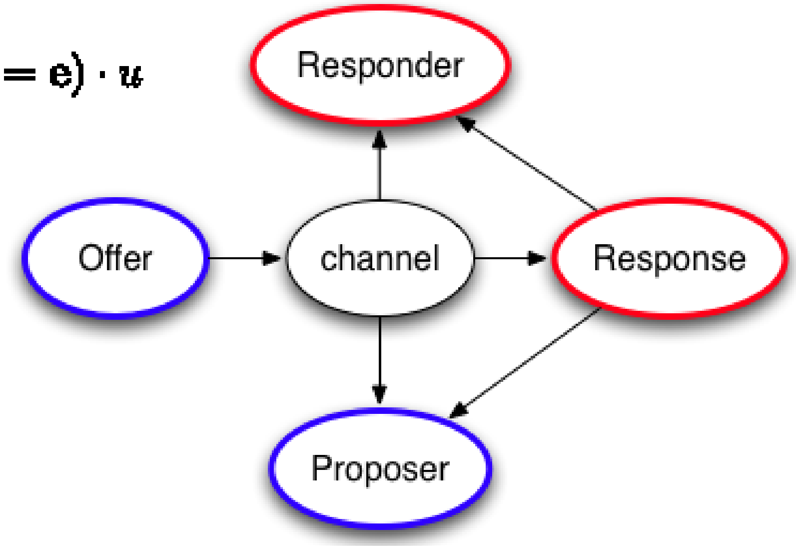
- Rectangles and diamonds represent decisions and utilities associated with agents; ovals represent chance variables.
- A strategy for a decision is a mapping from the informational parents of the decision to a value in its domain.
- A strategy profile includes strategies for all decisions.



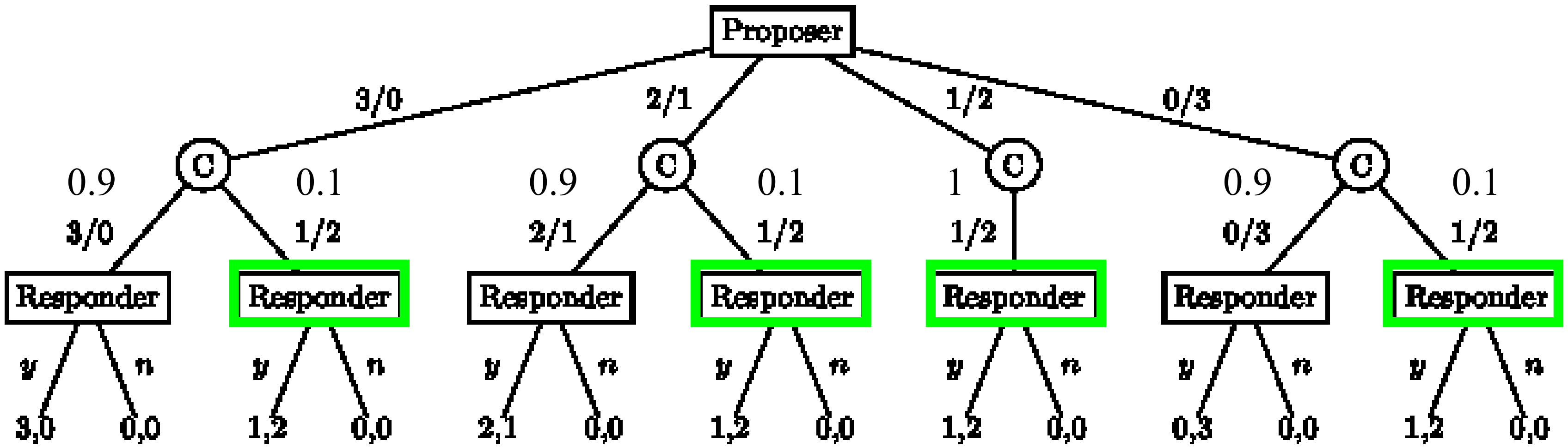
Computing Expected Utility in MAIDs

Let α be an agent, and ***s*** be a strategy profile for all decisions in a MAID. Let P^s be the distribution over the BN that implements ***s*** in a MAID. Let ***U*** be the set of utility nodes for α, and ***E*** be evidence nodes. The utility for α given ***s*** and evidence ***E*=e** is

$$U^{\mathbf{s}}(\alpha \mid \mathbf{E}=\mathbf{e}) = \sum_{U \in \mathbf{U}} \sum_{u \in \mathrm{Dom}(U)} P^{\mathbf{s}}(U = u \mid \mathbf{E} = \mathbf{e}) \cdot u$$



Conversion to Extensive Form Game



Information set for responder

MAID Equilibrium

A strategy profile $\mathbf{\Theta}$ in the MAID is a Nash equilibrium if for any decision D_i belonging to agent α, we have

$$\theta_i(. \mid \mathbf{pa}_i) \in \text{argmax}_{\theta_i \in \Delta(S_i)} EU^{(\theta_i . \mathbf{\Theta}_{-\mathbf{i}})}(\alpha \mid \mathbf{pa}_i)$$

If the strategy profile $\mathbf{\Theta}$ is a Nash equilibrium of a MAID, then $\mathbf{\Theta}$ is also the Nash equilbirum of the extensive form game.

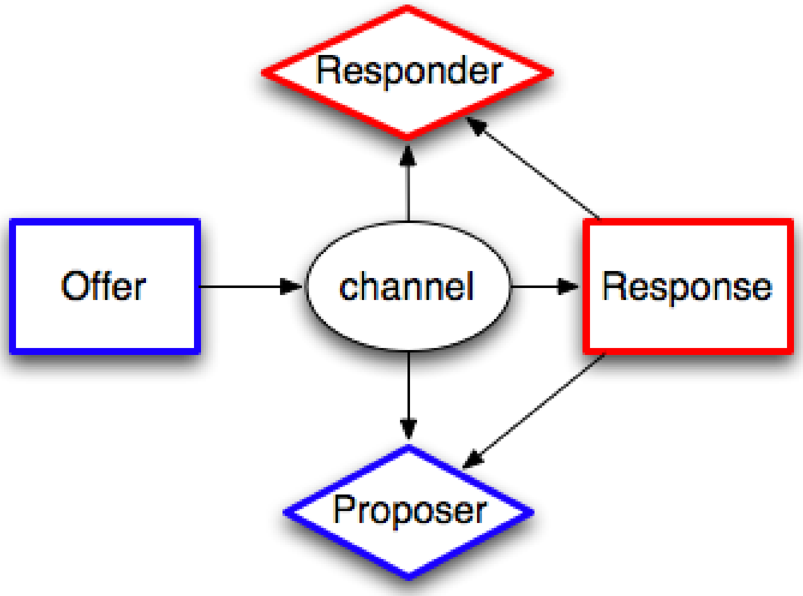
Solving MAIDs

- Naïve solution: Convert MAID to extensive form game, and solve it.
- …but lose the structure of the MAID.

- There is an alternative method that works directly on the MAID graph. We define a new graphical criterion for expressing dependence between decisions.

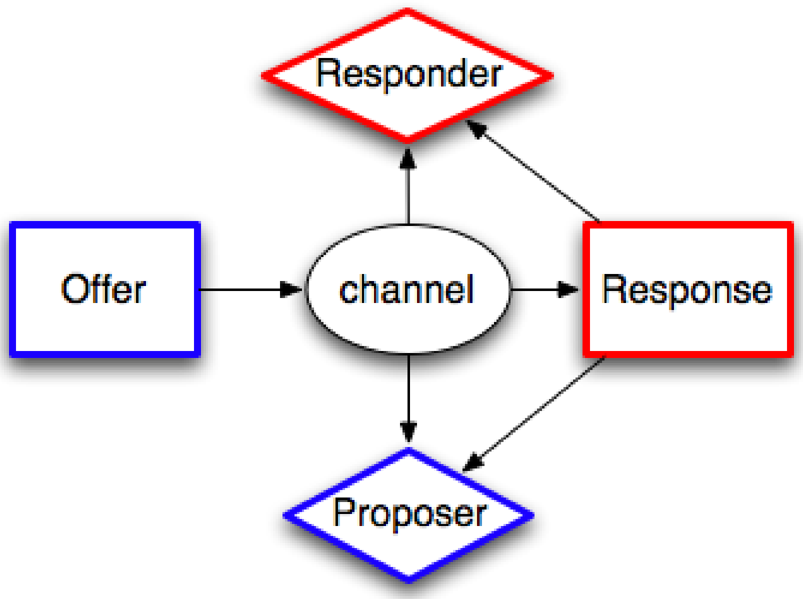
Strategic Relevance

- A decision *D* is strategically relevant to decision D' belonging to some agent if its utility depends on the strategy for D.
- Strategic relevance is a relation that holds between any two decisions in the MAID.



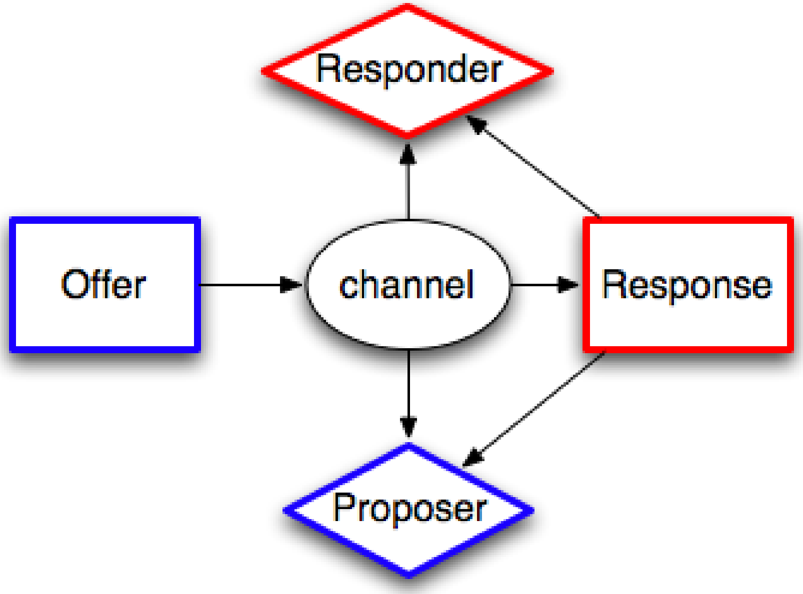
Strategic Relevance: Example

- Given a strategy for the responder
  - accept splits (1,2),(0,3).
- there exists an optimal strategy for the proposer
  - offer (1,2) split.
- Conclusion: Proposer could do well if it knew the responder’s strategy.



Strategic Relevance: Example

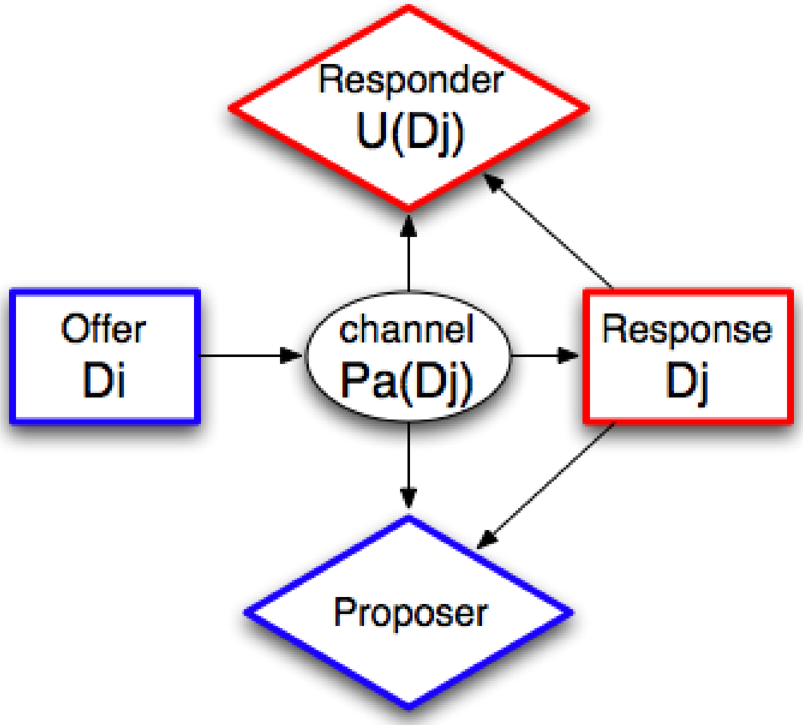
- Given strategy for Proposer
  - propose split (2,1).
- The optimal strategy for responder is
  - agree to beneficial offer reported by channel.
- The proposer's action affects the channel.
- Responder cares about proposer's *action*, but not its *strategy*.



Responder's strategy is relevant to Proposer
Proposer's strategy is not relevant to Responder

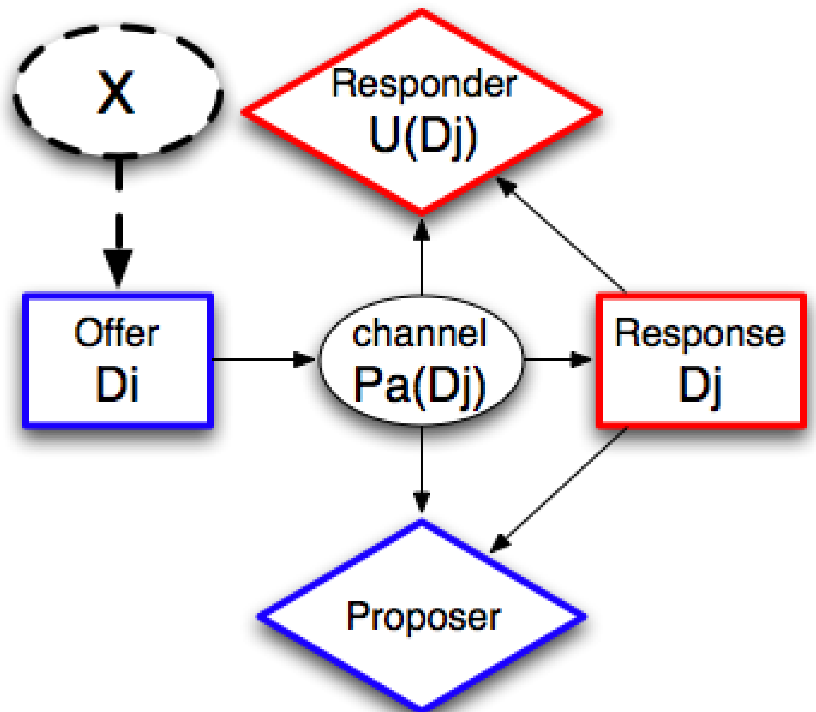
S-Reachability: A Graphical Criterion for Relevance

- Decision nodes D_j , D_i
- Informational parents for D_j, denoted ***Pa(Dj)***.
- Utility node for agent that owns *Dj*, denoted *U(Dj)*.



S-Reachability: A Graphical Criterion for Relevance

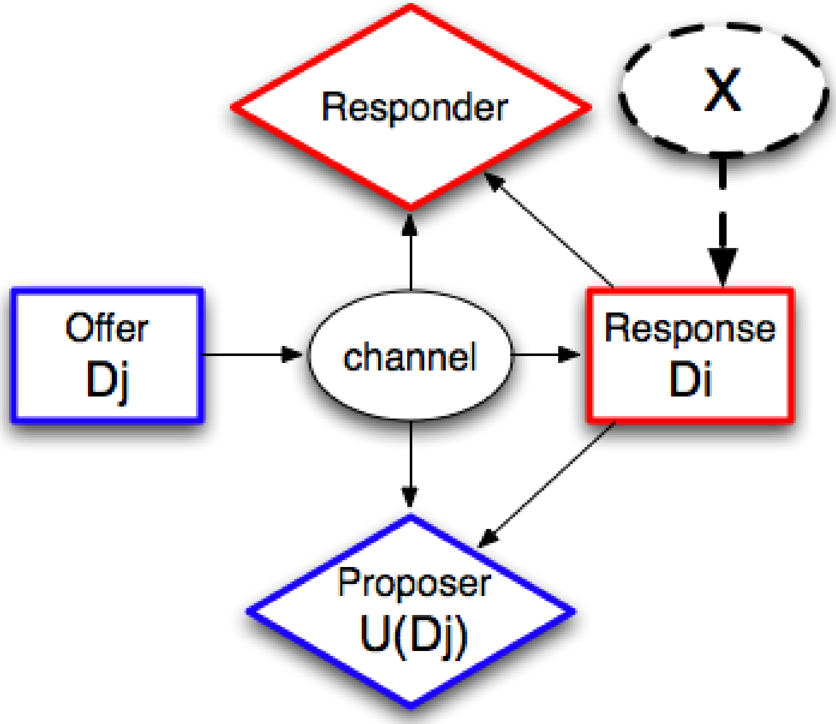
- A decision D_i is **S-Reachable** from D_j if
  - add a new informational parent *X* to *Di.*
  - the path from *X* to *U(Dj)* is *not* blocked given ***Pa(D$_j$)***, and D_j.



The parent of D_j is the channel.
The path from *X* to Responder is blocked by channel.
So, Offer is **not** S-reachable from Response.

S-Reachability: A Graphical Criterion for Relevance

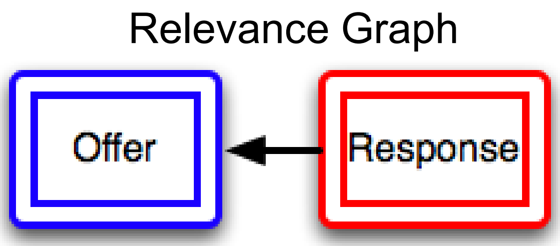
- A decision D_i is **S-Reachable** from D_j if
  - add a new informational parent *X.*
  - the path from *X* to *U(Dj)* is *not* blocked given ***Pa(D$_j$)***, and D_j.



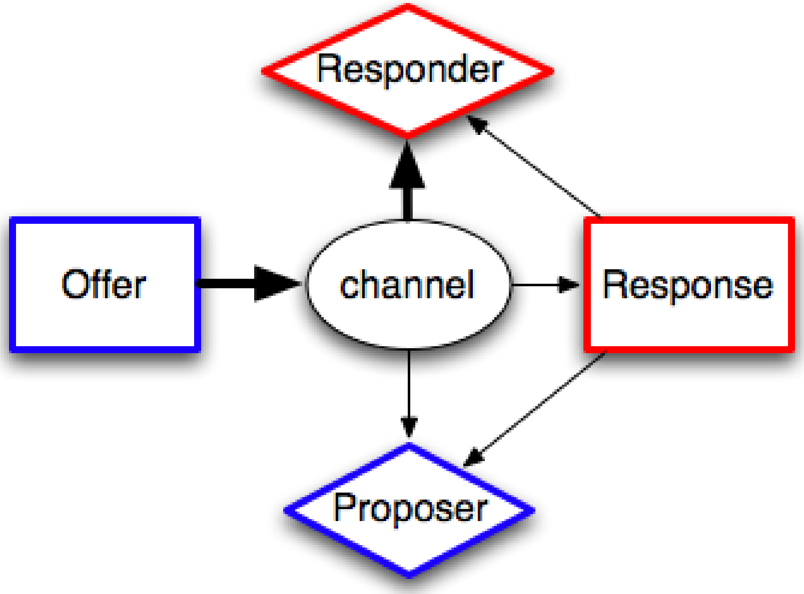
Offer does not have informational parents.
The path from *X* to Proposer is not blocked by Offer.
So, Response is S-Reachable from Offer.

Relevance Graph

- Nodes represent decisions; an edge from D1 to D2 means that D1

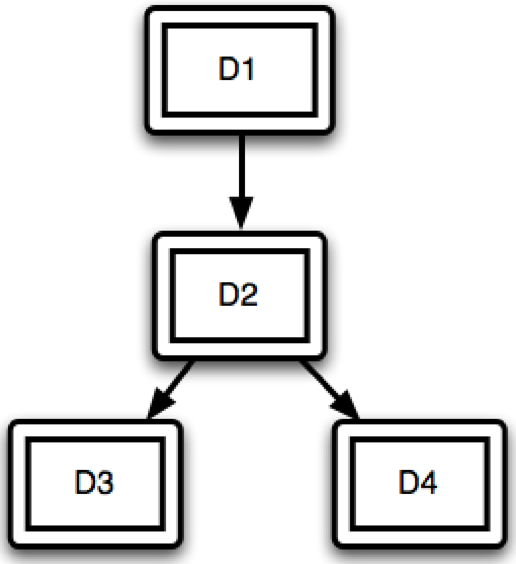


- Relevance graph for Ultimatum game:



MAID Alg.: Acyclic Relevance Graph

- Traverse the decisions by their topological order in relevance graph.
- All decisions that are not relevant to current decision can be implemented by chance nodes with uniform dist.
- Best-response strategies for decisions that are relevant to current decision already exist, and are implemented as chance nodes.
- Implement all utility nodes for these decisions as chance nodes.
- MAID is now ID. Can solve ID and extract the best-response strategy for the current decision.



Back to Robotics Context: Applications of Societal Importance

NurseBot

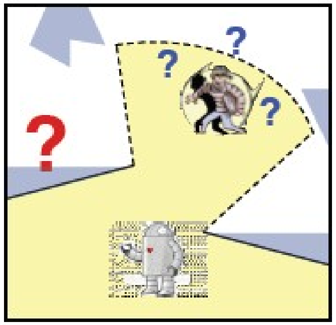
Ubiko, a hospital robot guide
Aizu Central Hospital, Japan

Essential Sub-problem

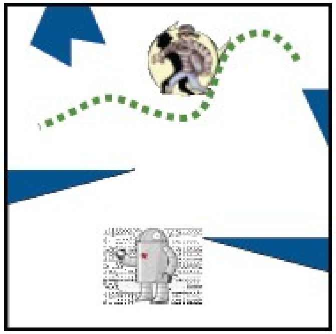
[Source: D. Hsu's talk at ICRA 2010 Workshop]

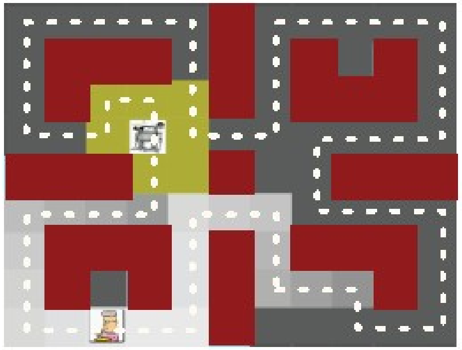
Two Ends of a Spectrum

local planning



global planning

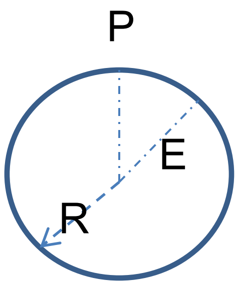




Local and Greedy Interaction

At this level, what is the principle for interaction?

A Classic: The Lady in the Lake



A ‘pursuit-evasion’ game:

- Lady E is swimming in a circular pond with max. constant speed $v_2 < 1$. She can instantaneously change direction.
- Sir P, who can’t swim, wishes to intercept the lady on shore and can run at *unit* speed (also can change direction instantly)
- E would eventually like to leave the lake without meeting P (E is faster than P on land, so we only solve the lake problem)
- E wishes to stay well away from P, maximizing angular distance PE (from centre) at terminal position, and vice versa

Lady in the Lake Problem

The kinematics of this game are as follows:

$$\dot{\theta} = \frac{v_2 \sin u_2}{r} - \frac{v_1}{R}$$
$$\dot{r} = v_2 \cos u_2$$

The cost function is simply $|\theta(T)|$

The solution (HJI equation) for this problem is as follows:

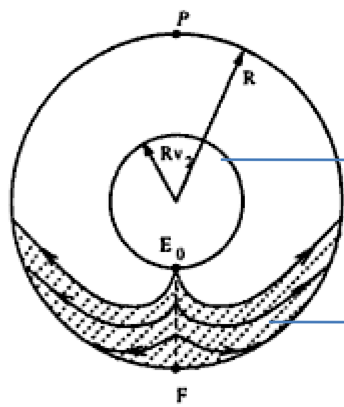
$$\min_{u_1} \max_{u_2} \left\{ \frac{\partial V(\theta, r)}{\partial r} v_2 \cos u_2 + \frac{\partial V(\theta, r)}{\partial \theta} \left(\frac{v_2 \sin u_2}{r} - \frac{v_1}{R} \right) \right\} = 0$$

$\therefore$ (after some algebra...)

$$u_1^* = \operatorname{sgn}\left(\frac{\partial V}{\partial \theta} \right)$$

$$\sin u_2^*(t) = \frac{R v_2}{r(t)} \operatorname{sgn} \theta(T)$$

Lady in the Lake - Solution



Simply max |θ|

HJI sol.

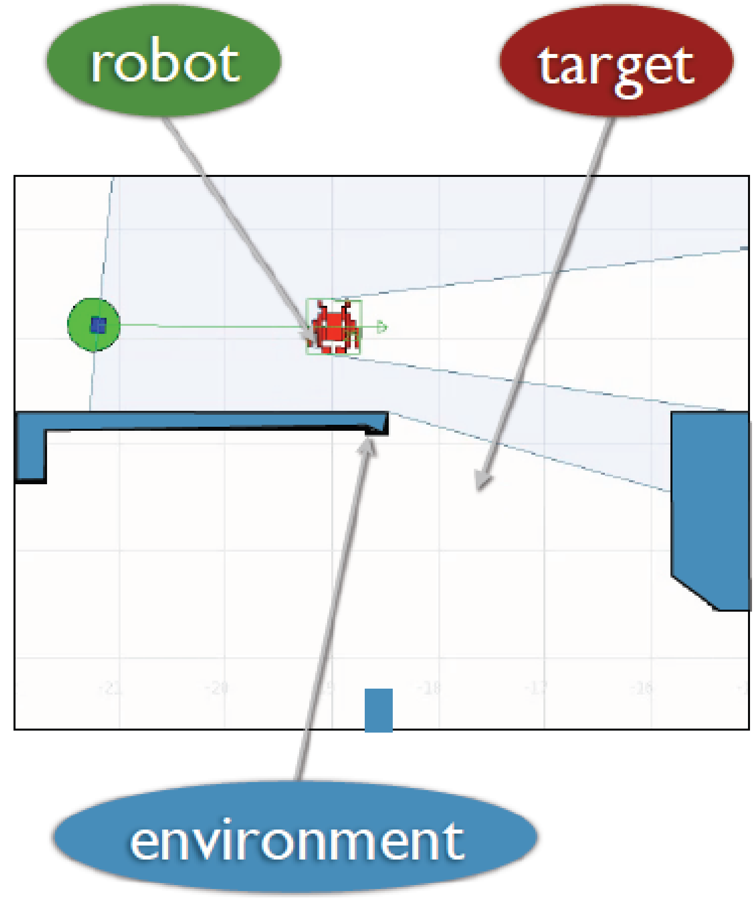
Optimal trajectories in the relative space.

Inside the Rv_2 circle, geometric considerations yield :

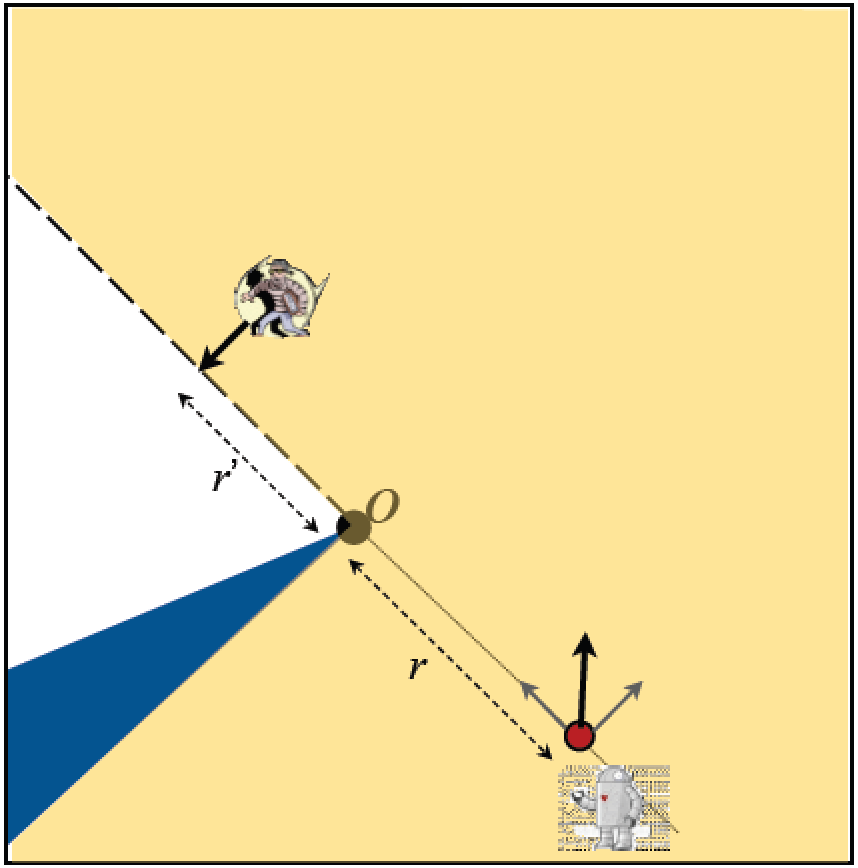
$$|\theta(T)| = \pi + \arccos v_2 - \frac{1}{v_2}\sqrt{(1 - v_2^2)}$$

Problem with Local and Greedy

- Simple visual servo control works well only if there are no or few obstacles.
- Obstacles obstruct the robot's
  - mobility
  - visibility

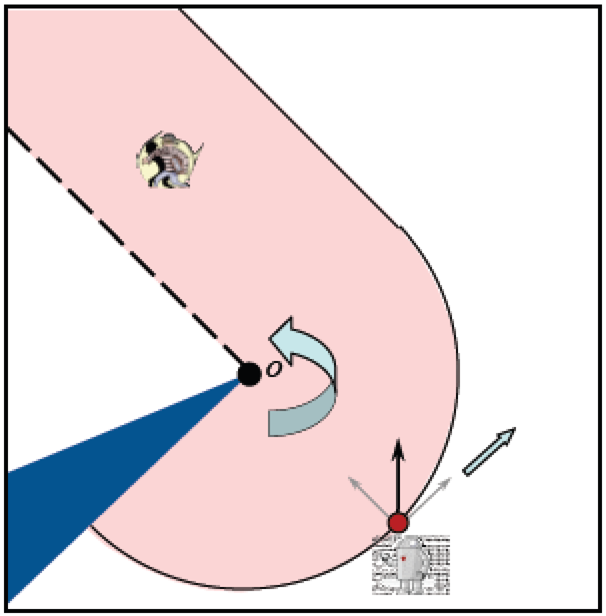


What are the Robot's Options?

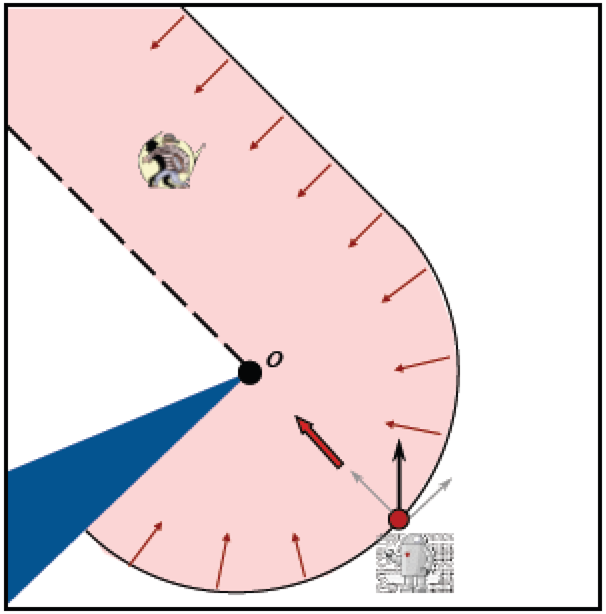


An Approach: Reason about Vantage Zones

swing



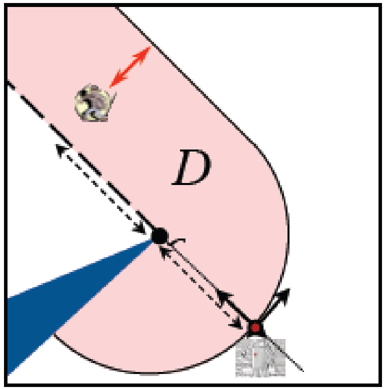
approach



Set up Optimization Problem over Risk

- Risk w.r.t. a single occlusion edge i

$$\varphi_i = \frac{Dist(\text{target}, D)}{RelVel(\text{target}, D)} = \frac{r - e}{v_r + (r'/r)/v_n - v'_n}$$

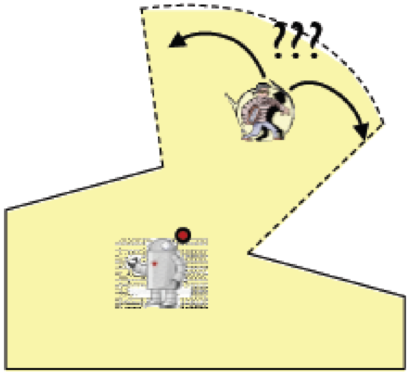


- Risk w.r.t. to multiple occlusion edges

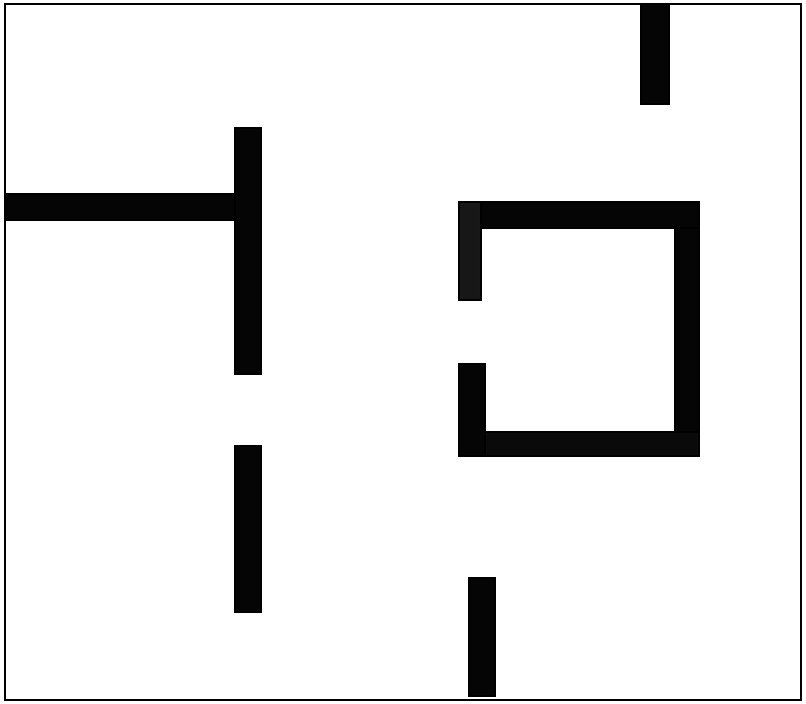
$$\varphi = \sum_i p_i \varphi_i$$

- Choose the action to minimize the risk

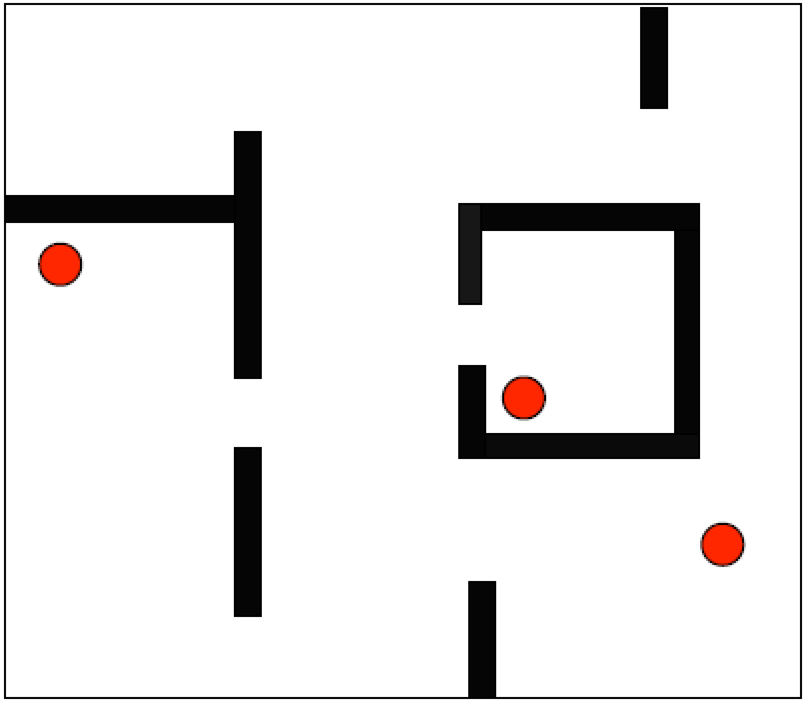
$$\nabla \varphi$$



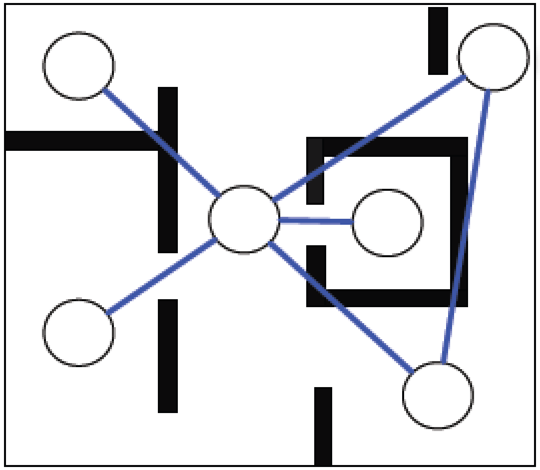
More General P-E Scenarios



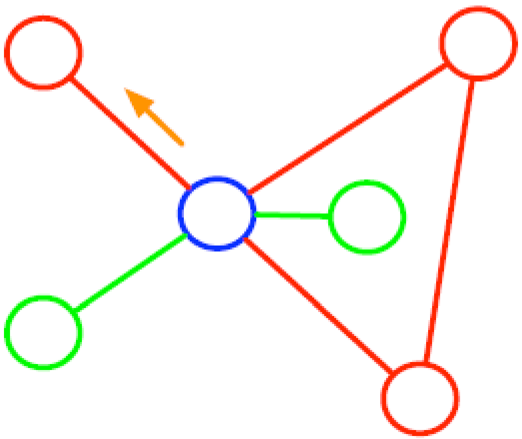
Find Intruders



Use Graph Abstractions



Perform Edge Search

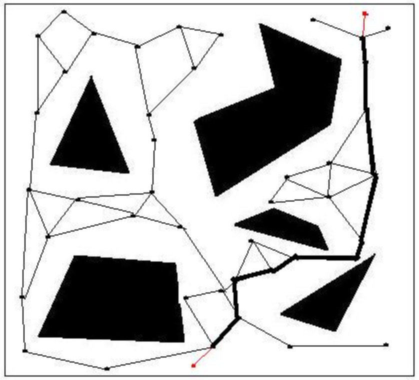


T. Parsons, *Pursuit-evasion in a graph*, 1976
L. Barriere et al. *Capture of an intruder by mobile agents*, 2002

Games on Graphs

Why is this a good model?

- Recall how we originally abstracted c-spaces for motion synthesis
- We could now play games over these structures



Many games on graphs with colourful names:

- Cops and robbers, hunters and rabbits, etc.
- They are models for **search** over graphs and discrete structures
- So, differences are in information structure and assumptions regarding capabilities

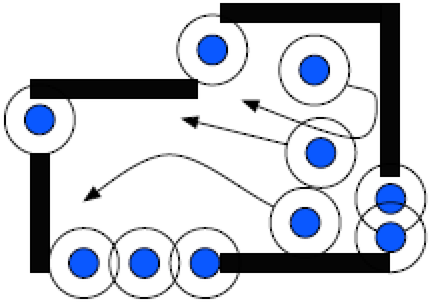
Graph-Clear Strategy: Clearing and Blocking

- robots may remove contamination applying two operations:
  - clear: remove all contamination from a vertex
  - block: prevents an intruder from passing unobserved through and edge
- an edge is blocked if a block operation is applied
- a cleared vertex/edge becomes recontaminated if there exist a path to a contaminated vertex/edge
- intruders move at unbounded speed and have full knowledge of the pursuers' strategy
  $\Rightarrow$ recontamination happens as soon as it is possible

Cost for Graph Clear

Robots have limited sensing capabilities:

- blocking an edge may need more than one robot
  $\Rightarrow$ cost of a block is the weight of a vertex $w(v)$
- clearing a vertex may need more than one robot
  $\Rightarrow$ cost of clearing is the weight of an edge $w(e)$
- Hypothesis: all edges emanating from a vertex must be blocked while clearing



$$s(v) = w(v) + \sum_{e_j \in edges(v)} w(e_j)$$