On the Reasoning Patterns of Agents in Games

Avi Pfeffer
Harvard School of Engineering and Applied Sciences

Ya'akov Gal
MIT CSAIL

Abstract

What reasoning patterns do agents use to choose their actions in games? This paper studies this question in the context of Multi-Agent Influence Diagrams (MAIDs). It defines several kinds of reasoning patterns, and associates each with a pattern of paths in a MAID. We asks the question, what reasoning patterns have to hold in order for an agent to care about its decision? The answer depends on what strategies are considered for other agents' decisions. We introduce a new solution concept, called well-distinguishing (WD) strategies, that captures strategies in which all the distinctions an agent makes really make a difference. We show that when agents are playing WD strategies, all situations in which an agent cares about its decision can be captured by four reasoning patterns. We furthermore show that when one of these four patterns holds, there are some MAID parameter values such that the agent actually does care about its decision.

Introduction

What reasoning patterns do agents use to choose their actions in games? A reasoning pattern is a chain of arguments that lead to and explain a decision. For example, modus ponens is a reasoning pattern in logic. One general reasoning pattern is the maximum expected utility principle. But that is a blunt instrument; it provides one generic explanation for all decisions. If we look deeper, we can provide a more nuanced answer, that may be more satisfying. For example, an agent might take an action in order to reveal information to another agent, thereby influencing behavior. This paper formally investigates and analyzes reasoning patterns in games.

There are a number of reasons it is important to understand reasoning patterns in games. One is that if we are designing an agent to play against another agent, it is important to model and understand the way the other agent works. This is particularly so when the other agent is a human. In order to understand human play, we need to understand the reasoning patterns that the human might use to make decisions. In particular, people might prefer some reasoning patterns to others. A second reason is explanation. If we are to design agents that work with humans, or especially agents that act as proxies for humans, we will want to explain their strategies and behavior to the people they work with. Using reasoning patterns will make for better, more intuitive explanations. Without reasoning patterns, we would have to say something opaque like "I computed the Nash equilibrium or best response strategy and it told me to do this". With reasoning patterns we could say "I knew that my action would have an effect on Bob's utility and he would have to react to what I did, so I chose my action to get him to react in the best way possible for me". A third reason is computational. By identifying reasoning patterns, we can also identify situations in which reasoning patterns do not hold. These can be used to simplify a model for the purpose of solving it. We identify criteria based solely on the graphical structure of a game through which the game can be simplified. A fourth reason is knowledge engineering. When we are designing a model, it is useful to know what reasoning patterns are possible in the model. For example, when designing a Bayesian network, it is important to understand the role of the "explaining away" reasoning pattern, in order to construct a network with the right structure. Similarly understanding the reasoning patterns of games will help to construct good models of games. A fifth reason is belief update. Suppose B observes A's action, and A has access to C which is important to B. How should B update his beliefs about C? Answering this question requires understanding what motivations A might have to report the value of C.

Our study is pursued in the context of Multi-Agent Influence Diagrams (Koller & Milch 2001), a graphical representation language for strategic situations. In addition to their compactness, MAIDs provide a natural description of the story of a strategic situation, in particular how different variables and decisions influence each other. Thus they provide the ideal context for studying reasoning patterns. We identify several basic reasoning patterns and associate each with a pattern of paths in a MAID.

The question naturally arises, what reasoning patterns will agents actually use? The answer to that depends on what strategies we allow for the other agents. If we allow other agents to adopt arbitrary strategies, then any situation in which an agent can either influence its own outcome directly, or influence another agent that influences its outcome, yields a reasoning pattern. However, the pattern may be only weakly justified. If the agent is taking an action to influence another agent, there should be a good reason the other agent should pay attention to the first agent's action. To capture this, we define a solution concept called well-distinguishing (WD) strategies, in which agents only make distinctions where those distinctions really make a difference. So, if the second agent behaves differently for different actions of the first agent, then it has good reason for doing so, and the first agent should take this into account. In essence, WD strategies disallow arbitrary behavior, that makes arbitrary distinctions that don't matter.

It turns out that four reasoning patterns are particularly important. We prove a completeness result, namely that these four patterns characterize all situations in which an agent might be motivated to choose an action, when the other players are playing WD strategies. A natural converse question is whether these reasoning patterns are really patterns that matter, i.e., whether in all cases where one of the reasoning patterns holds an agent actually cares about its decision. The answer to this question in general is no, but we are able to show a weaker converse, namely that in any MAID in which one of these four patterns is present, there is some assignment of MAID parameter values such that the agent is motivated to take an action. Thus the four reasoning patterns are a perfect characterization of the situations in which an agent cares about its decision (relative to WD strategies), based on graphical considerations alone.

Preliminaries

We assume that the reader is familiar with Bayesian networks (BNs) (Pearl 1988), and the concept of d-separation in particular. Multi-Agent Influence Diagrams (MAIDs) (Koller & Milch 2001) are variants of influence diagrams (Howard & Matheson 1984). A *MAID* is a directed acyclic graph containing three kinds of nodes: *chance* nodes denoted by ellipses, *decision* nodes denoted by rectangles, and *utility* nodes denoted by diamonds. Each chance node has an associated conditional probability distribution (CPD), as in a BN. A utility node has an associated determinsitic function from values of its parents to the real numbers. The parents of a decision node represent information that is known to the decision maker at the time of making the decision, and are called *informational parents*. Each decision and utility node is associated with a particular agent. For simplicity, in this paper we assume that each agent has one decision. This is for presentation purposes only. The case where an agent A has more decisions can be simulated simply by creating one agent A_i for each decision D_i of A, giving all the A_i the same utility function, and making sure all earlier decisions of A, as well as all information available for them, are available to A_i when making its decision. Also, we assume in this paper that each node has at least two values. We will use D_A to denote the decision node belonging to A, U_A to denote one utility node belonging to A, and $\mathbf{U}_A$ to denote the set of all utility nodes belonging to A.

A *strategy* σ_A for a node D_A is a function from values of the parents of D_A to probability distributions over the values of A. We use $\sigma_A(\mathbf{q})$ to denote this distribution for some value $\mathbf{q}$ of the parents of D_A, and $\sigma_A(\mathbf{q})(d_A)$ to denote the probability of a particular value d_A. A *strategy profile* $\boldsymbol{\sigma}$ consists of a strategy for every decision in a MAID. We use $\boldsymbol{\sigma}_{-A}$ to denote the strategies of all agents other than A in $\boldsymbol{\sigma}$. Given a strategy profile $\boldsymbol{\sigma}$, we can replace each decision node D_A in a MAID with a chance node whose CPD implements σ_A. This defines a BN, which defines a probability distribution over the values of all chance and utility nodes. We write $P^{\boldsymbol{\sigma}}(x)$ to denote the probability of x in the BN defined by $\boldsymbol{\sigma}$. We write $\text{EU}^{\boldsymbol{\sigma}}[A|\mathbf{c}]$ to denote the expected utility to A, under the distribution defined by $\boldsymbol{\sigma}$, given that variables $\mathbf{C}$ take the values $\mathbf{c}$. If we want to talk about agent A deliberately taking a particular action d_A, when the strategy profile is $\boldsymbol{\sigma}$, we will adopt Pearl's causal notation (Pearl 2000) and write $\text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A), \mathbf{c}]$.

Reasoning Patterns

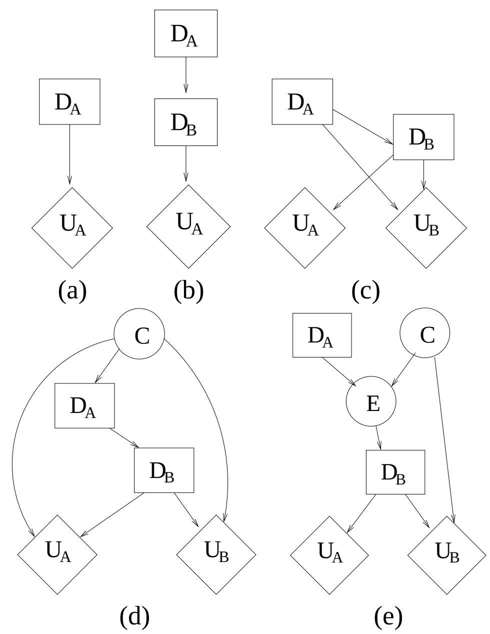

Figure 1: Patterns of reasoning: (a) direct effect; (b) influence for no reason; (c) manipulation; (d) signaling; (e) revealing/denying.

There are a number of different reasoning patterns an agent might use in a game. Here we present some of these patterns. We do not claim that this is an exhaustive list, but these patterns are qualitatively different from each other, and sometimes yield surprising insights.

The first kind of pattern is the most obvious, and also the most common. When an agent's action has a *direct effect* on its utility, obviously it will care about what to do. In MAID terms, this means that there is a directed path from the agent's decision to one of its utility nodes, that does not pass through the decisions of other agents. A canonical figure of this pattern is shown in Figure 1(a). Agent Alice takes a decision D_A that directly affects her utility U_A. This path may be mediated by several chance nodes along the way, but Alice will still have an influence on her outcome. Now, Alice's utility might also be affected by the actions of other agents — this does not change the fact that Alice has a direct effect on her own outcome.

More interesting are games in which an agent cannot affect its own outcome directly, but can *influence* other agents to act favorably toward itself. An example is shown in Figure 1(b). Here agent Bob, who takes decision D_B, affects Alice's utility. Alice's decision is known to Bob. Therefore, when she makes her decision she should consider Bob's response, and choose her action so as to make Bob take the action that is most beneficial to her. However, in this situation there is no particular reason why Bob should take a particular action in response to Alice. Specifically, there is no reason why Bob should take a different action in response to different moves by Alice. Thus, while this is a possible pattern of reasoning for Alice, it is a relatively poorly justified one.

There are several stronger patterns of reasoning, where Alice's strategy has an influence on Bob for good reasons; Bob is actually motivated to change his decision in response to Alice. One such pattern is *manipulation*. The canonical pattern is shown in Figure 1(c). Here, as before, Bob's decision has an effect on Alice's utility, and Alice's decision is known to Bob. In addition, both Alice and Bob have an influence on Bob's utility. All the paths in the graph may be mediated by other nodes. Here, Bob's optimal response may depend on Alice's move. Thus Bob has a motivation to respond differently to different actions of Alice. As a result, Alice is motivated to choose her action so as to elicit the best possible response from Bob.

An example payoff matrix illustrating this situation is as follows. Note that this is not a normal form matrix as commonly appears in game theory. In the normal form, players move simultaneously. Our game is a sequential game in which Alice moves first, and her action is known to Bob. In each entry, the first number is Alice's payoff, and the second is Bob's.

	Bob	
Alice	Left	Right
Up	(1,1)	(2,2)
Down	(1,2)	(2,1)

Here Alice's utility does not depend on her action. However, if Alice plays Up, Bob's best response will be to play Right, leading to a utility of 2 for Alice, while if Alice plays Down Bob will play Left, leading to a utility of 1 for Alice. Thus Alice is motivated to play Up.

More than one reasoning pattern may be active at a time. Consider a variation of Figure 1(c) where there is also an edge from D_A to U_A. Then both direct effect and manipulation may be active. Suppose the game is a sequential Battle of the Sexes game, with the following utilities.

	Bob	
Alice	Left	Right
Up	(1,2)	(0,0)
Down	(0,0)	(2,1)

Here Alice's maximum utility results from a combination of her own action and Bob's action, which depends on hers. Therefore she chooses her optimal action of Down partly because of its direct effect on her own utility, and partly to manipulate Bob to play right. Both reasoning patterns are active.

Manipulation is only one of the ways an agent can be influenced by another agent for strong reasons. It is also possible for one agent to influence another agent without having any direct influence on the second agent's payoffs. One of the ways this can happen is for the first agent to *signal* something she knows about and something the second agent cares about. An example of this comes from the game of bridge. Bridge is a four-player card game where the players are in two teams or "partnerships". In some situations, a player may have only small cards to play, such as the $2\heartsuit$ and $5\heartsuit$. Since the cards are small, they have no chance of winning at any point in the game. The choice of which card to play at a particular point has no direct bearing on the outcome of the partnership. However, a player may use the choice of card to signal to her partner something about her hand; this information might be useful to her partner. For example she may play the $5\heartsuit$ to signal that she has two cards in the heart suit.

The canonical MAID pattern for signaling is shown in Figure 1(d). Here Alice's action is known to Bob, and Bob has an influence on both his own and Alice's outcome. In addition, there is a variable C that is known to Alice at the time she makes her decision. This variable is relevant to Bob, as it influences his outcome. As before, all these paths may be mediated by other variables. Also, the path from C to U_B need not be direct. All that is required is that it not be blocked by a set of variables that will be specified later. In addition, notice that there is a path from C to U_A. This is important. It turns out that an agent has no motivation to signal something it does not care about. This can be understood as follows. Suppose Alice does not care about C. Suppose she communicates C truthfully to Bob. Then Bob will change his action based on what Alice communicates. But in that case, Alice should always communicate the message that causes Bob to play his best action for her. Therefore Bob should not believe Alice's communication. As a result, Alice has no reason to communicate anything in particular. On the other hand, if Alice's utility depends on C, she may want to Bob to change his response based on C, therefore she may want to communicate C to him.

Signaling means one agent communicating something it knows to another agent. It is also possible for an agent to influence another agent by causing it to find out information the first agent *does not know*. We call this pattern of reasoning *revealing/denying*. A canonical MAID pattern is shown in Figure 1(e). Here, Bob cares about a variable C. Alice's decision is a parent of a variable E, which is known to Bob. Alice's action can change Bob's knowledge of C. For example, there is a room with treasure, that may or may not have a tiger in it. Bob can go into the room to get the treasure. Bob will get negative utility if there is a tiger present and positive if not. If Bob gets the treasure, Alice will get a cut (whether or not a tiger was in the room). Alice's possible actions are to open the curtain to the room or not to open the curtain. If Alice opens the curtain, Bob can observe whether there is a tiger present before entering the room. Now, suppose Bob's expected utility, if he does not know whether there is a tiger present, is negative, so Bob's strategy will be to not enter the room. Then it is in Alice's interest to open the curtain, so that Bob will enter the room and get the treasure when no tiger is present. Here Alice is causing information to be revealed to Bob that Alice herself does not know. This pattern is called revealing/denying because it can also work in the other direction. Suppose Bob's strategy when he does not know whether there is a tiger present is to get the treasure. Then it is in Alice's interest not to open the curtain, so that Bob will always get the treasure and not only when there is no tiger. Here Alice is denying information to Bob.

Like in signaling, the path from C to U_B need not be direct. However, this pattern is qualitatively different from signaling, in that there need be no path from C to U_A. The reason is that Alice's action is not a communicative action — it is not reporting the value of a variable. Rather, it is doing something to change the state of the world, i.e. the value of variable E. Alice cannot lie when taking the revealing action; if she opens the curtain Bob will find out whether there is a tiger. Alice cannot deceive Bob into thinking there is no tiger when there is actually a tiger present.

Motivation with Respect to Well-Distinguishing Strategies

We are interested in the question: under what reasoning patterns is an agent motivated to take one action rather than another? This section begins by defining what it means for an agent to be motivated.

Definition 1: Let D_A be a decision node. We say that D_A is *motivated with respect to a class of strategies* if there exist values d_A^1 and d_A^2 for D_A, strategies $\boldsymbol{\sigma}$ in the class, and values $\mathbf{q}$ for the informational parents of D_A, such that

$$\text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A^1),\mathbf{q}] \neq \text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A^2),\mathbf{q}] \quad \blacksquare$$

Intuitively, this means "What I do might make a difference". In answering the question of when an agent is motivated, we need to decide what strategies we are allowing the other agents to play. One possibility is that all strategies are considered. Then the answer is that any situation in which an agent has any influence at all on its outcome is a possible reasoning pattern. More precisely, whenever there is a direct path from D_A to $\mathbf{U}_A$, D_A may be motivated relative to the class of all strategies.

If we place reasonable restrictions on strategies, however, we may obtain more nuanced behavior. One possible such restriction is to say that a strategy does not make arbitrary distinctions. In particular, if two different values of the parents of a decision node have no different impact on the outcome whatever the strategies of other players, the strategy should not make a distinction between those values. In addition, if two possible actions for a decision node have no different impact on the outcome whatever the strategies of other players, the strategy should not prefer one action to the other. Intuitively, an explanation that only involves strategies that make appropriate distinctions will be more convincing than one that does not. This leads to the definition of a well-distinguishing strategy.

Definition 2: Let $\boldsymbol{\Sigma}$ be a set of strategy profiles. A strategy $\sigma_A \in \Sigma_A$ for D_A is *well-distinguishing with respect to* $\boldsymbol{\Sigma}$ if

1. Whenever σ_A distinguishes between two sets of values $\mathbf{q}_1$ and $\mathbf{q}_2$ of the parents of D_A, i.e. $\sigma_A(\mathbf{q}_1) \neq \sigma_A(\mathbf{q}_2)$, there exist strategies $\boldsymbol{\sigma}_{-A} \in \boldsymbol{\Sigma}_{-A}$ for other players, and a possible action d_A for D_A, such that
$$\text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A),\mathbf{q}_1] \neq \text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A),\mathbf{q}_2]$$
where $\boldsymbol{\sigma} = (\sigma_A, \boldsymbol{\sigma}_{-A})$.
2. Whenever, for a given set of values $\mathbf{q}$ of the parents of D_A, σ_A distinguishes between two possible actions d_A^1 and d_A^2 for D_A, i.e. $\sigma_A(\mathbf{q})(d_A^1) \neq \sigma_A(\mathbf{q})(d_A^2)$, there exist strategies $\boldsymbol{\sigma}_{-A} \in \boldsymbol{\Sigma}_{-A}$ for other players, such that
$$\text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A^1),\mathbf{q}] \neq \text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A^2),\mathbf{q}] \quad \blacksquare$$

Trivially, any time a distinction is a not made the conditions are satisfied. We then define the following sets of strategies for all decisions:

$$\begin{aligned} \text{WD}_0 &= \text{all strategies} \\ \text{WD}_k &= \text{well-distinguishing strategies wrt } \text{WD}_{k-1} \\ \text{WD} &= \cap_{k=0}^{\infty} \text{WD}_k \end{aligned}$$

It is clear that WD is non-empty, since it trivially includes all strategies that do not distinguish amongst different cases of the parents. In fact, the following fact holds, reassuring us that WD is a useful solution concept.

Proposition 3: *The set of WD strategies always includes a Nash equilibrium.*

Proof: Let G be the given game. Let H be the *continuous* game in which the set Σ_A of *pure* strategies for decision A is the set of *mixed* strategies for A in G that are WD. The payoff functions u_A in H, given a pure strategy profile $\boldsymbol{\sigma}$, assign payoffs equal to the payoffs that would have been received for playing the mixed strategies $\boldsymbol{\sigma}$ in G. We can verify the following:

- S_A is nonempty. The mixed strategy that assigns uniform probability to all actions no matter what the value of the parents is WD.
- S_A is compact, because it is a closed and bounded subset of a Euclidean space.
- S_A is convex. Let σ_1 and σ_2 be two strategies for A in H, i.e. two WD mixed strategies for A in G. Let σ be a mixture of σ_1 and σ_2. Suppose σ differentiates between two values $\mathbf{q}_1$ and $\mathbf{q}_2$ of the parents of D_A. Then either σ_1 or σ_2 must also have differentiated between them. Since σ_1 and σ_2 are WD, there must exist strategies $\boldsymbol{\sigma}_{-A}$ for other decisions such that $\text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A),\mathbf{q}_1] \neq \text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A),\mathbf{q}_2]$. Therefore σ satisfies condition 1 of Definition 2. Now suppose σ differentiates between two possible actions d_A^1 and d_A^2. By a similar argument, either σ_1 or σ_2 must have so differentiated, and therefore σ satisfies condition 2 of Definition 2. Therefore σ is WD.
- u_A is continuous in $\boldsymbol{\Sigma}$, because it is linear.
- u_A is quasi-concave in Σ_A. This is true because Σ_A is convex and u_A is linear.

Thus H satisfies all the conditions of Theorem 1.2 of (Fudenberg & Tirole 1991), due to (Debreu 1952), (Glicksberg 1952) and (Fan 1952). Therefore H has a pure strategy Nash equilibrium $\boldsymbol{\sigma}^*$.

Now, let τ_A be a non-WDmixed strategy for A in G. We claim that there exists a WD strategy τ'_A that achieves the same payoff against $\boldsymbol{\sigma}^*_{-A}$ as τ_A.

Let $\mathbf{q}_1, ..., \mathbf{q}_n$ be values of the parents of D_A such that $\text{EU}^{\tau_A, \boldsymbol{\sigma}^*_{-A}}[A|\text{Do}(d_A), \mathbf{q}_i]$ is equal for all i. For all i, let

$$\tau'_A(\mathbf{q}_i) = \frac{\sum_{j=1}^{n} P^{\tau_A, \boldsymbol{\sigma}^*_{-A}}(\mathbf{q}_i)\tau_A(\mathbf{q}_i)}{\sum_{j=1}^{n} P^{\tau_A, \boldsymbol{\sigma}^*_{-A}}(\mathbf{q}_i)}$$

Note that the above probabilities and expected utilities do not actually depend on τ, and are equal if τ is replaced by τ'. Then the total time each strategy is played when any of the $\mathbf{q}_i$ happen is the same for τ'_A and τ_A, and since it does not make a difference which of the $\mathbf{q}_i$ hold, τ'_A and τ_A achieve the same total payoff over all cases where any of the $\mathbf{q}_i$ hold.

Now let $\mathbf{q}$ be a value for the parents of D_A, and let $d_1, ..., d_n$ be actions for D_A such that $\text{EU}^{\tau_A, \boldsymbol{\sigma}^*_{-A}}[A|\text{Do}(d_i), \mathbf{q}]$ is equal for all i. For all i, let $\tau'_A(\mathbf{q})(d_i) = \frac{1}{n}\sum_j \tau_A(d_j)$ Then, by an analagous argument, the total amount of time spent playing any of the d_i is the same for τ' as for τ, and since it does not make a difference which of them is played, τ and τ' achieve the same total payoff. By construction τ' is WD.

Now since σ^*_A is a best response to $\boldsymbol{\sigma}^*_{-A}$ in H, σ^*_A must be at least as good as τ'_A against $\boldsymbol{\sigma}^*_{-A}$. It follows that it must be at least as good as τ_A. Thus σ^*_A is at least as good as all mixed strategies in G, thus it is a best response to $\boldsymbol{\sigma}^*_{-A}$ in G. Since this is true for all decisions A, $\boldsymbol{\sigma}^*$ must be a Nash equilibrium of G.

∎

We say that a strategy profile is WD if all of its components are WD. A WD equilibrium is a NE that is also WD. Not every NE is a WD equilibrium, and not every WD strategy profile is a NE. There is a subtlety here. While the set WD always contains a NE, it is not true that there is always a NE that is WD with respect to WD *equilibria*, rather than with respect to all WD strategies. That is, there may be a situation in which the NE is WD if we consider all possible WD strategies for other players, but it is not WD if we only allow the other players to play WDNash equilibria. A variant of matching pennies in which the unique NE is for players to randomize non-uniformly between Heads and Tails is an example. In this NE, players are indifferent between the pure strategies in the support of the mixed strategy, when the other player is playing the NE strategy. So the NE strategy is not WD when only the NE strategy is considered for the other player. However, in this case, the pure strategies in the support of the mixed strategy are also WD, and the NE strategy is WD with respect to the pure strategies, so the NE strategy is WD.

We can now ask, under what circumstances do agents have a reason to prefer one strategy over another, when all agents are playing WD strategies. It turns out that these circumstances correspond exactly to the four cases described in the previous section. Each of these cases can be characterized by a certain pattern of paths in a MAID, as we capture with the following definition:

Notation and terminology: Let X and Y be two nodes. $\mathbf{V}_Y^X$ are the parents of Y that are descended from X, and $\mathbf{W}_Y^X$ are the other parents of Y. Let π be a path in a BN. A *V-node* is a node in π such that both edges in π go into the node. A $\wedge$*-node* is a node in π such that both edges in π come out of the node. The *key node* of π is the first $\wedge$-node in π. A *back-door path* is a path in which the first edge goes into the first node. A *front-door path* is a path in which the first edge comes out of the first node.

Definition 4: A decision node D_A is *effective* if one of the four following cases holds:

1. **Direct effect** There is a directed decision-free path from D_A to a utility node in $\mathbf{U}_A$.
2. **Manipulation**
   (a) There is a directed decision-free path from D_A to an effective decision node D_B.
   (b) There is a directed, effective path from D_B to $\mathbf{U}_A$ (an *effective path* is a path in which all decision nodes, except possibly the initial node, and except $\wedge$-nodes, are effective).
   (c) There is a directed, effective path from D_A to $\mathbf{U}_B$ that does not pass through D_B.
3. **Signaling**
   (a) There is a directed decision-free path from D_A to an effective decision node D_B.
   (b) There is a directed, effective path from D_B to $\mathbf{U}_A$.
   (c) There is an effective back-door path π from D_A to $\mathbf{U}_B$ that is not blocked by $D_B \cup \mathbf{W}_{D_B}^{D_A}$.
   (d) If C is the key node in π, there is an effective path from C to $\mathbf{U}_A$ that is not blocked by $D_A \cup \mathbf{W}_{D_A}^C$.
4. **Revealing/denying**
   (a) There is a directed decision-free path from D_A to an effective decision node D_B.
   (b) There is a directed, effective path from D_B to $\mathbf{U}_A$.
   (c) There is an effective indirect front-door path π from D_A to $\mathbf{U}_B$ that is not blocked by $D_B \cup \mathbf{W}_{D_B}^{D_A}$. ∎

The definition above is circular, because whether or not one node is effective depends on whether other nodes are effective. We therefore define the set of effective nodes to be the largest self-consistent set according to the above definition. A simple algorithm for finding this set is as follows.

$\mathbf{S} \leftarrow$ all nodes
Repeat until no node found
  Find a node $X \in \mathbf{S}$ that does not satisfy any of the four conditions, assuming nodes in $\mathbf{S}$ are effective
  Remove X from $\mathbf{S}$
Return $\mathbf{S}$

Have we described all the possible reasoning patterns that make an agent care about its decision? In other words, does the definition of effective node capture all situations in which a node is motivated? The answer to this question, if all players are playing WD strategies, is yes, as we now prove. We begin with a useful lemma. In the ensuing, we will use the idea of *severing edges* in a MAID between informational parents and decision nodes *and maintaining an equivalent network*. This means that for all strategy profiles $\boldsymbol{\sigma}$ in a class of strategies that we are considering, the probability distribution defined by $\boldsymbol{\sigma}$ can be implemented in a BN without the severed edges.

Lemma 5: *Let C be a node, D_A the decision node belonging to A and $\mathbf{U}_A$ be all utility nodes belonging to agent A. Then, if all players are playing WD strategies, and C is d-separated from $\mathbf{U}_A$ by $D_A \cup \mathbf{W}^C_{D_A}$, we can sever the links between $\mathbf{V}^C_{D_A}$ and D_A, and maintain an equivalent network.*

Proof: First we show that for any $U_A \in \mathbf{U}_A$, $\mathbf{V}^C_{D_A}$ is d-separated from U_A by $D_A \cup \mathbf{W}^C_{D_A}$. This is because any path π from some $V \in \mathbf{V}^C_{D_A}$ to U_A that is not blocked by $D_A \cup \mathbf{W}^C_{D_A}$ can be extended to the path π' beginning at C, descending directly to V, and continuing along π. The only additional nodes in π' are non-V-nodes that are not observed and do not block π', and V itself. If π is a front-door path from V, π' is not blocked by V, since V is not observed. If π is a back-door path from V, π' is not blocked by V, since V is a V-node in π', and the child D_A of V is observed. Therefore π' will not be blocked by $D_A \cup \mathbf{W}^C_{D_A}$. Thus, since we have assumed that all paths from C to U_A are blocked by $D_A \cup \mathbf{W}^C_{D_A}$, so must all paths from $\mathbf{V}^C_{D_A}$ to $\mathbf{U}_A$. Therefore $\mathbf{V}^C_{D_A}$ is d-separated from $\mathbf{U}_A$ by $D_A \cup \mathbf{W}^C_{D_A}$. It follows that for any strategies $\boldsymbol{\sigma}_{-A}$ for other decisions, and for all d_A, $\text{EU}^{\boldsymbol{\sigma}_{-A}}[A|\text{Do}(d_A), \mathbf{V}^C_{D_A}, \mathbf{W}^C_{D_A}] = \text{EU}^{\boldsymbol{\sigma}_{-A}}[A|\text{Do}(d_A), \mathbf{W}^C_{D_A}]$. Therefore any WD strategy σ for D_A does not distinguish between different values of $\mathbf{V}^C_{D_A}$, and we can sever the links between $\mathbf{V}^C_{D_A}$ and D_A and maintain an equivalent network. ∎

Theorem 6: *If D_A is motivated with respect to the set of WD strategies, and in particular, with respect to WD equilibria, then D_A is effective.*

Proof: We will show that if D_A is not effective, it is not motivated with respect to WD strategies. The proof is by induction on the order in which nodes are marked as non-effective. The base case where there are zero non-effective nodes is trivial. For the inductive step, note first that the induction hypothesis implies that for any node previously marked as non-effective, any strategy that makes distinctions among the parents is not WD. Therefore we can sever the links between such nodes and their parents, and maintain an equivalent network.

Now, let D_A be a node newly marked as non-effective. Assume, by way of contradiction, that D_A is motivated, i.e. there exist $\mathbf{q}$, $\boldsymbol{\sigma}_{-A}$, σ_{A_1} and σ_{A_2} such that $\text{EU}^{(\sigma_{A_1}, \boldsymbol{\sigma}_{-A})}[A|\mathbf{q}] \neq \text{EU}^{(\sigma_{A_2}, \boldsymbol{\sigma}_{-A})}[A|\mathbf{q}]$. Since D_A is non-effective, there can be no directed decision-free path from D_A to any U_A. But there must be some directed path from D_A to some $U_A \in \mathbf{U}_A$, or else the expected utility of A could not depend on her strategy in any way. Therefore there must be a directed decision-free path from D_A to some D_B, and a directed path from D_B to U_A. Furthermore, D_B and the path from D_B to U_A must be effective, because otherwise the path would include severed links. Therefore conditions (a) and (b) of cases 2–4 of Definition 4 are satisfied.

Since we have shown that there must be a directed path from D_A to D_B, it cannot be the case that we can sever the links between $\mathbf{V}^{D_A}_{D_B}$ and D_B. By Lemma 5, had D_A been d-separated from $\mathbf{U}_B$ by $D_B \cup \mathbf{W}^{D_A}_{D_B}$, we could have severed the links between $\mathbf{V}^{D_A}_{D_B}$ and D_B. It follows that there must be an effective path from D_A to some $U_B \in \mathbf{U}_B$ that is not blocked by $D_B \cup \mathbf{W}^{D_A}_{D_B}$. The only way in which this can be the case without satisfying the conditions of cases 2–4 of Definition 4 is if the path is a back-door path π from D_A, and there is no effective path from the key node C in π to $\mathbf{U}_A$ that is not blocked by $D_A \cup \mathbf{W}^C_{D_A}$. But then, by Lemma 5, we can sever the links between $\mathbf{V}^C_{D_A}$ and D_A.

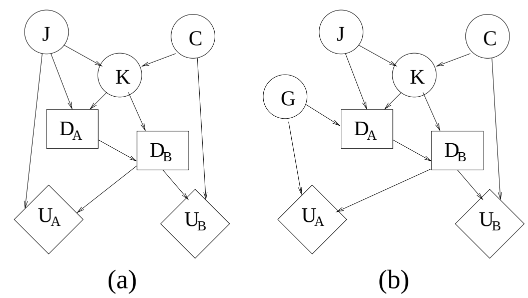

Figure 2: Examples for proof of Theorem 6

We claim that as a result, there can be no path π from D_A to $\mathbf{U}_B$ that is not blocked by $D_B \cup \mathbf{W}^{D_A}_{D_B}$ that does not contain severed links. Since cases 2 and 4 do not hold, there is no such directed or front-door path. Here we must be careful. The most natural thing to say is that any such path π must begin by going from D_A to a parent $V \in V^C_{D_A}$ and this link is severed, so π is not active. However, this need not be the case. The type of situation we must take into account is shown in Figure 2(a). Here the path π from D_A to U_B that goes via J, K and C is not blocked by $D_B \cup \mathbf{W}^{D_A}_{D_B}$. However, in this case, observe that we can make J a key node. There is an active path from J to U_A, and we have supposed that the path from A to U_B that goes via J is active. Hence the conditions of case 3 are satisfied, contrary to assumption. This argument can be generalized. The path from D_A to U_B via J must begin with a parent $W_1 \notin \mathbf{V}^C_{D_A}$. There are two cases to consider. In the first case, there is a path from J to U_A that is not blocked by $D_A \cup \mathbf{W}^J_{D_A}$, and thus the conditions of case 3 are satisfied, contrary to assumption. In the second case, as in Figure 2(b), J is d-separated from U_A by $D_A \cup \mathbf{W}^J_{D_A}$. In that case, by Lemma 5, we can sever the links between $\mathbf{V}^J_{D_A}$ and D_A. Since $W_1 \in \mathbf{V}^J_{D_A}$ and $W_1 \notin \mathbf{V}^C_{D_A}$, we are severing at least one new link. After this has been done, we can continue with the same argument. Either there is still an active path from D_A to U_B that does not contain severed links, or else we are done with the claim. If there is still an active path, either case 3 is satisfied contrary to assumption, or we can sever more links. Eventually we will run out of links to sever, so the claim is proved.

Thus there is no path from D_A to U_B that is not blocked by $D_B \cup \mathbf{W}_{D_B}^{D_A}$. Appealing once again to Lemma 5, we can sever the links between $\mathbf{V}_{D_B}^{D_A}$ and D_B. But then there can be no directed path from D_A to D_B that does not contain severed links. We have already shown that if D_A is non-effective and motivated, such a path must exist. Therefore we conclude that D_A, the node currently being marked as non-effective, is non-motivated. ∎

This theorem has computational implications for solving MAIDs. We can begin by identifying the non-effective nodes, one by one. Each such node is then replaced by a chance node that randomizes uniformly over its decisions, and the links from its parents to the node are severed. We can then also sever additional edges as identified by Lemma 5. This may lead to more nodes becoming non-effective, so the process iterates. Once these modifications have been made, the simplified MAID can be passed to a MAID solution algorithm such as (Koller & Milch 2001; Vickrey & Koller 2002). This has benefits both because the resulting MAID is sparser, and because there are fewer decisions to make.

Now that we have shown that the four patterns of Definition 4 completely characterize all the situations in which a node is motivated, the next question to ask is whether the converse is true. If a node is effective, does that necessarily mean that it is motivated? Unfortunately, as stated, the answer is no. We can design networks in which a node is effective but the action at that node makes no difference as to the outcome. However, we can prove a weak converse, that says that if a node is effective, then there is some set of parameter values for the network such that the node is motivated. The situation is similar to Bayesian networks, where d-separation between two nodes implies that they are independent, but lack of d-separation does not imply that they are dependent. However, if two nodes are not d-separated they are dependent for some parameter values.

Theorem 7: *If a node D_A in a MAID is effective, then there exist parameter values for the MAID such that D_A is motivated with respect to WD strategies.*

Proof: The proof is by construction. First we note that in any network, we can select a subset of the nodes in the network as important, and render the rest of the nodes irrelevant. For a chance node, we can render it irrelevant by setting its CPD to be uniform whatever the value of its parents. For a utility node, we set the utility to be the same for all values of its parents. For a decision node, we set the utility nodes of the same agents to assign the same utility in all situations. Because the decision is WD, the strategy for the node must assign equal probability to all actions for all values of the parents. We can therefore sever each such node from its parents. Also, we can make its chance-node children ignore the node in their CPDs, thus allowing us to sever the link from the node to these children. Similarly for utility-node children. As for decision-node children, we can force them only to consider strategies that do not depend on this node. If we succeed in showing that there exist parameters such that a node is motivated under this restriction, then it is certainly the case without the restriction. Therefore we can sever all the links between the node and its parents and children, and safely remove it from the network. Thus we can focus our attention to the patterns appearing in the definition, without having to worry about extraneous nodes. Furthermore, if there are multiple redundant patterns, we eliminate all except one. It suffices to show that if there is one such pattern, we can assign paramters to it to make the node motivated.

Now we assign parameters to the network as follows. We select a distinguished value x^1 for each node X. For each chance node X, we assign its CPD as follows. Let $\mathbf{q}$ be an assignment of values to the parents of X. If the number of parents that take on their distinguished value is even, the CPD assigns probability $\frac{p-1}{p}$ to x^1, where p is a prime, with the remaining probability distributed uniformly to the remaining values of X. If the number of times is even, the CPD distributes probability $\frac{p-1}{p}$, where p is a different prime, uniformly among the non-distinguished values of X, with the remaining probability assigned to x^1. A different prime is used for all different parameters of the network. For a utility node U, a utility of 1 is assigned if the number of parents that take on their distinguished value is even, 0 otherwise.

Let $\boldsymbol{\sigma}$ be the strategy profile in which each agent with an effective node plays a near-parity strategy like the ones just defined for chance nodes, and each agent with a non-effective node randomizes uniformly for all values of its parents. We claim that in this network, $\boldsymbol{\sigma}$ is WD. This is proved by induction. Obviously it is in WD_0 since that contains all strategies. Assume it is in WD_i. We can prove the second condition of Definition 2 by showing that for all effective nodes D_A, and for all values $\mathbf{q}$ of their parents,

$$\text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A^1), \mathbf{q}] \neq \text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d'_A), \mathbf{q}]$$

where d'_A is a non-distinguished value of D_A. Since D_A is effective, there must be a directed path from D_A to U_A in which every decision node is effective. At each node X along the path, the probability of x^1 will be the sum of products of rational numbers, in which the primes in the denominator will be different for d_A^1 and d'_A. Therefore the distribution over U_A will also be different, as required.

For the first condition of Definition 2, consider an effective node D_A, and a parent V. By our construction, D_A distinguishes between different values of V. We must show that

$$\text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A^1), V, \mathbf{W}] \neq \text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d'_A), V, \mathbf{W}]$$

where $\mathbf{W}$ are the other parents of D_A. Since we have removed all extraneous nodes, there are the following cases to consider:

1. V is part of a path to an ancestor decision D_B. Now there are three subcases.
   (a) There is a directed path from D_B to $\mathbf{U}_A$ that does not pass through D_A. In this case, since we have removed redundant paths from the network, the path cannot also pass through a different parent $W \in \mathbf{W}$.

(b) There is a back-door path from D_B to $\mathbf{U}_A$ that is not blocked by $D_A \cup \mathbf{W}$.

(c) There is a front-door path from D_B to $\mathbf{U}_A$ that is not blocked by $D_A \cup \mathbf{W}$.

In all three sub-cases, the path from V via D_B to $\mathbf{U}_A$ is not blocked by $D_A \cup \mathbf{W}$, and the statement holds.

2. V is part of a back-door path from D_A to a utility node U_B of some other decision. In this case, by condition 3(d) of Definition 4, there must be a path from the key node C to U_A that is not blocked by $D_A \cup \mathbf{W}$, so again the statement holds.

Therefore $\sigma_A \in \text{WD}_{i+1}$ for all decisions D_A, and therefore $\boldsymbol{\sigma} \in \text{WD}_{i+1}$. Thus $\boldsymbol{\sigma} \in \text{WD}$. Since for an effective strategy D_A, and values $\mathbf{q}$ of its parents, $\text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d_A^1), \mathbf{q}] \neq \text{EU}^{\boldsymbol{\sigma}}[A|\text{Do}(d'_A), \mathbf{q}]$, D_A is motivated with respect to WD strategies. ∎

Conclusion

This paper has made a number of contributions. The most important is to open up the question of analyzing and identifying the reasoning patterns agents use in games, on a more fine-grained level than simply maximizing expected utility. We have identified a number of interesting patterns, some of which have surprising properties. We have shown that a subset of these patterns strongly characterizes situations where agents are motivated to take an action given that other agents are playing strategies that only make distinctions where those distinctions make a difference.

The reasoning patterns captured by the WD solution concept are highly-justified ones for which there is a very good explanation, but it would certainly be interesting to find a solution concept that includes a wider range of reasoning patterns, while not going as far as general strategies to include all of them. It would also be interesting to analyze established solution concepts such as rationalizability (Bernheim 1984; Pearce 1984), and to examine other forms of graphical games (Kearns, Littman, & Singh 2001).

Our results have implications for all the purposes discussed in the introduction. For explanation, the reasoning patterns described here provide the basis for intuitive explanations that make sense. Furthermore, explanations that use WD strategies will be more convincing than those that do not. For knowledge engineering, an engineer can ask questions like "Do I expect signaling to be present in this scenario? If so, I must make sure the paths supporting the signalling patterns are present in my model." Computationally, Lemma 5 and Theorem 6 provide the basis for simplifying models before passing them to a MAID solution algorithm. For belief update, understanding the circumstances in which one agent affects what another knows helps understand when and how an agent should update its beliefs. In particular, if signalling and revealing/denying are not present, beliefs should not be updated.

Finally, understanding these reasoning patterns can form the foundation for understanding how people reason in games that can be described by MAIDs. Here, there are many interesting open questions to explore. Are real agents more likely to use some patterns of reasoning than others? For example, one might conjecture that "revealing/denying" is less likely to be used, at least until it has been explicitly pointed out. How can we study these questions in the lab or the field? This is an interesting direction for future work.

There are other implications of this research to explore. For example, are real agents (human and computer) more likely to use some patterns of reasoning than others? For example, one might conjecture that "revealing/denying" is less likely to be used, at least until it has been explicitly pointed out. Another question regards belief update. Suppose B observes A's action, and A has access to C which is important to B. How should B update his beliefs about C? Answering this question requires understanding what motivations A might have to report the value of C.

Acknowledgements

This work was sponsored by AFOSR under contract FA9550-05-1-0321. Thanks to David Parkes and Sevan Ficici for useful comments and suggestions.

References

Bernheim, B. D. 1984. Rationalizable strategic behavior. *Econometrica* 52:1007–28.

Debreu, D. 1952. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences* 38:886–893.

Fan, K. 1952. Fixed point and minimax theorems in locally convex topological linear spaces. *Proceedings of the National Academy of Sciences* 38:121–126.

Fudenberg, D., and Tirole, J. 1991. *Game Theory*. MIT Press.

Glicksberg, I. L. 1952. A further generalization of the Kakutani fixed point theorem with application to Nash equilibrium points. *Proceedings of the National Academy of Sciences* 38:170–174.

Howard, R., and Matheson, J. 1984. Influence diagrams. In *The Principles and Applications of Decision Analysis*. Strategic Decisions Group, Menlo Park, California. 721–762.

Kearns, M.; Littman, M.; and Singh, S. 2001. Graphical models for game theory. In *Uncertainty in Artificial Intelligence*.

Koller, D., and Milch, B. 2001. Multi-agent influence diagrams for representing and solving games. In *International Joint Conference on Artificial Intelligence*.

Pearce, D. G. 1984. Rationalizable strategic behavior and the problem of perfection. *Econometrica* 52:1029–50.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann.

Pearl, J. 2000. *Causality: Models, Reasoning and Inference*. Cambridge University Press.

Vickrey, D., and Koller, D. 2002. Multi-agent algorithms for solving graphical games. In *National Conference on Artificial Intelligence (AAAI)*.