RESOLVING CONFLICTS IN A TREE:
DRAMA THEORY IN THE EXTENSIVE FORM

Nigel Howard (dramatec Ltd)

ABSTRACT

Game theory assumes that it must be possible for a conflict represented by a game to be satisfactorily resolved without re-defining the game itself. Drama theory asserts that full conflict resolution generally requires players to engage in a rational-emotional process of re-defining both the game and their "positions" in it until there exists a fully satisfactory resolution on which they all agree. In re-defining the game and their positions in it, players must eliminate six “dilemmas”, each of which therefore tends to cause emotions and rationalisations tending toward its elimination. For example, a party needing to make a threat credible tends to produce negative emotions and rationalisations that increase its preference for carrying out the threat. This process of dilemma-elimination is assumed to take place as players communicate with each other prior to playing the game. It cannot be called “rational” in the sense of “instrumental rationality”, since that takes preferences and beliefs as fixed. Dilemma-elimination requires changing preferences and beliefs, and can therefore only be called “rational” in a wider sense.

So far, dilemma-elimination has been analysed assuming that at each stage, the conflict exists within a "strategic game"— ie, a game in which all final choices are simultaneous and independent. This raises the question: how is the process affected if it takes place within an "extensive game"—one in which some choices are made after others have become fixed? In such cases, conflict resolution becomes a contingent, time-dependent process. This paper answers this question in a manner that allows real-world players to react to the game-theoretic problems raised by the extensive form while avoiding some of the difficulties encountered by the concept of “backward induction.”

DRAMA THEORY AND THE CHALLENGE TO INSTRUMENTAL RATIONALITY

Drama theory challenges the game-theoretic approach to resolving conflict. This is basically the approach used by economists, decision theorists and many others. It’s an essential part of our way of understanding the world.

Game theorists try to resolve conflicts *without re-defining the game*. And what does this mean? “The game” is defined by players’ preferences and their beliefs about their opportunities. So game theory assumes conflicts can be resolved without changing these.

Actually, games are about the only kind of human interaction in which these *are* fixed. When you're playing a game, your preferences and opportunities are fixed by the rules of the game. In most interactions, they're not. You spend time debating what they should be.

Drama theory proves that, in general, a satisfactory resolution can't be found if preferences and beliefs (about others' preferences and about opportunities) are fixed. Players' "hearts and minds" have to change. Otherwise it's generally impossible for the following criteria of "satisfactoriness" to be met:

- no disagreement—ie, all parties propose the same solution;
- no distrust—ie, the common solution is a *strict, strong equilibrium*[1].

How is this proved? Within what framework?

The drama-theoretic framework assumes the following. First, each party openly states a *position*; this is the solution it advocates. Then, if positions differ, it states a *fallback strategy*; this is what it will unilaterally do if positions don't change. These statements of positions and fallback take place before decisions are made. They occur in a period of *pre-play communication*.

This is the framework assumed by drama theory. Within this framework, it's proved that "no disagreement" and "no distrust" require the non-existence of six independent *dilemmas*. Drama theory predicts, therefore, that each party will use emotion and argument to try to change hearts and minds in a way that eliminates the six dilemmas. It will naturally try to eliminate them *in its favour*—ie, so that a common solution is implemented in line with its objectives.

Despite this element of objectives-seeking, dilemma-elimination isn't and can't be "rational". Not, at least, in the sense of "instrumental rationality". This interpretation of rationality assumes preferences and opportunities to be fixed. Given this, the only problem a player faces is the technical one of optimizing over a given set of possibilities.

This may be analytically convenient. But if "rationality" requires fixed preferences and opportunities, then *re-defining* one's preferences and opportunities can't be rational.

It can, however, be rational in a wider sense. In this sense, some ways of arriving at beliefs and preferences are more rational than others. Rationality in this wider sense involves, we'll see, emotional appeals and argumentation between parties. Such activities have no role in game theory.

The paradigm of instrumental rationality is deeply embedded. It's seen in the ever-hardening distinction between consumer and producer. Consumer "preferences" are arbitrary. They require and permit no justification. Producer "strategy choices" are ever more rigorously determined—by the requirement to maximize exchange value for consumers.

Instrumental rationalists see this as right. They also argue that emotion and debate should play no part in arriving at justified beliefs. These should emerge from a disinterested survey of facts.

Impossible. What facts? Students of scientific method, from Popper [1] to Kuhn[2], stress that "facts" are too many to be disinterestedly collected. We need a reason to select the facts we collect.

Arguments too need a purpose more specific than establishing "the truth". To evaluate an argument we must know what it aims to prove. This principle is even found in statistical methodology (in the distinction between Type I and Type II error).

Thus, scientific method finds that rational methods of belief formation involve collecting facts and making arguments *in order to prove or disprove something*. This is also the view of drama theory.[2] And what are rational methods of preference formation? Partly they are a matter of forming beliefs about *consequences*. Thus they overlap with methods of belief formation.

In both cases, emotion is important. Without it, the essential element of motivation is lacking. There has to be a reason for wanting to prove or disprove something.

A MIDDLE EAST EXAMPLE

To illustrate the drama-theoretic process of dilemma-elimination, we'll use a model of the Arab-Israeli conflict. The model remains current, though built over a year ago.

Consider the "card table" in Figure 1. Each column shows a selection of "cards" played. Look first at column p. "p" stands for "present intentions"; this column serves to calibrate the model. It states that the future *if present intentions continue to be carried out* will be that the Palestinians won't stop terrorism or recognize/accept Israel; the Arab states will continue to fund terrorism and not recognize or accept Israel; and Israel won't accept a "viable" Palestinian state (as defined by Palestinians)[3] but will raid and suppress Palestinians.

Now consider the other columns. These show parties' *positions* and, in between the positions, a column f. This "f" column is the *fallback*–the future that will eventuate if each party implements its *fallback strategy*.

Look at positions first. The Palestinian/Arab position (P+A) is that they will stop terrorism and recognize/accept Israel if Israel will accept a "viable" Palestinian state and cease raids and suppression. The Israeli position is the same, except for acceptance of a "viable" Palestinian state. (Israel rejects this as it is defined by the Palestinians; or, to put it another way, Israeli's offer of a state is seen by the Palestinians and Arabs as derisory). The fallback is actually the same as column p; thus, what each side says it will do if positions don't change is continue the present future, with terrorism on one side, suppression on the other.[4,5]

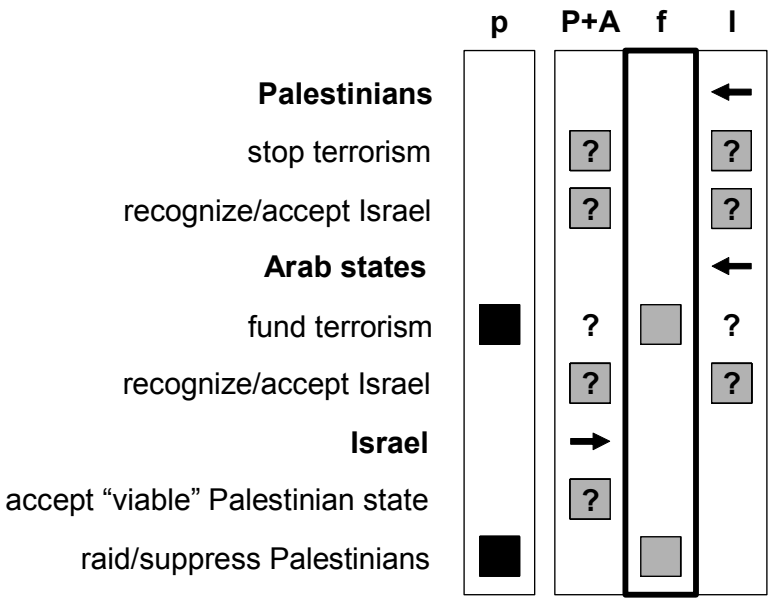

Figure 1: Card table model of Arab-Israeli conflict

What do the arrows and question marks mean? The arrows stand for *perceived preferences* of the party in whose row they're drawn. *These preferences are always between positions and the fallback.* Here, they show that each side is perceived to prefer the fallback to the other's position. A question mark in a column stands for a *doubt* on the part of relevant parties that the intention stated in that part of the column would actually be carried out. Thus, they show (a) Israeli doubts that Palestinians and Arabs would stick to any agreement to stop terrorism and recognize/accept Israel, (b) Palestinian/Arab doubts that Israel would carry out any agreement to accept a "viable" Palestinian state.

Note that doubts are assumed to affect preferences. For example, Israeli doubts of Palestinian/Arab contingent intentions are a large part of their reason for preferring the fallback to acceptance of a "viable" Palestinian state. Note too that doubts may appear in the fallback as well as in parties' positions; they do so when a party's threat isn't credible. In this case, however, each party's threats are credible; they're carried out on a daily basis. There are therefore no doubts in the fallback.

Next, consider the "tug of war" diagram in Figure 2. This presents the data in Figure 1 in a different way, so as to highlight the dilemmas faced by parties. Ovals are positions; the rectangle in the middle is the fallback. Horizontal arrows represent the preference arrows appearing in Figure 1; vertical arrows represent the doubts (question marks).

How do you win a tug of war? In terms of our diagram, how do you get others to accept and implement your position? Quite simply, you're certain you will win a tug of war if both (or all) horizontal arrows point toward you and no vertical arrows go from your position.

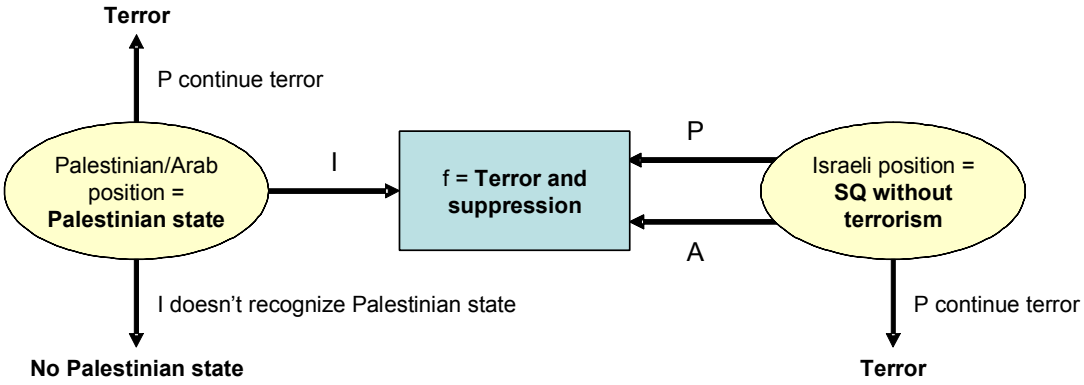

Figure 2: Arab-Israeli tug of war

How can we say this? Because:

- The fact that no vertical arrows leave your position means it's credible – ie, if accepted by all parties it would be implemented.[6]
- Preferring the fallback to any other party's position, you're under no pressure to accept it.
- Preferring your position to the fallback, all other parties are under unilateral pressure to accept it.

Note that pressure is one thing; it can be resisted by counterbalancing pressure, just as one side's pull in a tug of war is balanced by another's. Unilateral pressure is something else. It's pressure against which there's no resistance.

Now, each party knows this (consciously or unconsciously). Drama theory therefore predicts that each will try to make the relevant arrows point its way. How? Various means are open to it. It can add or delete cards (by arguing for their relevance or irrelevance). It can argue for hitherto unseen consequences of card selections. It can express emotions (eg, of benevolence, defiance or hatred) that indicate its preferences are changing. It can take concrete actions that change perceived preferences, consequences and opportunities. Such measures try to change or maintain the direction of arrows both for direct reasons (I want to make horizontal arrows point my way and no vertical arrows leave my position) and indirect ones (creating or abolishing vertical arrows may, we've seen, affect the direction of horizontal arrows).

How do parties finally achieve a common position, if they do? Well, one way to make your horizontal arrow go my way is to alter my position to make it more acceptable to you; in the extreme, I may abandon my position for yours. In general, trying to make their positions more attractive to each other tends to move parties toward a common position.

Whether parties converge to a common position or escalate toward conflict, it's this process of change that drama theory looks at. Drama theory, in other words, doesn't make assumptions about parties, preferences and options in order to draw final conclusions from them. It makes assumptions about them in order to see how these assumptions are under pressure to change. A drama-theoretic confrontation is

dynamic. It evokes emotions that drive parties to try to *change* assumptions, however initially correct they may be, in their favour.

The point of drama theory for a party involved in a confrontation is to see where it needs to take measures to make arrows point its way and thwart others' efforts to make arrows point their way. Its point for a party analyzing someone else's confrontation is to understand the dynamic, emotional forces that are at work causing beliefs, preferences and attitudes to change, and propelling parties toward cooperation or conflict. This can also help an involved party. It's useful to understand your own and others' feelings and motivations.

Dilemma analysis

How does this relate to the "dilemmas" mentioned above and discussed in Howard (1999) or Bryant (2003)? Quite simply, each arrow in a tug of war represents a specific dilemma facing a party. Thus,

- I face a *threat* dilemma if a vertical arrow from the fallback has my name. *My threat isn't credible.*
- I face a *rejection* dilemma if a horizontal arrow to another's position has my name. *I'm under pressure to accept their position.*
- I face a *persuasion dilemma* if a horizontal arrow from my position has another's name. *They're under no pressure to accept my position.*
- I face a *trust* dilemma if a vertical arrow from my position doesn't have my name. *I can't trust others to implement my position, even if they agree to it.*
- I face a *cooperation* dilemma if a vertical arrow from my position has my name. *Others can't trust me to implement my position, even if they agree to it.*[7]

Thus, in looking at measures for changing arrows we're considering how parties may seek to eliminate dilemmas. We're also thinking how other parties, for whom the dilemmas operate in their favour, may resist those attempts.

Applying this to our example, consider the following measures:

<u>Palestinians and Arabs:</u> Position: a viable Palestinian state not dependent on Israel.

- To overcome potential Rejection Dilemmas (ie, make right horizontal arrows continue to point left): Show/feel preference for Terror/suppression over Status quo with no viable Palestinian state.
- To overcome a Persuasion Dilemma (ie, make left horizontal arrow point left): Determine to make terror worse for Israel, so as to make Israel prefer a viable Palestinian state to Terror/suppression.
- To overcome a Cooperation Dilemma (ie, remove arrow going up from Palestinian State): Argue that a US-led intervention force will convince Israel that terror will cease and Israel will be accepted/recognized once there is a Palestinian state.
- To overcome a Trust Dilemma (ie, remove arrow going down from Palestinian state). Argue that a US-led intervention force will compel Israel to abandon settlements.

<u>Israel:</u> Position: cessation of terror without a "viable" Palestinian state.

- To overcome a potential Rejection Dilemma (ie, make left horizontal arrow continue to point right): Show/feel distrust of Palestinians, hence preference for Terror/suppression over Palestinian state.
- To overcome a Persuasion Dilemma (ie, make right horizontal arrow point right): Determine to make suppression worse for Palestinians and the status quo better, so as to make them prefer Status Quo without a Palestinian state to Terror/suppression.
- To overcome a Trust Dilemma (ie, remove the arrow going down from Status Quo without terrorism): Have policy of fierce retaliation against terrorism, amelioration of conditions when terrorism lessens and separation of Palestinian and Israeli populations.

This list of measures could, of course, be extended. Dilemma analysis uncovers the reasons for parties' attitudes, emotions and actions. An important distinction is between measures already taken and future measures; drama theory is helpful in forecasting the latter, as it focuses on *why* parties will look for new measures.

Note that because drama theory requires you to think how parties will use emotion and creativity to change the assumptions initially put into the model, it requires you to "think outside the box" – ie, to think how parties will react to your analysis, thus changing it. Dilemma analysis is a tool for showing where the model itself is likely to be "wrong" in that, while presumably correct at present, it's likely to change as players react to it. However, information as to exactly how it will change (as distinct from the direction of and motivation for change) cannot, by definition, be present in the model itself. It must be sought in the real-world context of the model—what it represents. This, therefore, is where the user must look for answers to the questions the model puts.

DILEMMAS IN A TREE

Drama theory, so far, has implicitly assumed that parties will eventually make their decisions simultaneously and independently, as in a so-called "strategic" game.

These simultaneous decisions are assumed to follow a period of pre-play communication, in which parties use emotion and argument to try to re-define the game. While this is going on, no final decisions are made.

When this period of communication ends, where do parties stand? They may or may not have agreed on a common position. If they have, their stated intention is to carry it out. Each must simultaneously and independently decide whether to do so. If they haven't, their stated intention is their fallback strategy. They must decide whether to implement that.

Their decisions are, at this point, game-theoretic. In deciding what to do, they'll have to decide what they think others will do. The question they must now answer is the game-theoretic one: is there a convergence of beliefs about intentions? Stated intentions may be the starting point for this analysis. They may focus expectations on one of a number of possible equilibria, thus solving a problem that arises in game theory when there are multiple equilibria. But their main function has been to assist

parties in their efforts to re-define the game. For this, they are no longer relevant. However parties may or may not have succeeded in redefining the game, it is what it is now. It will be played as such.

According to game theory, this means that if stated intentions are a strict, strong equilibrium, they will be carried out. If merely an equilibrium (so that no individual can gain by an individual defection, though an individual or coalition may have a defection under which no member loses), they *may* be carried out. If they aren't even an equilibrium (so that some individual can gain from an individual defection), they won't be carried out.

Drama theory does not dispute this analysis. But its predictions from it are not the game-theoretic ones. It's interested in the period of pre-play communication, when "stated intentions" are stated. To make these intentions credible, players may try to overcome problems in implementing them. Needing to make others' intentions incredible may lead them to point out and emphasise implementation problems.

In general, the presence or absence of game-theoretic "defections" will be arguments drama-theoretic players can use.

How? Players' conditional beliefs about defections are indicated in a card table by doubts (question marks). A doubt raised in a column indicates a belief that a party would defect from that position or fallback. If there are no doubts in a column, it's believed that there will be no defections. Thus, doubts refer to perceptions, taken as data. We *observe* doubts. We do so by taking note of what parties say. A doubt exists if a doubt is expressed.

Now, a game-theoretic analysis finds formal, logical reasons as to why doubts *should* exist. The drama-theoretic question is: are these formal, logical reasons for doubts judged to be causing actual doubts? If so, then they point to characteristics of the current environment that parties will try to change. If not, then it's realistic to ignore them.

In other words, there may be formal reasons for a doubt—but no doubts. This is realistic. Paradoxes of "backward induction" found in game theory show how it's possible for formal doubts to be of a kind that real players ignore.

It's also possible to have doubts without any formal justification. In that case, if no formal reasons exist for a doubt, then pointing this out is an argument that players can use to try to eliminate the doubt. Paradoxes of backward induction show that there is no guarantee that they will succeed: absence of formal reasons for a doubt may not be sufficient to eliminate it. Nevertheless, players can try.

In sum, there are two ways in which game-theoretic analysis serves the drama-theoretic purpose of predicting how parties will try to change the current environment. First, it points to characteristics of the environment that ought, logically, to give rise to doubts. If these are giving rise to actual doubts, they're characteristics that some parties will try to change; if not, they're characteristics that other parties may use as arguments that doubts ought to exist. Second, it points to characteristics of the environment that mean there is no logical reason for a doubt. These provide

arguments that are doubt-eliminating, hence likely to be used by parties with that aim; at the same time they're characteristics that other parties, wishing to cause doubts, might try to change. In each case, the game-theoretic analysis is used to track back from the abstract model to the concrete characteristics of the environment that provide arguments players will want to use.

In Figure 1, for example, column P+A is, in the current environment, beset by doubts. Palestinians and Arab states doubt that, if they gave up terrorism, the Israelis would withdraw from their settlements comprehensively enough to allow a viable Palestinian state; Israelis doubt that the Arabs, given a viable Palestinian state, would give up terrorism. Thus, both sides perceive that P+A would not be implemented. This—assuming that doubts result from formal, game-theoretic reasoning—would be only the point made by game theory. The additional drama-theoretic point would be that the Palestinians and Arabs (whose position this is) will try to change the environment to eliminate these doubts. As said, this is how we explain their efforts to get American involvement, which they see as giving assurance to each sides that the defections they fear would not occur. It also explains their frequent offers, in exchange for a viable, independent state, to recognize Israel and forgive what they see as Israeli crimes. On the other side, it explains why the Israelis (whose position differs from P+A) oppose American involvement and reject Arab offers of recognition and forgiveness as deceptions—arguing that they contradict their continuing support for terrorism.

Dilemmas of confrontation and implementation

All this shows how game-theoretic analysis is relevant to the vertical arrows in a tug of war—the arrows that represent doubts as to whether positions or the fallback would be implemented. These doubts correspond to three drama-theoretic dilemmas—the trust, cooperation and threat dilemmas. They are the *dilemmas of implementation*.

Game theory is, however, irrelevant to the dilemmas represented by horizontal arrows—the persuasion, rejection, and positioning dilemmas. These may be called *dilemmas of confrontation.* They show pressures on parties to abandon their positions, not because they are unimplementable, but because of their benefits or disbenefits compared to other positions or the fallback.

Both types of dilemma are so called because of the hard choices they present during pre-play communication. But the hard choices are different. The hard choice in a dilemma of confrontation is whether to keep pressing for a position that seems untenable, given the costs and benefits, for each party, of each position and of the fallback. The hard choice in a dilemma of implementation is whether to keep making or requesting a promise or threat that seems incredible.

Eliminating dilemmas of implementation

Now let's ask: given that dilemmas of implementation are game-theoretic, what happens if we lift the assumption of simultaneous strategies? This would mean replacing the strategic game with one in extensive form; that is, it would mean allowing some decisions to be made, and become irreversible, before others. What difference would this make to methods of eliminating threat, cooperation and trust dilemmas?

The first point is that a good analysis can be made without lifting this assumption. A model is only a model, not reality. This fact is particularly significant in drama theory, where we analyse the model in order to track back to reality and see what real-world characteristics the parties may try to change. But since a model is only a model, the boundary between model and reality is variable. Different models locate it differently. What the formal assumption of simultaneous implementation really means is that considerations about the sequencing of decisions are located in the environment—the reality being modelled—not the model.

But we've said that considerations not in the model are just what drama theory leads us to focus on. It allows a player to give such considerations as real-world reasons for eliminating or retaining a doubt, even though they can't be looked at formally, within the model.

All the same, it may be advantageous to model non-simultaneous decisions formally—ie, bring them into the model. Consider, for example, some arguments Israel might use for retaining the doubts in the P+A column of Figure 1. How might it justify its own doubts as to whether the Arab side would give up terrorism? It could point out that it's being asked to trust the Arabs after, not merely before, abandoning its settlements and allowing the creation of a viable Palestinian state capable of independent action. Wouldn't the Arabs, having got some of what they wanted by temporarily giving up terrorism, revert to terrorism in the hope of getting more? Next, Israel could further undermine the Arab position by arguing for the validity of doubts as to whether it itself would abandon its settlements in order to enable the creation of a viable Palestinian state. It wouldn't want to argue this directly. A party won't want to undermine its general credibility by arguing that it couldn't be trusted carry out an agreement. Instead, Israel might argue that it would be politically impossible for it to abandon its settlers in exchange for Arab promises that are obviously untrustworthy. That would be another way of saying the same thing. In answer to the Arab argument that the US can and should pressure Israel into such action, Israel might further argue that, ultimately, it trusts the US not to abandon Israel simply because Israel finds it impossible to trust the Arabs.

The first argument cites the fact that if P+A were implemented, a time would come when Israel's decisions to withdraw would have become irreversibly fixed while the Arabs' remain reversible. This argument is modelled in the left-hand part of the game tree in Figure 3, which shows the Arab decision whether or not to give up terrorism following *after* a putative Israeli withdrawal.

The argument that Israel could trust the US not to abandon it is based on the assumption that if Israel doesn't withdraw, the US will eventually have to decide whether or not to carry out such a threat. This is modelled in the right-hand part of the tree, where the US decision follows the Israeli decision.

The preferences attributed to the parties at the tree's end-points are those cited by Israel in the above arguments. Note that the model focuses on two particular arguments against the agreement P+A. To do so, it has altered the details of Figure 1, adding the US as a player, lumping the Arab parties together as a single player, and omitting other details shown in Figure 1.

Figure 3: Model of a US-supported Arab-Israeli agreement

A game-theoretic analysis of this tree would point out that the agreement (shown by heavy lines in the tree) is impractical for two reasons. On the left, because Israel would see that if it allowed a viable Palestinian state, the Palestinians and Arabs would have no reason not to follow their preference by continuing terrorism (thereby choosing their Best future rather than their Next Best). On the right, because

Israel would see that if it refused to allow a viable Palestinian state, the US would follow its preference by not turning against Israel (thereby choosing its Next Best future rather than its Next Worst). Israel, therefore, would see that its real choice was between its Worst future(continuing terrorism from a viable Palestinian state) and its Next Best (no Palestinian state and continuing US support). Given this choice, it would obviously choose the latter. Thus, the agreement would not be implemented.

This game-theoretic analysis essentially sets out the logic of the two arguments that we've said Israel might use against the P+A agreement. This is because the doubts in Figure 1 are given formal justification in Figure 3. Of course, since this is game theory, the analysis is stated as a prediction as to what cards will be played. But the same analysis can be used for drama-theoretic prediction of how the model is likely to be altered. To see this, compare it with the analysis in Figure 4.

Figure 4 shows the kind of card table we use to analyse doubts (dilemmas of implementation) using extensive-form arguments. This is the general procedure. Take a "commitment" (a position or fallback); write it as a column; analyse the doubts that beset it. In Figure 4, the commitment is column a. It's the same as the agreement in Figure 3. Two doubts are analysed. They're the reasons in Figure 3 for doubting the agreement. The first doubt is whether the Israelis would allow a viable Palestinian state, given a promised cessation of Arab terrorism. The second is whether, if they did, the Arabs would in fact cease terrorism. Points to note are as follows.

- Each doubt is assigned an "objective". This is the future that doubtful characters are thought to be aiming for in implementing that doubt, or doubts.[8]

Its function in the analysis is to show why doubt exists. Columns oI and oA show suspected objectives of Israelis and Arabs. In each objective column, the doubts implemented to seek that objective are shown by arrow-heads (“>”).

- To each “deviation” (set of simultaneously implemented doubts with a common objective), there’s a stated response. This is what others say would follow that deviation. Columns rI and rA show these responses. They’re paired with the objectives they respond to.

- In each response column, some card-plays are surrounded by two vertical lines (“| |”). These are the card-plays regarded as “fixed” when the response is made. The deviation (set of implemented doubts) that triggered the response are surrounded by an arrow-head and a vertical line (“> |”). They too are seen as fixed when the response is made.

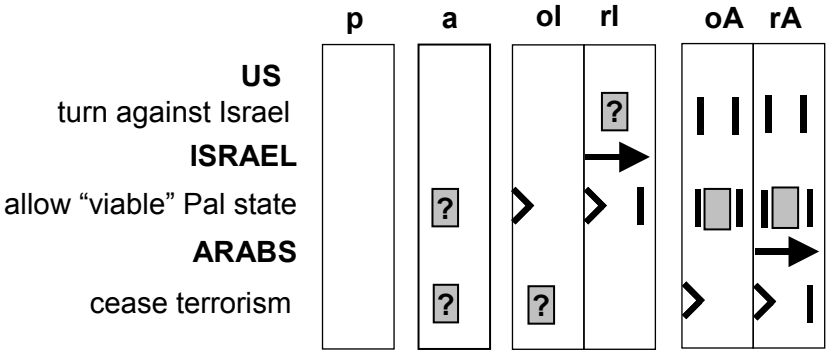

Figure 4: Card table modelling Arab-Israeli agreement

Note two assumptions, implicit in Figure 3, that are explicit in Figure 4. One is in column rI . It’s the assumption that if Israel doesn’t allow a “viable” Palestinian state, then the Arabs won’t cease terrorism. The other is in columns oA and rA . It’s the assumption that if it’s fixed that Israel will allow a Palestinian state, then it’s fixed that the US won’t turn against Israel.

Arrows in response columns indicate preferences, as they did in Figure 1. Question marks indicate doubts. The rule is: doubts affect preferences. Thus, the arrows in Figure 4 point to the drama-theoretic equivalent of the prediction in Figure 3—that the responses don’t eliminate the doubts. The reasoning is:

- Israel prefers column rI to column a because of the doubt in rI that the US would in fact turn against Israel and the doubt in a that the terrorism would cease.
- The Arabs are judged to prefer column rA to a simply because they’re judged to prefer to continue terrorism. Doubts don’t enter into this preference, as no further deviation are possible from rA, all card-plays being fixed.

Figure 5 shows the tug-of-war representation of Figure 4. From the “Land for Peace” agreement, an arrow representing doubts about Israel goes up to the objective, “Peace, no withdrawal”. From the same agreement, an arrow representing doubts about the

Arabs goes down to the objective "Withdrawal, no peace". From the first arrow, a dotted-line arrow goes to the response "No peace, US turns against Israel." The fact that this response (given its doubts) is preferred by Israel to the agreement is indicated by an arrow from the agreement to the response. Meanwhile, another arrow goes up from the response. This indicates doubt that the US would actually respond by turning against Israel.

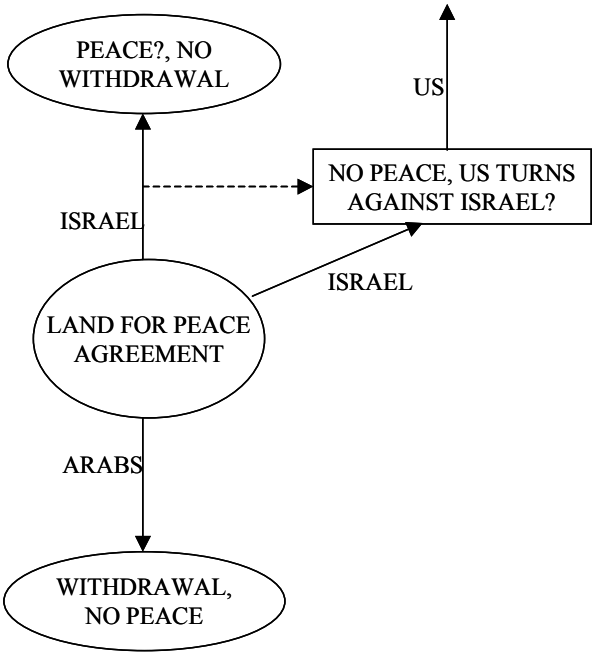

Figure 5: Tug of war modelling Arab-Israeli agreement

If desired, the analysis can be taken further. The doubt in column rI can be analysed in the same way as the agreement in Figure 4.

This is done in Figures 6 and 7. It's simple. We're analysing the commitment that if Israel doesn't allow a viable Palestinian state, the US will turn against it—even though, by assumption, Palestinian terrorism won't have ceased. This commitment is column c in Figure 6. It contains a doubt that the US would in fact turn against Israel. To confirm this doubt, we show the objective column oU. There's no possible response to pursuit of this objective, as all other card-plays are fixed. So the response column is the same as the objective. The arrow shows the US preference. As with the Arabs in Figure 4, the US prefers not to turn against Israel because: (a) it has this preference assuming no further response; (b) no further response is possible, all card-plays being fixed. (As in Figure 4, the assumption that Arab non-cessation of terrorism is fixed makes explicit an assumption implicit in Figure 3.)

Figure 7 is the corresponding simple tug of war.

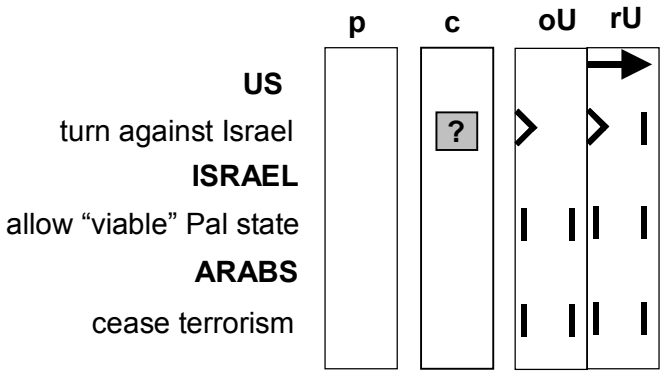

Figure 6: Analysis of doubt that US would turn against Israel

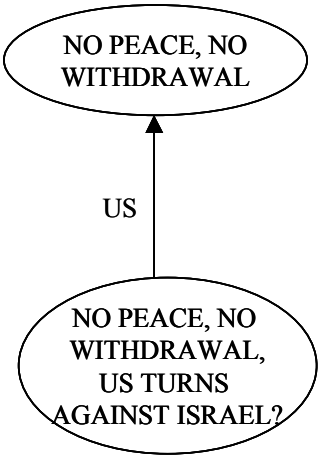

Figure 7: Tug of War modelling doubt that US would turn against Israel

DRAMA THEORY IN A TREE COMPARED TO GAME THEORY

We've shown, by an example, how to extend drama-theoretic analysis of dilemmas of implementation to make it "extensive"; that is, make it take into account the fact that some options become fixed before others.

What is the general extensive procedure? To analyse the doubts besetting an commitment (ie, an agreement or fallback):

- Write the commitment as a column *c*.
- Indicate doubtful cards in *c*—cards parties believe would be changed. (By a "change" we mean a change from "played" to "not played" or vice versa.)
- For each doubtful card, ask what other doubtful cards, decided simultaneously with it, would need to be changed to meet its objective. In this way, form sets of cards. Call these sets "suspected deviations"—or simply "deviations". (As an example, in Figure 1 the Arab states' card "fund terrorism" can't meet its objective if the Palestinian card "stop terrorism" is played. Hence, in the column P+A changing (playing) the card "fund terrorism" is not a suspected

deviation on its own, but becomes so when combined with changing (not playing) the card "stop terrorism". Thus, in column P+A the set {fund terrorism, stop terrorism} is a suspected deviation; but the one-element set {fund terrorism} is not.)

- For each suspected deviation, write a column showing the objective that would be sought by it. This objective column necessarily contains implementation of the deviation; other changes will be implemented in it if they're changes the deviators (the parties that control the deviation) think would follow.
- Next to each objective column, write a column showing the stated response to the deviation it contains. This is the response that non-deviators say would follow.
- Indicate doubts (if any) in each response column.
- Judge whether each deviation in the commitment is *deterred*. To do this, compare the commitment to the stated response to the deviation. Ask: is there some defector that prefers the commitment to the response? If so, the deviation is deterred. If not, not.

The output of this analysis: a statement as to whether and how deviations are deterred. If a deviation is deterred, a rationale for eliminating the corresponding dilemma has been found—ie, an argument as to why it shouldn't exist. The argument is to point out this deterrent. If not, a rationale will continue to be sought as parties try to eliminate the dilemma. Thus, the analysis shows how parties will try to redefine their situation.

Note that the objective columns are theoretically unnecessary. They're there to confirm doubts and construct deviations (simultaneously implementable sets of doubts with a common objective). What really matters is the deviations and the responses to them. If deviations were directly observable, we could dispense with doubts and objectives. Given deviations and responses to them, we could judge whether the responses deter the deviations. In practice, doubts are what we most easily observe. We check them and construct suspected deviations by attributing objectives to doubtful players.

After analysing doubts in an agreement or fallback, we can use the same method to analyse doubts in responses—and so *ad infinitum*. But we need not do so. Each decision whether or not to take the analysis further in this way is a modelling decision. It's a decision whether or not to use formal modelling to answer questions raised by doubts—questions that take us outside the present model. They could be answered using formal methods or simple intuition. Sometimes formal methods will be useful, sometimes not. In any case, sooner or later formal methods must give way to direct knowledge or intuition about the environment.

Comparison with game theory

This, then, is drama-theoretic analysis of dilemmas of implementation. How does it compare with game-theoretic analysis?

To make the comparison, assume that the two approaches are used to model and analyse the same real-world situation, and that both make completely accurate empirical observations of it. Under this assumption, the differences between the two

are set out in Figure 8. This shows that they differ in what they try to predict, what they take as given and why and when they make empirical observations.

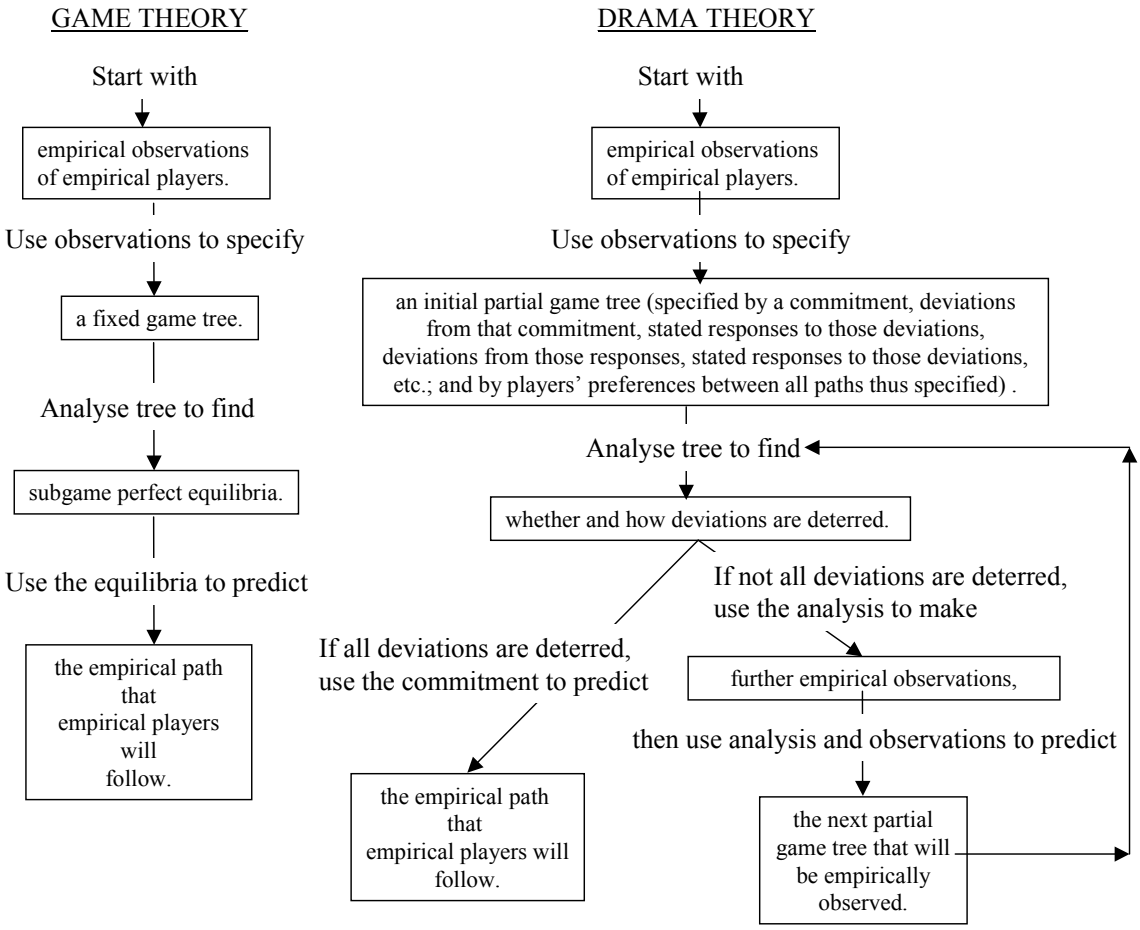

Figure 8: Comparison between game-theoretic analysis and the drama-theoretic analysis of dilemmas of implementation

To begin with, both game theory and drama theory make empirical observations. But their observations are different.

Game theory builds a complete game tree. It does so by observing all the sequences of actions available to the players and their preferences between them. Drama theory builds a partial tree of the same tree (ie, a tree comprised of a subset of the tree's paths). It does so by observing only players' stated commitments regarding their actions, stated doubts regarding those commitments, stated responses to those doubts, stated doubts regarding those responses, and so on as far as players' statements go. It also observes their preferences between the sequences of actions (the commitment and responses) that their statements refer to.

Figure 9 illustrates. It gives an example of a game tree. This is a game with perfect information and simultaneous moves. In the tree, a drama-theoretic analysis has been picked out in heavy lines.[9] This defines (in the obvious way) a partial game.[10]

To make all clear, Figure 10 shows how the same tree would be modelled by a card table. Here, columns represent paths through the tree, 1s denote cards that are played,

0s denote cards that are not played, and tildes ("~") denote cards that are undecided (neither played nor not played) in that particular column. The drama-theoretic analysis is in the first three, boxed columns.

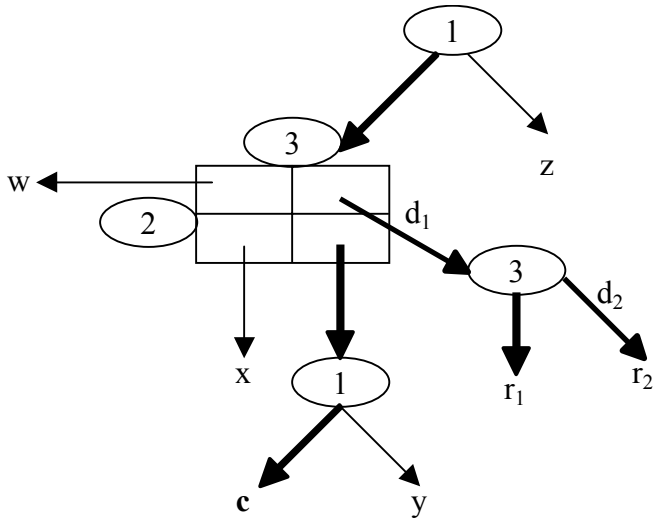

Figure 9: Example of a game tree showing commitment, deviations and responses

Player 1	Left	**1**	**|1|**	**|1|**	**1**	**1**	**1**	**0**
2	Top	**0**	**>1|**	**|1|**	**1**	**0**	**0**	~
3	Left	**0**	**0**	**|0|**	**1**	**1**	**0**	~
1	Left again	**1**	~	**|~|**	~	~	**0**	~
3	Left again	~	**1**	**>0|**	~	~	~	~
		c	r_1	r_2	w	x	y	z

Figure 10: Card table representation of tree and partial tree in Figure 8

Observing players' common knowledge

Now we've said that drama theory listens to ("observes") players' statements—ie, what they say about their commitment and their responses to deviations. These are their threats and promises to each other. By listening to them, drama theory observes the partial tree that consists of certain paths—viz, the commitment and responses.

This, we would argue, is the only the directly observable part of the tree. This is our argument. The game tree is the players' subjective tree; it's what the players take to be common knowledge among themselves. But how do we know what they take to be common knowledge? *The meaning of a threat or promise has to be common knowledge; if not, it fails to be a threat or promise.* Thus, by listening in to their threats and promises, we are observing what they take to be common knowledge. The rest of the tree may be estimated in various ways. It isn't directly observed, because we don't directly observe it to be common knowledge.

"Solving" the model: predicting what will happen

So far we've discussed the different models built by game theory and drama theory. Game theory builds a complete tree. Drama theory builds a partial tree of the same tree. It picks out certain paths—a commitment and a set of responses.

After building their models, both approaches "solve" them. That is, they analyse them for a mathematical object. Again, they differ. They solve for different objects.

Game theory solves for "equilibria".[11] These are choices at each node of the tree that are optimal for the chooser when at that node, given the other, subsequent choices. Having found these, it predicts the empirical path that empirical players will follow: *an equilibrium path*. This is a path followed by making a choice, at each node, that belongs to an equilibrium.

Multiple equilibria are possible. They pose the problem of which equilibrium path to predict. If, as we assume, players have committed themselves to a particular equilibrium, this solves the problem. Game theory predicts the corresponding equilibrium path—the commitment.

Drama theory looks for a more complex—and incomplete—solution. First, it solves for "deterrence". It asks: are the responses deterrents? That is, does each response deter the deviation it is a response to?

If so, then drama theory, like game theory, predicts the empirical path that empirical players will follow: the commitment.

Otherwise, drama theory looks for a different kind of solution and makes a different kind of prediction. But before we discuss this, let's compare the predictions made when responses do or do not deter.

If responses deter, does drama theory, within the partial tree that it observes, give the same prediction as game theory? In other words: *if responses deter, is the commitment an equilibrium path of the partial game defined by the commitment and the responses?*

Answer: Yes, because if the responses deter, the commitment is more than an equilibrium path. It is a *strict, strong, equilibrium path*: a path determined by a *strict, strong, equilibrium* of the partial tree.[12]

Proof: Partition the paths belonging to the partial tree into sets $R_1, R_2, \ldots$ as follows:

$R_1 = \{c\}$, where c is the commitment;
$R_n = \{r \mid r$ is a response to a deviation from an element of $R_{n-1}\}$.

Now select a choice from each node of the partial tree as follows:

> At each node that belongs to a path in R_1, choose the arc that is not a deviation from that path;

at each remaining node that belongs to a path in R_2, choose the arc that is not a deviation from that path;
…
at each remaining node that belongs to a path in R_n, choose the arc that is not a deviation from that path;
and so on.

Then this selection of choices constitutes a strict, strong equilibrium that determines the path *c*.

In Figure 9, for example, we can see that the *c* is a strict, strong equilibrium of the partial tree picked out in heavy lines provided that player 2 prefers *c* to r_1 and 3 prefers r_1 to r_2.

"Solving" the model when responses don't deter

What if the stated responses to deviations are not deterrents?

In this case, we might expect the commitment not to be an equilibrium path of the partial tree. There are, however, three reasons why it may actually be an equilibrium path, even though the stated responses are not deterrents.

First, the "strictness" criterion may fail. A deviation (or rather, the stated response to it) may be *indifferent*, for some player, to not deviating. Drama theory, unlike game theory, considers this to be a problem. If a player is indifferent to whether it deviates, it is just as likely to deviate as not. Players that want to make the commitment credible will want to eliminate this possibility. Drama theory must try to predict how they will do so.

Second, the "strength" criterion may fail. Some subset containing *more than one player* may be able to gain by a deviation, even though no individual can. Again, drama theory treats this as a problem that players trying to make the commitment credible will want to rectify. It must try to predict how they will react.

Finally, it may be the case that the *rationally anticipated response* is a deterrent, even though the *stated response* is not. For example, suppose that in Figure 9, player 3 prefers r_2 to r_1. Then the stated response r_2 does not deter the deviation d_2 from the response r_1. Yet if player 2 prefers *c* to r_2, the rationally anticipated response r_2 does deter the deviation from *c*, so that *c* is an equilibrium path in the partial tree. Game theory would therefore predict that it will be followed.

Drama theory does not. Instead of assuming, like game theory, that the tree is fixed, it assumes that players may try to make their commitments credible. In trying to do so, they may *change* the tree—though how far they're able to do so depends on properties of the environment not represented in the model. Now other players know this. Hence, the fact that 3 has committed itself to r_1 will, in general, make 2 hesitate to take for granted that 3 will choose r_2 instead.

Thus, drama theory again sees a diminution of credibility where game theory sees none. It sees the credibility of the commitment diminished by the fact that a *stated*

response fails to deter, even though the rationally expected response does so. It therefore tries to predict the reactions of players to this lack of credibility.

So there are three cases when drama theory, seeing credibility deficits that game theory doesn't, questions a game-theoretic prediction about an equilibrium. They are: when the equilibrium isn't "strict"; when it isn't "strong"; and when it isn't "stated".

Now the first and last cases may seem easy to rectify.

Take the first. If a player is indifferent between keeping its commitment or not, the fact that it has made a commitment would, it seems, tip the balance. The "normal" drama-theoretic prediction will, indeed, be that an otherwise-indifferent player will prefer to carry out its commitment, once made, and that this preference will be common knowledge.

Doesn't this amount to adopting the game-theoretic "non-strict" criterion? Not quite. First, it's a drama-theoretic prediction—a prediction as to how players will change the game. Such a prediction is not game-theoretic. Second, it isn't absolute. It does depend, in certain cases, on the environment. A player may actually have a flirtatious "meta-preference" for remaining indifferent in order to keep others guessing.

Take Scarlett O'Hara, heroine of the film *Gone with the Wind*, as she plays with the feelings of three beaux. She has to go the ball with one, but makes it clear she doesn't care which. Having made her selection, she doesn't want that commitment to make her seem any the less indifferent. Figure 11 models her situation.[13]

Scarlett	choose 1st beau	**1**	**>0|**	**>0|**
	choose 2nd beau	**0**	**>1|**	**|0|**
	choose 3rd beau	**0**	**|0|**	**>1|**
1st beau	offer himself	**1**	**|1|**	**|1|**
2nd beau	offer himself	**1**	**|1|**	**|1|**
3rd beau	offer himself	**1**	**|1|**	**|1|**

		c	r_1	r_2
Payoffs to:	Scarlett	1	1	1
	1st beau	2	1	1
	2nd beau	1	2	1
	3rd beau	1	1	2

Figure 11: Scarlett O'Hara at the ball

Game theory, given players' commitment to *c*, would predict *c*. After all, it's an equilibrium singled out by the players. Drama theory notes that Scarlett may switch to another beau: *c* isn't strict. Drama theory therefore predicts a situation full of emotion, with the beaux trying to change the game (ie, change Scarlett's preferences; it should be easy; the smallest change will suffice). This, of course, is just what Scarlett wants. This illustrates the point that players in a drama may not want complete resolution. They may want the emotional buzz that final resolution drains away.

Next, consider equilibria that aren't "stated". Surely these can be made credible by stating as responses the deterrent subpaths that *are* credible. Take Figure 9. If player 3

prefers r_2 to r_1 (so that the response r_1 does not deter the deviation d_1), yet player 2 prefers c to r_2 (so that c is an equilibrium path), then all that's necessary is for player 3 to state r_2 (instead of r_1) as its response to d_1.

And indeed, such a change of stated response will be the "normal" drama-theoretic prediction. It's one that requires no change in the game, in positions or in the fallback, but only in player's statements about their responses. But this prediction also is not absolute. It may fail in certain environments. A Scarlett O'Hara type player may again want to "tease" others—create confusion and mixed feelings—by stating a response that not only fails to support its position, but is worse for itself than one that does. It may do so simply in order to tease, or because it has more serious reasons for confusing others and postponing a resolution.[14]

The middle case discussed above—when the commitment is an equilibrium path, but not a "strong" one—isn't so easily dealt with. However, there's no real difference with game theory here. It's true that much (not all) game theory restricts deviations from an equilibrium to deviations that can be carried out by individuals. But game theory naturally admits that an equilibrium vulnerable to profitable deviation by a *coalition* is to that extent less stable. Drama theory merely adds that if the tree isn't fixed, players will use emotion and rationalisation to try to remedy such vulnerability.

The essential difference between game theory and drama theory lies in what drama theory does in the general case when responses fail to deter—including the case when the commitment is not an equilibrium path.

This is the case when both game theory and drama theory predict that the commitment will not be implemented. It is the case, for example, when player 3 in Figure 9 prefers r_2 to r_1, while 2 prefers c to r_1 but r_2 to c. Game theory now not only predicts that the commitment won't be implemented. It asserts that the players know it, and know that each other knows it, etc. Their commitment therefore evaporates. They cannot be said to make it, knowing as they do that it won't be believed (and that others know that they know this, etc). They will make some other commitment—or none at all.

What does drama theory do? It identifies the vulnerabilities in the commitment—the responses that are not deterred—and the players that face dilemmas due to them. In the case of a position, these will be the players that take that position; they face trust or cooperation dilemmas. In the case of a fallback, the players that face dilemmas are the "undeterred deviators" from the fallback. These players face threat dilemmas.

Drama theory then examines the environment and asks what these players can do to eliminate the vulnerability. The players—it asserts—will be driven by emotion to look critically at the assumptions underlying the asserted vulnerabilities and ask whether they are supported by evidence and reason. The analyst must empathise, put itself in the players' places and do the same. In this way it can predict player reactions.

This is just what a scientist does in questioning assumptions that forbid a theoretical advance. Just as a scientist may overthrow assumptions and discover a new theory, so players may overthrow assumptions and eliminate vulnerabilities. Analysts, therefore, can follow them and foresee what they will do.

Whether players can succeed in overthrowing assumptions obviously depends crucially upon the environment, not merely upon their desire to succeed. The latter, however, is all that formal analysis of the *given* tree can predict.

Thus, drama theory makes predictions as to how players will react to dilemmas—but only after making further empirical observations. It may predict, like game theory, that players will choose another commitment, finding this one impossible to sustain. This, of course, involves changing the partial tree they communicate about, though not necessarily the complete tree. But this is only one possibility. Other possibilities involve changing the game (the complete tree) to make the commitment sustainable.

In general, as indicated in figure 8, drama theory predicts the next partial tree that will be constituted by players' communications to each other. It does so by predicting the emotions they will feel and the rationalisations these emotions will motivate, *given* the environment they have to contend with.

REFERENCES

1) Baan A, Howard N, Tait A, ' "COM as Shooter" - Operational Planning using C2 for Confronting and Collaborating'. CCRTS Symposium 2003; available from www.dodccrp.org.

2) Bryant, J. The Six Dilemmas of Collaboration.Wiley, 2003.

3) Crannell, M., Howard, N; Norwood, G.W. and Tait, A., 'How to assess and exit strategy: Measures of Merit for compliance'. CCRTS Symposium 2004; available from www.dodccrp.org.

4) Howard, N. 'Drama Theory and its Relation to Game Theory. Part 2: Formal Model of the Resolution Process.' Group Decision and Negotiation, 3 207-235. 1994.

5) Howard, N. 'How to Win Peace Operations: Theory Vs Practice.' CCRTS Symposium, 2000; available from www.dodccrp.org.

6) Howard, N. 'n-person "soft" games.' Jour. of Operational Research Soc, 49 144-150. 1998.

7) Howard, N. Confrontation Analysis: How to Win Operations Other Than War. CCRP Press, 1999; available from www.dodccrp.org.

8) Howard, N. Paradoxes of Rationality: Games, Metagames and Political Behavior. MIT Press 1971. Available as MIT Classic.

9) Murray-Jones P., Howard N. 'Co-ordinated Positions in a Drama-theoretic Confrontation: Mathematical Foundations for a PO Decision Support System.' CCRTS Symposium, 2001; available from www.dodccrp.org.

10) Murray-Jones P., Howard N. 'Foundations of a Decision Support System for Peace Operations: Designing Experiments to Test the Psychological Hypotheses

of Confrontation Analysis', CCRTS Symposium, 2001; available from www.dodccrp.org.

11) Murray-Jones, P., Stubbs, L. and Howard N. 'Confrontation and Collaboration Analysis: Experimental and Mathematical Results'. CCRTS Symposium, 2003; available from www.dodccrp.org.

12) Osborne, M.J. and Rubinstein, A. A Course in Game Theory. MIT Press, 1994.

13) Rosenhead, J., ed. Rational Analysis for a Problematic World. Wiley, 1989.

14) Smith, R., Howard, N. and Tait, A. 'Commanding Anti-Terrorist Coalitions: A Mid-East Illustration.' CCRTS Symposium, 2002; available from www.dodccrp.org.

15) Smith, R., Howard, N. and Tait, A. 'Confrontations in War and Peace.' CCRTS Symposium, 2001; available from www.dodccrp.org.

FOOTNOTES

[1] Defined as a point from which no coalition can defect without at least one member becoming worse off.

[2] The controversy about intelligence-gathering prior to the Iraq war illustrates the depth of our society's resistance to the principles of scientific method, as understood by those that have studied them. Intelligence services were criticised for collecting intelligence to prove a case, rather than in a disinterested way. A valid criticism might be that intelligence services did not, in their efforts to prove their case, try hard enough to undermine it (on the "devil's advocate" principle that to show how hard it is to undermine a case is the best way to prove it). But no one made this criticism. Everyone criticised them for having a case to prove—ie, for not being disinterested.

[3] Quotes, in a card table, are put round words the use of which might be challenged by some parties.

[4] Obviously, different Palestinians and Arabs, like different Israelis, advocate different positions, fallback strategies and preferences. We nevertheless see our parties as having overall positions, fallback strategies and preferences. These are arrived at as a resultant of their internal conflicts and agreements. These internal conflicts and agreements can be analysed, if desired, using the same methods as are used to analyse the overall conflict or agreement.

[5]The playing of a card (shown by a shaded rectangle) has a somewhat different meaning in column p than in the other columns. That's why, in p, it is shaded black. In p it refers to *actual* intentions, those actually being carried out, regardless of what is stated. In the other columns, the playing (or not) of a card represents a *stated, contingent* intention of the form "if you do this, I'll do that."

[6] At least, you believe it would be implemented; if you didn't believe this, we'd have to say there were doubts attached to your position. In the absence of doubts, you are certain. (Of course, you may be wrong, even though certain. You are not infallible.)

[7] These are the five main dilemmas. In addition, there are two subsidiary dilemmas—the positioning dilemma and the sincerity dilemma These are faced, respectively, by a party that prefers another's position, or the fallback, to its own position. They are connected to each other and to the rejection dilemma. The relationship is as follows: I must have a positioning dilemma if I have both a sincerity dilemma and a rejection dilemma; I cannot have a positioning dilemma if I have neither of these.

[8] In general, a "deviation" (a set of simultaneously implemented doubts) is required in order to meet an objective. In Figure 4, however, each deviation involves a single card.

[9] A game with perfect information and simultaneous moves is defined by

- a tree;
- a set of I of players with preferences over the end points of the tree;
- for each node x of the tree, a set
  $S^x = \prod_{i \in I} S_i^x$
  of simultaneous, independent choices by the players together with a function f^x specifying, for each s in S^x, an arc $f^x(s)$ leading from x that is followed if s is chosen.

[10] A partial tree of the game tree does not generally define a game with simultaneous moves unless we specify the simultaneous choices at each node x that lead to each arc of the partial tree that begins at x. In the case of the partial tree defined by a commitment together with a set of deviations and responses, the corresponding partial game is defined in the obvious way. That is, if the commitment-or-response at node x requires making the choices $(\bar{s}_i \mid i \in I)$, while a deviation at this node consists of the choices $(s_i' \mid i \in I)$, then in the partial game the set of choices for player i at node x is $\{\bar{s}_i\} \cup \{s_i'\}$; the choices $(s_i' \mid i \in I)$ lead to the deviation; and all other choices lead to the arc that belongs to the commitment-or-response. In other words, a deviation is chosen if and only if all deviators choose it: and if there is no deviation, none is chosen.

[11] By "equilibrium" we will always mean what game theorists call "subgame perfect equilibrium". We will not discuss any other kind.

[12] "Strict" denotes "requiring compliance to be strictly preferred—rather than preferred-or-indifferent". "Strong" denotes "after allowing for actions by coalitions (subsets) of players—rather than only by individuals". Thus, a strict, strong equilibrium is the choice, at each node, of an arc such that *each subset of players* that could make an alternative choice has *at least one member* that finds (assuming no other deviating choice will be made) that the given choice leads to a path *strictly preferred* to the path the alternative leads to. Clearly, a strict, strong, equilibrium is an equilibrium that obeys some further conditions—ie, it is strict and strong.

[13] Scarlett has committed herself to the first beau, but two deviations are discussed by the players: she might switch to another beau. No beau ever discusses ceasing to offer himself, so in the partial tree they are given no choice and the game is effectively a one-person game. It's modelled as a simultaneous-choice game; that is why all choices, following a deviation by Scarlett, are fixed.

[14] "Teasing" may be defined as the deliberate postponement of the satisfactory resolution of a conflict or collaboration problem. Like Scarlett, people tease because they enjoy the emotional buzz of an unresolved situation. They prefer it to the absence of any need for emotion or debate that characterizes complete and perfect resolution. Marriages and relationships are often kept alive by refusing to be "serious" in this way—ie, by deliberately not resolving certain fundamental conflicts.