Co-ordinated Positions in a Drama-theoretic Confrontation: Mathematical Foundations for a PO Decision Support System

Peter Murray-Jones
(DERA)
Nigel Howard
(dramatec)
12 Chesham Road, Brighton BN2 1NB, UK
Tel.: +44 1273 67 45 86
e-mail: nhoward@dramatec.com)

This report has been funded under the Corporate Research Programme, TG11, of the Ministry of Defence, UK

Abstract

Confrontation Analysis (CA) has an essential mathematical foundation. It is based on Drama Theory, a development of Game Theory. Instead of analysing a game-type model for 'solutions', CA analyses it for 'dilemmas'. These show where and how the model is likely to prove inadequate as players re-define their situation in order to eliminate their dilemmas. Fundamental theorems proved about dilemma-elimination have assumed that each player's 'position' (the solution it advocates) is a single outcome. A typical confrontation in Bosnia illustrates the fact that a player's position may offer others a number of outcomes. Research reported in this paper has generalised the fundamental theorems to the case of 'general' positions that may contain more than one outcome. We argue that general positions may be assumed to be 'co-ordinated'; this concept is defined and discussed. Finally, we show how our results are used in designing a C2 system for confronting Non-Compliant Parties in a Peace Operation.

Introduction

Confrontation Analysis is an approach to modelling situations involving multiple parties with conflicting objectives. The technique can be used to:

- Assist a particular party in negotiating a desirable resolution to a confrontation; or
- Support a third party attempting to mediate between warring factions.

The technique makes use of a mathematical framework to expose *'dilemmas'* in a confrontation. Reactions by which parties try to resolve these dilemmas are identified, giving rise to a new confrontation 'frame'. Confrontation Analysis is based on Drama Theory, a development of *Game Theory* that analyses situations in which players communicate prior to the game. It analyses the effect of their communications in changing their beliefs and preferences. It does so using the assumption that changes are brought about when players face 'dilemmas'.

Confrontation Analysis[1] has been endorsed by senior NATO serving officers as a method for analysing Peace Operations (PO). A decision support/C2 system for PO has been proposed based upon confrontation-analytic techniques[2]. The aim of this first year (of a three year) research project, funded under the Ministry of Defence Corporate Research Programme (CRP) Technology Group (TG) 11, was to further develop the mathematics required to provide a basis for the proposed system.

This project, then, is to extend the mathematical foundation of Confrontation Analysis to provide techniques needed to build a multi-level information and analysis system that would provide appropriate support to a PO commander and his staff.

Dilemma analysis

Dilemmas are faced when players arrive at a 'moment of truth' – a point when each has taken a 'position' (i.e., a recommended solution) that it regards as final. At this point, positions taken may be compatible or incompatible.

- If positions are compatible, players go into *collaborative* mode, where they face dilemmas of 'trust' and 'co-operation' roughly describable as follows
    - I face a 'trust' dilemma if I believe you would be better off *not* implementing the common position, were I to do so.
    - I face a 'co-operation' dilemma if you face a trust dilemma in relation to me – i.e., if you believe I would be better off not implementing our common position.

- If positions are incompatible, players go into *confrontational* mode and adopt 'fallback strategies' (policies that they claim they will implement if an agreed position cannot be reached). They then face 'threat', 'deterrence', 'inducement' and 'positioning' dilemmas, roughly describable as follows.
    - I face a 'threat' dilemma if you believe I would be better off *not* carrying out my fallback strategy, were you to carry out yours.
    - I face a 'deterrence' dilemma if I believe you would be better off at the 'fallback' future (the future that would result from us both carrying out our fallback strategies) than at my position.
    - I face an 'inducement' dilemma if you do not face a deterrence dilemma in relation to me – i.e., if I might be better off at your position than at the fallback.
    - I face a 'positioning' dilemma if I actually prefer your position to my own (but am forced to argue against it on the grounds that it is 'unrealistic').

Drama theory asserts that changes in players' beliefs and attitudes are brought about, during pre-play communication, by rational-emotional attempts to eliminate these dilemmas. This assertion is based on the fact that only by eliminating the dilemmas they face can players arrive at a true resolution of their interdependent decision problem.

In support of the assertion, it has been proved[4] that if no dilemmas exist, then players have a common position that is a 'strict, strong equilibrium' – meaning in game-theoretic terms that it has a high degree of stability and is strongly self-reinforcing. If, on the other hand, players do not succeed, during the process of pre-play communication, in eliminating all their dilemmas, then they must go either into conflict (declaring that they will carry out their fallback strategies, and either doing so or failing to do so if the conflict is 'flunked'). Or they must go into a scenario of 'false' co-operation (agreeing to carry out a common position without intending to do so or trusting one another to do so).

A decision support/C2 system for PO commanders

This, in brief, is the theoretical framework underlying the proposed decision support/C2 system. The paradigm proposed for PO is that the Main Effort for a PO commander lies in confronting NCPs (Non-Compliant Players) in order to bring them into compliance with the will of the International Community. Hence, by analysing the dilemmas that face him and the NCPs involved in such confrontations, he can plan a more effective strategy for achieving his objectives.

The proposed system will be based upon the use of 'card tables' as the appropriate tool for planning and storing information about confrontations. It will be an email-based system linking commanders with each other and with civilian agencies so as to enable a more co-ordinated approach to the problem of obtaining NCP compliance. It will be a large, networked system for storing, updating, analysing and passing on information relevant to confrontations.

This paper sets out some basic mathematical definitions and theorems that will be required by the system. Models of two drama-theoretic 'moments of truth' are used for illustration. Though simple, these models are nevertheless realistic, as moments of truth *must* be simple so that characters can be sure they understand each other. The drama-theoretic approach is to model the simple, stripped-down frames within which understandings are reached, in order to see the pressures for change that arise within them, and thereby to look for changes that are realistic because they respond to pressures experienced by characters. Simple models of moments of truth, because they are realistic, are readily appreciated by decision-makers, yet are capable of supporting complex information structures. The proposed system will use simple models to communicate with decision-makers, backed up by complex ones used by staff to store, analyse and pass on information.

In this paper we:

- Re-prove the basic theorems of drama theory for the case when positions are 'general'
- Explore the implications and justification of the assumption that in real-world applications, 'general' positions are always 'co-ordinated'

- Do the above within a card-table model, rather than the usual game-theoretic model
- Discuss the implications of the above mathematical work for command and control of PO.

Positions: are they single futures or sets?

Drama theory asserts that characters reach a '*moment of truth*', at which there is pressure for beliefs and values to change – when they have taken 'final' positions within a '*common reference frame*'. Which is to say a minimal common set of beliefs about the 'frame' they currently perceive themselves to be in. As said, a model of a moment of truth must be simple, as each character must make sure that it knows that the other knows… etc… what it means – and they can only be sure of this within a simple common reference frame. What is decided within each simple frame has, however, complex consequences and implications for other players; hence, the military requirement is for a complex, multi-level support system built upon the use of simple models.

In previous treatments, a character's 'position' at a moment of truth has consisted of a single future within the frame. Often, however, this does not make sense: the Bosnian peace-keeping example in Figure 1 shows why a position must sometimes be defined as a *set* of futures.

In this card-table model of a confrontation between opposing positions, a Serb mayor has the responsibility for re-connecting utilities to a deserted village previously occupied mostly by Croats. The Croats in this were 'ethnically cleansed' by the Serbs after committing atrocities against them and now wish to return, as provided for in the Dayton agreement. The mayor, in order to secure reconstruction aid, has said he is prepared to support the returns, provided they are delayed until he has re-connected utilities. The Croats, distrusting the mayor, will not agree to this delay, but insist on returning at once, while utilities are being connected.

The card table used to model this situation is based on the 'tableau' method introduced in Howard[3]. Characters are listed at left with the 'cards' they can play (i.e., their yes/no policy options) listed beneath their names. For example, the international community – a coalition of SFOR (NATO's Stabilisation Force) and various civilian agencies – is a character in Figure 1. It decides whether or not to play the card 'stop reconstruction aid'. Having listed players, the playing-or-not of a card can be shown in various ways. In Figure 1, playing a card is shown by a 1 on an upturned card, not playing it by a 0 on a face-down one. A column of 1s and 0s is then interpreted as representing a possible future – that expected if the policies with value 1 are implemented, while those with value 0 are not.

This allows us to represent a position containing just one future; for example, the Serb mayor's position (first column) is that he will connect up the utilities and support returns provided these are delayed till after connection of utilities; aid from the IC should then not be stopped. However, each of the other characters does not take a position on one of the issues – the Croats take no position on aid, the IC take no position on delaying returns. This is shown by the tilde '~' in place of these cards. Thus, their positions contain more than one possible future.

	SERB	CROAT	IC	threat	default
SERB MAYOR	2	3,5	2,3	4	1
connect utilities to village	1	1	1	0	0
support returns	1	1	1	0	0
CROAT ETHNIC PARTY	5	1,2	1,5	4	3
delay returns (till after connection)	1	0	~	0	0
IC (INTERNAT'L COMMUNITY)	2	1,3	1,2	5	4
stop reconstruction aid	0	~	0	1	0

- A deserted village, where Croats were previously in the majority, has had its houses repaired by the International Community (IC), but utilities have not yet been connected. This has to be done by the Serb mayor, who will not support the Croats returning. His lack of support would mean violence if they were to try.

- However, after the IC threatened to stop reconstruction aid, the mayor agreed to their return provided it is delayed until he has connected utilities. His present position (column **SERB**) is acceptable to the IC, but not to the Croats, who do not trust the Mayor. It is that if the IC does not stop the aid he will connect utilities and support returns provided the Croat ethnic party (which decides when to send returnees) will delay sending them until after connection of utilities.

- The Croat position (column **CROAT**) is that they (the Croat ethnic party) should not delay sending returnees, while the Mayor proceeds with connection of utilities. They take no position on reconstruction aid. The IC position (**IC**) is that they are willing to continue aid whether the refugees return before or after connection of utilities, provided the Mayor does connect the utilities and supports returns.

- The threatened future, or fallback, (column **threat**) is that the Mayor, not receiving aid, will not connect utilities nor support returns, while the Croats will not delay sending the returnees. This is also the default future (column **default**), except that aid has not yet been stopped. The IC is hoping not to have to take this step, which would be difficult and embarrassing, though not impossible, to reverse.

Figure 1: Card table showing a Bosnian moment of truth

Character's *preferences* between futures on a card table are indicated by numbers written on the same line as their names, number 1 being assigned to the most preferred future shown, 2 to the next most preferred, and so on. A column containing multiple futures is given two rankings – the rankings of the most and least preferred futures in that column.

Figure 1 is a simple model of a moment of truth. When used to support decision-making, such a model acts as a simple structure within which to store all kinds of information about the situation; this information is accessed by clicking on a computer screen showing the card table in order to open up windows of information about particular aspects. In addition, the complex information system being considered would require this model to be developed and elaborated by support staff, who would explore its implications for lower-level commanders and their coalition partners by adding large numbers of characters and cards. Thus, the system would require large, complex models of this type, in addition to the simple ones used to support decision-making.

Here, the example illustrates the point that a character's 'final' position need not involve taking a position on every issue: it may not care much about certain issues, or may feel it has no right to take a position on them. Even if it cares about them and feels entitled to take a position, it may be willing to be flexible on certain issues in order to reach an agreement. Thus, the Croats take no position on whether the 'aid' card should be played. They prefer aid to be given if their refugees return peaceably, not if not. But they do not demand this; they simply take no position on this issue. Similarly, the IC takes no position on the question whether returns should be delayed until after utilities have been connected. Being anxious for the other two parties to agree they are willing to go along with any agreement on this point. Hence, two different futures are compatible with (i.e., fall within) the IC's position, two with the Croat position.

Should positions be representable by single columns?

Thus, there is a need to extend the theory to the case of 'general' positions that may contain more than one future. This requires defining a character's position as a *set* of futures, and re-proving the theorems of reference 4 to cover this case.

But an arbitrary set of futures will not generally be representable by a single column, like the positions in Figure 1. Can we expect the set that stands for a character's *position* to be representable in this way – particularly in a large, complex model? Figure 2 gives an example of a position – that of the Defence Department in an interoperability problem – that is not single-column representable. Is it a reasonable model of the department's position? (As before, the simple model in Figure 2 represents a complex reality, information about which would be accessed by clicking on the model to bring up details on particular aspects.)

A related question about general positions is this. If each player chooses a selection from its own cards that is compatible with a position P, is the future that results necessarily compatible with P?

If so, we may call the position in question *co-ordinated.* All the positions in Figure 1 are co-ordinated. However, the department's position in Figure 2 is not. The result is that the characters acting together may fail to implement this position while each claiming to do so; each service may use just its own system and claim to be implementing the department's position, even though the future that results falls outside that position.

Compounding this problem is another that the department faces because its position is uncoordinated. The positions of all three characters in Figure 2 are incompatible – i.e., no future belongs to all of them. Yet no two positions are incompatible; any two characters are able to agree.

	A	B	DEPT		threat	default
SERVICE A	1,2	2,4	1	3	4	2
discontinue A's system	0	~	0	1	0	0
use B's system	0	~	0	1	0	0
SERVICE B	3,5	1,3	4	1	2	3
discontinue B's system	~	0	1	0	0	0
use A's system	~	0	1	0	0	0
DEFENCE DEPT	2,5	1,4	2	1	4	3
oppose A's system	0	~	0	0	1	0
oppose B's system	~	0	0	0	0	0

Comments on this card table

- Services A and B have developed systems that are not interoperable. Each service uses its own system, demands that the Defence Department not oppose it, and takes no position on which system the other service should use or whether the Department should oppose the other's system. (For the Department to oppose a system will make it hard to get funding.)

- To achieve interoperability, the Department takes the position that either system is acceptable, provided both services use the *same* system; thus it takes the uncoordinated position shown in the two columns headed **DEPT**. If this position is not accepted, the department threatens to oppose A's system, Thus its fallback puts pressure on A to accept its (the Department's) position. Having got A's acceptance, it plans to put pressure on B.

Figure 2: Interoperability conflict showing the effect of an uncoordinated position

This is a problem for the department. The services agree at the default future, in which things carry on as they are. The department needs to change this. Yet it cannot find any point on which any one party disagrees with another. This makes it difficult to put pressure on them, or ask them to change. Where do you ask a character to change, when there is no point on which it disagrees with any other?

These are problems with uncoordinated positions. To deal with them, we shall in the next section prove a number of theorems that show, firstly, that uncoordinated positions can always be modelled (we suggest more satisfactorily) as co-ordinated ones and, secondly, that when this is done, the problems discussed do not arise. In the final section, we shall re-prove the basic theorems of drama theory for the case of general, co-ordinated positions.

Definition of a frame

We begin our mathematical treatment by defining a frame – the drama-theoretic equivalent of a strategic-form game.

Formally, define a *frame* as a pair

$(h, >)$,

where h is a function (the *holding* function) from a set of cards to a set of characters, and > is a family (or "indexed set") of preference relations, one for each character. The function h and the family > have the following properties.

<u>The holding function</u> $h: D \to C$ assigns *cards* in a set D called the *deck* to *characters* in a set C called the *cast.* The interpretation is as follows: hd (the "holder" of card d) is the character that controls (decides whether to play) d. In Figure 1, for example, (using shortened names for both cards and characters):

D = {connect, support, delay, stop aid}; C = {Mayor, Croats, IC};

h(connect) = h(support) = Mayor; h(delay) = Croats; h(stop aid) = IC.

Write h^* for the inverse function from the subsets of C to the subsets of D. Thus h^*G (where G is any group of characters, i.e., any subset of C) is the set of cards "held", or controlled, by members of G. For example, h^*\{Croats, IC\}= {delay, stop aid}.

A subset s of D (i.e., a subset of cards) is called a *selection* when thought of as a set of cards selected by some or all of the characters and a *future* when thought of as representing the particular future determined when that set of cards is selected by all the players. We write S for the set of all possible selections s – i.e., the set of subsets of D. As said, S may also be seen as representing the set of possible futures.

<u>The family</u> $> = (>_c \mid c \in C)$ is a family (or 'indexed set') of *preference relations* over the set S of futures. It contains one relation for each character c in the cast C. In Figure 1, characters' preference rankings for the futures shown are indicated by numbers. The statement '$s >_c t$' says that character c prefers selection s to selection t. For example, from Figure 1 we see that

$\{\text{connect, support, delay}\} >_{\text{Mayor}} \{\text{connect, support}\}$.

The negation of $>_c$ is written '$\leq_c$'. Thus '$s \geq_c t$' means that c is either indifferent between s and t or prefers s. In assuming that preference rankings can be indicated by assigning numbers in a card table we are assuming that the relation $\geq_c$ has the three properties of an ordinal ranking – i.e., *reflexivity* ($s \geq_c s$ for any selection s), *transitivity* (whenever we have $s \geq_c t \geq_c u$, we also have $s \geq_c u$) and *completeness* (for all s, t, we either have $s \geq_c t$ or $t \geq_c s$).

If G is a *group* of characters (meaning any subset of the cast C) we write '$s >_G t$' to mean 's is preferred to t by all members of G' and '$s \geq_G t$' to mean that every member of G either prefers s to t or is indifferent between them. For example, from Figure 1:

$\{\text{connect, support, delay}\} >_{\{\text{Mayor, IC}\}} \{\text{connect, support}\} >_{\{\text{Croats, IC}\}} \varnothing$.

From this definition, we always have s $>_{\varnothing}$ t (where $\varnothing$ is the empty group), since it is vacuously true that every member of $\varnothing$ prefers *s* to *t*, $\varnothing$ having no members. (The word 'group' is used in place of the game-theoretic term 'coalition' – despite the fact that 'group' has another, well-known mathematical usage – to avoid any connotation of common interests or collusion between the characters 'grouped' together.)

Note that a frame is fully specified by merely specifying a holding function *h* and a family > of preference relations, as we have done; the sets *C* and *D* are specified in specifying *h* (they are its domain and codomain) and *S* is simply the set of subsets of *D*. If we were to write out the specification (*h*, >) a little more in full, it would be

$$(h = (h\colon C \to D), > = (>_c \mid c \in C))$$

where each $>_c$ is a preference relation on the set of subsets of *D*. A frame need not be finite, although in applications it usually is. Characters and cards can both be infinite in number. If each hand is countably infinite, each character selects from a copy of the real line. If then we have an infinite cast, *S* is an infinite-dimensional Euclidean space.

This is how we model a drama-theoretic frame. Formally, the model is equivalent to a strategic-form game. This, it is assumed, is played as a non-co-operative game – i.e., the final future is determined by characters simultaneously and independently choosing sub-selections of cards from their hands, thus determining a total selection. However, these final choices are preceded by a period of communication, during which characters may re-define their perceived game. Re-definitions take place at successive 'moments of truth'.

Definition of a moment of truth

When players are in confrontational mode, a *moment of truth* is defined as a triple

$$(F, p, f),$$

where $F = (h, >)$ is a frame (the common reference frame), $p = (p^c \mid c \in C)$ is a family of *positions*, one for each character in the cast *C*, and *f* is the *fallback*. We have $f \in S$, since *f* is a particular future in the frame *F*, composed of the fallback strategies $f \cap h^*\{c\}$ that each character *c* declares it will follow if its position p^c is not accepted. But for each *c* we have $p^c \subseteq S$, since p^c is a 'general' position. Thus, p^c is a set of selections (i.e., a set of sets of cards).

This definition may also be used to for when players are in collaborative mode – i.e., when the intersection $\cap p$ of the positions p^c is non-empty –, provided that in this case we regard the fallback *f* as being a member of $\cap p$. (We shall later prove that under certain assumptions, this determines *f* uniquely, i.e. $\cap p$ is a singleton set).

The selections belonging to c's position p^c will be called c's *proposals*. In Figure 1, for example, we have p^{CROATS} = {{connect, support}, {connect, support, stop aid}}; the Croat proposals are {connect, support} and {connect, support, stop aid}.

A moment of truth within a finite frame, if regarded as a simple, uninterpreted mathematical object, can be represented by a diagram like those in Figure 3. Figure 3(a) represents the moment of truth in Figure 1, Figure 3(b) that in Figure 2. In these diagrams, cards are represented by rows, characters by horizontal divisions of rows. The playing of a card is represented in a column by a 1, not playing it by a 0, while a dash stands for '1 or 0'. If there are n characters, there are $n+2$ vertical divisions of the columns, representing the n positions $p^1, \ldots, p^n$, the fallback future f and the default future (although actually the default future plays no part in the definition of a moment of truth). Formally, we write $D = \{1,\ldots m\}$, where m is the number of rows, and $C = \{1,..., n\}$, where n is the number of row divisions. The specification of a moment of truth is completed by specifying a family $> = (>_i| \ i = 1,\ldots,n)$ of n preference relations over the set S of subsets of $\{1,\ldots m\}$; in Figure 3 preference rankings for the columns in the table are set out in n rows below the table, with best and worst rankings shown for columns containing more than one selection. A selection $s \in S$ is represented by a number in the set $\{0,1,\ldots,2^m–1\}$, written in binary notation as a column; for example, the set {1} is represented by 100…0, the set {1,3} by 10100…0, and so on.

1	1	1	0	0
1	1	1	0	0
1	0	–	0	0
0	–	0	1	0
2	3,5	2,3	4	1
5	1,2	1,5	4	3
2	1,3	1,2	5	4

(a)

0	–	0 1	0	0
0	–	0 1	0	0
–	0	1 0	0	0
–	0	1 0	0	0
0	–	0 0	1	0
–	0	0 0	0	0
1,2	2,4	1 3	4	2
3,5	1,3	4 1	2	3
2,5	1,4	2 1	4	3

(b)

Figure 3: Mathematical structures of the card tables in Figure 1 and Figure 2

Co-ordinated positions

A position (indeed, any set x of selections) will be called *co-ordinated* if it meets the following condition. Suppose the set D of all cards is partitioned into subsets D_i in any way whatever. Fix this partition, and call a selection s_i made from the subset D_i the *ith local selection.* Then a co-ordinated set x is such that *local compatibility implies global compatibility*; that is, if each *i*th local selection $s_i \in D_i$ is compatible with x (in the sense that x contains an s such that $s \cap D_i = s_i$) then the resultant global selection $\cup s_i$ will also be compatible with x (i.e., $\cup s_i \in x$). Thus the positions in Figure 3(a) are co-ordinated, whereas the third position in Figure 3(b) is not.

We now have:

Theorem 1: A non-empty set of selections $x \subseteq S$ is representable by a single column if and only if it is co-ordinated. Moreover, co-ordination and single-column representability are true of x if and only if

$$\exists t, u \in S:\ x = \{s|\ t \subseteq s \subseteq u\}. \tag{1}$$

Proof: We shall show that if a non-empty x obeys (1), then x is representable by a single column, which in turn implies that x is co-ordinated, which in turn implies that x obeys (1).

First, if a non-empty x obeys (1), then we can represent it by a single column with a 1 (for 'played') in rows belonging to t, a 0 (for 'not played') in rows not belonging to u, and a dash in all other rows. (Note that we must have $t \subseteq u$, or x would be empty.) Next, an x represented by a single column must be co-ordinated, since if the ith local selection s_i includes all cards in D_i that are played in the column, and none that are not, then the global selection $\cup s_i$ will include all in D that are played in the column and none that are not. Finally, if x is co-ordinated, it is co-ordinated when independent choices are made within a partition, every set in which contains just one card d in D. In this case a compatible local selection of a card d is such that if d is played, it must be played in *some* future belonging to x and if d is *not* played, it must be *not* played in some future belonging to x. If x is co-ordinated, we therefore have

$$[(\forall d \in s\ \exists t \in x{:}\ d \in t)\ \&\ (\forall d \notin s\ \exists\ t \in x{:}\ d \notin t)] \ \Rightarrow\ s \in x, \tag{2}$$

or, equivalently,

$$[(s \subseteq \cup x)\ \&\ (D - s \subseteq D - \cap x)] \Rightarrow\ s \in x,$$

or

$$\{s|\ \cap x \subseteq s \subseteq \cup x\} \subseteq x. \tag{3}$$

But the right hand side of (3) must be a subset of its left hand side, since an s belonging to x must contain any card contained in *every* s belonging to x, and can only contain cards contained in *some* s belonging to x. Hence (3) is equivalent to

$$x = \{s|\ \cap x \subseteq s \subseteq \cup x\}. \tag{4}$$

But (4) implies (1), since if (4) is true the t required to exist in (1) may be chosen as $\cap x$ and the u may be chosen as $\cup x$.

Can we, in general, assume co-ordinated positions? It would be good if we could, as the single-column representation is a clear and economical way to specify a position. We can justify such an assumption on the grounds that a co-ordinated position is a proposal for resolving some of the issues while leaving others (those marked by tildes or dashes) unresolved – and this is a natural way of defining a position. An uncoordinated position, by contrast, makes a number of disjoint alternative suggestions as to how certain issues should be resolved, rather than proposing a particular way. It is as if the department in Figure 2 were to say: "You two disagree. One wants X, one wants Y. My solution: do either X or Y!" That is to propose no solution at all.

Rather than say this, a character might propose a formula, procedure or set of criteria for choosing between X or Y. This way of moving toward a solution is normally followed in real life. But then this character's position should be modelled not by listing disjoint alternatives, but by defining cards that specify how it proposes the unresolved issues should be dealt with. For example, the uncoordinated position in

Figure 2 should be modelled as in Figure 4, where it is now represented as co-ordinated. Attaching clickable information to this model would allow us to specify the proposed joint co-ordination process in as much detail as the department intends. If, however, the department has no detailed proposals, we would argue that merely by proposing the uncoordinated position in Figure 2, it is implicitly proposing some co-ordination mechanism.

The procedure followed in Figure 4 for replacing an uncoordinated position by a co-ordinated one can be stated generally. First, however, we introduce a definition that allows us to look at co-ordination in another way.

Define a card d as being *independent in* a set x if one of the following is true: (i) d is played in every $s \in x$; (ii) d is not played in any $s \in x$; (iii) wherever d is played in some $s \in x$, the future $s-\{d\}$ also belongs to x, and wherever d is not played in some $s \in x$, the future $s\cup\{d\}$ also belongs to x. That is, the set of cards independent in x is:

$$\{d \mid d \in \cup x - \cap x \;\Rightarrow\; \forall s \in x:\; s\cup\{d\},\, s-\{d\} \in x\} \tag{5}$$

From this definition, it is clear that if a card d is independent in x, and x is transformed into another set x' by first deleting certain cards from, and then adding certain cards to, every selection in x, then d will be independent in the set x'. That is, we have

Theorem 2: If d is independent in x, then d is independent in the set

$$x' = \{(s-D')\cup D'' \mid s \in x\},$$

where D', D'' are any sets of cards.

Proof: From inspection of (5).

We now prove:

Theorem 3: A set x is co-ordinated if and only if every card is independent in x.

Proof: If a set x is co-ordinated, then using theorem 1, it is represented by a column in which every row $d \in \cap x$ contains a 1, every row $d \notin \cup x$ contains a 0, and every other row contains a dash. Now if $d \in \cap x$ or $d \notin \cup x$, d is independent within x since the antecedent in the defining condition of (5) is false. If, however, $d \in \cup x-\cap x$, then it is true by virtue of the single-column representation of x that given any s belonging to x, $s\cup\{d\}$ and $s-\{d\}$ also belong to x; hence again d is independent in x. On the other hand, if every card is independent in x, then suppose we build up, card by card, a selection s belonging to x. As we do so we must obviously include each $d \in \cap x$ and exclude each $d \notin \cup x$; as for any other d, from (5) we are free to choose whether or not to include it. Hence $x = \{s \mid \cap x \subseteq s \subseteq \cup x\}$, and from theorem 1, x is co-ordinated.

	A	B	DEPT	threat	default
SERVICE A	1,2	2,4	3	4	2
discontinue A's system	0	~	0	0	0
use B's system	0	~	0	0	0
join co-ordination study with B	0	0	1	0	0
SERVICE B	3,5	1,3	4	2	3
discontinue B's system	~	0	0	0	0
use A's system	~	0	0	0	0
join co-ordination study with A	0	0	1	0	0
DEFENCE DEPT	3,6	2,5	1	5	4
oppose A's system	0	~	0	1	0
oppose B's system	~	0	0	0	0

Comments on this card table

The position "we should use system A or B" modelled in Figure 2 as an uncoordinated position is replaced by a proposal to set up a process – a joint co-ordination study between the two services – for deciding which system to use. This is modelled by giving each service a card called "join co-ordination study". In this way, the uncoordinated position in Figure 2 is replaced by a co-ordinated one.

Figure 4: A co-ordinated way of modelling the uncoordinated position in Figure 2

Now consider the following general procedure (Procedure COORD) for modelling an uncoordinated position as co-ordinated. Suppose a character c takes an uncoordinated position p^c. Consider the set N of cards that are not independent in p^c; it is non-empty by theorem 2, which also tells us that it is

$$N = \{d \in \cup p^c - \cap p^c \mid \exists s \in p^c:\ s \cup \{d\} \notin p^c \ \ or\ \ s - \{d\} \notin p^c\}. \qquad (6)$$

(For p^3 in Figure 3(b), for example, the set N is {1,2,3,4}.)

Now consider the set $h[N]$ of characters that hold cards in N. (For p^3 in Figure 3(b), $h[N]$ = {1,2}). Add to the hand of each character b in $h[N]$ a card called "agree to co-ordinate within $\{s \cap N \mid s \in p^c\}$", or some wording to the same effect that is, in context, more appropriate. (For p^3 in Figure 3(b), the additional cards would be "agree to co-ordinate within {0011, 1100}". In Figure 4, however, the additional cards, instead of being called "agree to co-ordinate within {{discontinue B's system, use A's system}, {discontinue A's system, use B's system }}", are called "join co-ordination study with B (respectively, A)".)

Write D' for this set of additional cards. By making appropriate modifications to the family > of preference relations, we obtain a new model that has $D \cup D'$ as its deck. In this model, define c's position as the set formed by deleting from each selection $s \in p^c$ all the cards in N and adding all those in D'. Thus the redefined position, p'^c, is

$$p'^c = ((s-N) \cup D' \mid s \in p^c\}. \tag{7}$$

This redefined position evidently has the same effect as the old one, bearing in mind the descriptions (names) attached to the new cards. The effect is now achieved by requiring characters to co-ordinate their selection of cards within the set N of cards. Our argument is that character c, in putting forward the uncoordinated position p^c, must implicitly mean to put forward the position p'^c, since a position is meant to be implemented, and implementing p^c requires some form of co-ordination between the players in $h[N]$. Thus replacing p^c by p'^c merely makes explicit what was implicit in the position p^c.

It remains to prove:

Theorem 4: The redefined position arrived at by Procedure COORD is co-ordinated.

Proof: We shall show that each card in $D \cup D'$ is independent in p'^c; the theorem will then follow from theorem 2. First, consider each card in N; it is not in $\cup p'^c$, from the construction of p'^c; hence it is independent in p'^c. Similarly, each card in D' is in $\cap p'^c$, hence is also independent in p'^c. Finally, consider the cards in $D-N$. These are independent in p^c. From theorem 2, they are therefore independent in p'^c.

Lessons learned and some further problems and theorems

What have we learned from these theorems? First, that for any position, the property of being representable by a single column is the same as that of being co-ordinated: these two desirable properties go together. Second: it is permissible to assume that characters' positions are co-ordinated since, if they are not, there is a re-modelling procedure (procedure COORD) that re-defines them as such.

Three further problems remain. First, concerning the intersection of positions taken by a group of characters; if non-empty, we would like to take this as the joint position of the group. But to do so, we must be sure that an intersection of co-ordinated positions is co-ordinated. Second: suppose no character takes a position in relation to a particular card – i.e., it is neither demanded nor excluded by any position. For simplicity, in order to have a simple model of the moment of truth, we would like to delete this card from the deck, since it is not an issue between the characters at this moment (though it may need to be in the deck of the large, complex models developed by staff to look at the implications of simple, general positions). What are the effects of excluding it? Third: concerning a group of characters whose positions are incompatible; is there (assuming co-ordinated positions) necessarily a pair within the group that disagree over a specific card? If not, it will be hard to find a particular issue to motivate change.

To decide these questions, we first have:

Theorem 5: The intersection of a family of co-ordinated sets is co-ordinated.

Proof: Consider the intersection $\cap X = \cap(x_i \mid i \in I)$ of a family X of co-ordinated sets. From Theorem 1, each x_i in X is such that $x_i = \{s \mid t^i \subseteq s \subseteq u^i\}$ for some t^i, u^i. Form the union t of all the t^i and the intersection u of all the u^i. Then each future in $\cap X$ is a superset of u, being a superset of every u^i, and a subset of t, being a subset of every t^i. Moreover, any s that is both a superset of u and a subset of t belongs to $\cap X$, since it satisfies the membership conditions for every set in X. Thus $\cap X = \{s \mid t \subseteq s \subseteq u\}$. But this is co-ordinated by Theorem 1.

To illustrate: in Figure 1, the intersection of the Croat and IC positions is the co-ordinated set ({connect, support}). The intersection of the Serb and Croat positions is the empty set. This too is co-ordinated.

In light of this theorem, we define the *joint position* p^G of a group $G \subseteq C$ as the intersection of their individual positions:

$$p^G = \cap(p^c \mid c \in G). \tag{8}$$

This means that in the case of a singleton group $\{c\}$, we have $p^{\{c\}} = p^c$, but for any non-empty, non-singular group G, p^G may be empty. However, $p^{\varnothing} = S$.

A group will be called *compatible* if its joint position is non-empty.

Next, define a deck D as *minimal* if it contains no card that is not either demanded or excluded by some character's position. That is, if

$$\forall d \in D: (\exists c: d \in \cap p^c \;\; or \;\; \exists c: d \notin \cup p^c). \tag{9}$$

The deck in all our examples is minimal. Our suggestion is that decks should always be minimal to give simple models of moments of truth. We have:

Theorem 6: If the deck is minimal and positions are co-ordinated, the joint position p^C of the whole cast is either empty or singular — i.e., contains just one future.

Proof: Suppose $p^C = \cap p$ is non-empty. Since D is minimal with respect to p, every card is either demanded by some character's co-ordinated position — in which case it is in every future belonging to $\cap p$ — or excluded by some character's position — in which case it is in none of the futures belonging to $\cap p$. No card can be both (i.e., demanded by some character's position and excluded by another's), or $\cap p$ would be empty. Hence, we have determined, concerning any card, whether or not that card is in a future belonging to $\cap p$ — i.e., we have determined a unique future belonging to $\cap p$.

Finally, we can resolve the problem of finding, in an incompatible group, a specific incompatibility between two characters – the problem that arises in Figure 2. We have:

Theorem 7: If positions are co-ordinated, an incompatible group contains a pair that disagrees over the playing of a particular card – i.e., if G is incompatible

$$\exists\, a, b \in G, d \in D: \;\; d \in \cap p^a - \cup p^b.$$

Proof: Suppose G is incompatible – i.e., $\cap(p^c | c \in G) = \varnothing$. If now each p^c is co-ordinated, we have

$$\cap(p^c | c \in G) = \{s \mid t \subseteq s \subseteq u\},$$

where t is the union of the sets $\cap p^c$ $(c \in G)$ and u is the intersection of the sets $\cup p^c$ $(c \in G)$. Therefore $\cap(p^c | c \in G) = \varnothing$ implies that t is not a subset of u; that is, that there exists a card in t that does not belong to u, which is to say,

$$\exists\, d: (\exists\, a \in G: d \in \cap p^a \text{ and } \exists\, b \in G: d \notin \cup p^b).$$

This states that for some card d there exist $a, b \in G$ such that $d \in \cap p^a - \cup p^b$. Equivalently, that there exist characters a and b in G such that a demands a card excluded by b.

The basic theorems in the case of general positions

We now have to re-prove the basic theorems of drama theory for the case of general, co-ordinated positions. It has already been proved[4] that at a moment of truth, characters with singular positions *either* find that they have reached a 'satisfactory resolution' of their joint decision problem (in that they agree on a strict, strong equilibrium) *or* they face dilemmas. We have to prove something like this for general positions.

To begin with, we will re-define the above terms for the case of general positions – having already re-defined 'moment of truth'. First, what is meant by 'satisfactory resolution'? This will again consist of agreement on a 'strict, strong equilibrium' – but this term itself will be re-defined.

Strict, strong equilibria

First, we introduce a useful notation. Given a selection s and a group G, write s_G for the sub-selection of cards selected, within s, by members of G. That is, write, for any s and G:

$$s_G = s \cap h^*G.$$

We now define a set x of selections as a *strict, strong equilibrium* if it is co-ordinated and no group G has a 'potential improvement' from it — where the set $\mathrm{Imp}_G(x)$ of *potential improvements for G from x* is defined by

$$\mathrm{Imp}_G(x) = \{s \in S - x \mid \exists t \in x:\ s_{-G} = t_{-G};\ s \geq_G t\}. \qquad (10)$$

(Note here that we write '$-G$' for the group $C-G$ – i.e., the set of characters in the cast C that are not in the group G)

Thus, a potential improvement from x is a selection outside x to which a group G can move 'unilaterally' (i.e., given that those not in G don't change their selections) from a selection inside x without any member of G losing utility from the move.

Accordingly, a strict, strong equilibrium will be a co-ordinated set of selections such that no group G can move, unilaterally, from a selection in it to a selection outside it without loss to some member. Hence, if all characters in the cast C agree to implement a non-empty, strict, strong equilibrium x, each individual or group within C must mean and is able to do so (i.e., can be trusted not to break the agreement), since the following points hold:

- x, being co-ordinated, has the characteristic that if and only if each character c chooses a selection $s_{\{c\}}$ compatible with x (i.e., such that $s_{\{c\}} = t_{\{c\}}$ for some $t \in x$) then the total selection s belongs to x.

- Consequently any group G that plans to break the agreement while expecting those not in G to keep it must expect to move from a point in x to a point outside it, which means that at least one of them must expect to lose.

Because of these characteristics, a non-empty, strict, strong equilibrium *that is the joint position of the whole cast* is indeed a 'satisfactory resolution', in that, first, all characters accept every future in it and, secondly, no group can be suspected of intending to defect from it. Finally, if the deck is minimal, it is singular (Theorem 6). It therefore represents full and complete dramatic resolution. (Note: such resolution is 'satisfactory' in a dramatic and technical sense; it does not necessarily make characters happy. It may be more like a tragic ending than a happy one; either kind of ending gives complete resolution, the tragic one through hopes being shattered, the happy one through hopes being realised.)

In order for the intersection of a group's positions to be a strict, strong equilibrium, it is sufficient (though not necessary) for each individual's position to be itself a strict, strong equilibrium. That is, we have

Theorem 8: Any intersection of string, strong equilibria is a strict, strong equilibrium.

Proof: First, such an intersection is co-ordinated by Theorem 5. Next, a potential improvement from an intersection of SSEs (strict, strong equilibria) would, by the definition of a potential improvement, be a point s not belonging to at least one of those SSEs such that, for some group G

- s is at least as good for all members of G as a point t belonging to every SSE,

- s is reachable by G from t (i.e., G can move to it from t).

Such an s would be a potential improvement for G from the SSE to which it does not belong; but there are no potential improvements from an SSE, proving the theorem.

Dilemmas faced at a moment of truth

Our next step is to define the 'dilemmas' that characters holding co-ordinated positions may face at a moment of truth. These are the 'change factors' that, by generating emotion followed by rationalisation, lead characters to re-define the frame; they tend to do so in such a way as to eliminate the dilemmas. Having defined them, we will prove that *a cast of characters that face no dilemmas are compatible, their joint position being a strict, strong equilibrium*. They have, therefore, satisfactorily resolved their problem. We will prove a number of other theorems concerning the dilemmas, but this – which asserts that dilemma-elimination leads to a satisfactory solution – is central.

To define the dilemmas we specify, for each character, six sets, called ‘gradients’, the non-emptiness of each of which puts that character in a specific dilemma. Though inspired by the gradient sets defined in reference 10, these are defined somewhat differently. As well as being simpler, the new definitions reflect important distinctions that became clear when the concepts were generalised to the case of general positions.

One distinction is that between dilemmas of *communication* and of *implementation*. The former include the inducement dilemma (“I’m under pressure to give in to you, because your position is no worse for me than the threatened future”), the deterrence dilemma (“You’re under no pressure to give in to me, preferring the threatened future to my position”), and the positioning dilemma (“I find it hard to argue with you because I prefer your position to my own, but am forced to reject it because it’s not realistic”). These dilemmas put rational-emotional pressure on a character to change its position in favour of someone else’s during the period of pre-play communication. By contrast, the dilemmas of *implementation* (the threat, co-operation and trust dilemmas) put pressure on a character after communications have ceased, when it must decide whether to carry out the commitments (threats or promises) it has made, or whether to believe others’ commitments, given a preference for ‘reneging’ on them. This pressure is, of course, foreseen during the period of communication, and therefore puts pressure on characters during the communication period also; but it is a different kind of pressure.

The difference between the two kinds of dilemmas appears as a difference in the way we should define the ‘gradient’ set, the non-emptiness of which shows the existence of each dilemma. Gradients specifying *c*’s dilemmas of communication are best defined as sets of *characters* (in relation to whom *c* has that dilemma), whereas gradients specifying *c*’s dilemmas of implementation need to be defined as sets of *potential improvements* that make *c*’s position or fallback incredible. These definitions best model the underlying psychological realities.

A different kind of distinction is between dilemmas of *confrontation* and *collaboration*. The first are relevant only when characters are in confrontation mode – i.e., when their positions are incompatible. Dilemmas of collaboration, by contrast, are important primarily in collaboration mode; they are relevant in confrontation mode only insofar as confronting characters look forward to the fact that, if a certain position is accepted, they will have to deal with it in collaboration mode.

The sets whose non-emptiness yields a dilemma are called ‘gradients’ in analogy with the gradient of a dynamical system, which is an object that represents the system’s tendency to change in various directions; ‘dilemmas’, we assert, motivate characters to attempt to change (redefine) their moment of truth so as to eliminate the dilemmas. We discuss the six dilemmas and their gradients one by one, beginning with the dilemmas of incompatibility.

Four dilemmas of confrontation

1. The *threat dilemma* is a dilemma of implementation. A character facing this dilemma might not implement the fallback strategy (or ‘threat’) it is committed to. The *threat gradient* for

character c contains all potential improvements for c from the fallback future – i.e., all potential improvements for the one-person group containing c. Thus it is

$$\begin{aligned} \text{ThGrad}(c) &= \text{Imp}_{\{c\}}(\{f\}) \\ &= \{s \in S - \{f\} \mid s_{C\text{-}\{c\}} = f_{C\text{-}\{c\}}\,;\ s \geq_c f\}. \end{aligned} \tag{11}$$

If this set is non-empty, it would be rational for c, if negotiations break down and it expects the others to implement $f_{C\text{-}\{c\}}$, not to implement its fallback strategy $f_{\{c\}}$ but to implement a different strategy $s_{\{c\}}$. We say that c's *fallback is incredible* and that c is *blustering.*

We also say that the threatened future f is incredible; that is, f will be called *incredible* if

$$\cup(\text{ThGrad}(c) \mid c \in C) = \cup(\text{Imp}_{\{c\}}(\{f\}) \mid c \in C) \neq \varnothing. \tag{12}$$

In words, f is incredible if the union of all threat gradients, which is the same as the set of *individual* potential improvements from f, is non-empty. Using game-theoretic terminology, we would say that f is incredible unless it is a *strict equilibrium*.

This definition of the threat dilemma differs from that in reference 10 in two ways.

- First, only *individual* potential improvements from f are regarded as members of the threat gradient. The argument for this is that, in forming a card-table model, a decision has to be made as to how to aggregate organisations to form characters (e.g., should the various sub-characters comprising the IC in Figure 1 be shown separately, or lumped together as in the table?) The best answer is to lump sub-characters together as a single character when it can be assumed that, should negotiations break down, they will consult together in deciding what to do next. It follows that only individual potential improvements should be considered in the event of a breakdown.

- Second, potential improvements from f that lead to one or another character's position are included in the threat gradient – whereas previously an attempt was made, by excluding them, to define gradients as mutually exclusive. There seems, however, to be no good reason for this: it seems reasonable that the same improvement should pose two different kinds of dilemma. If A prefers B's position to the threatened future, there seems no reason to hide the fact that this may give A both an inducement dilemma (pressuring A to 'give in' and accept B's position during communications) and a threat dilemma (tempting A to move unilaterally from the threatened future to B's position, should negotiations fail).

Example: In Figure 3(a), 3 has a has threat dilemma, since ThGrad(3) contains the selection $\varnothing$ (0000, in the fifth column). Why? Because $\varnothing$ is ranked 4 in 3's preferences, whereas f (0001, in the fourth column) has rank 5. Hence we have $\varnothing >_3 f$, while $\varnothing_{\{1,2\}} = f_{\{1,2\}} = \varnothing$. Thus, 3 is blustering. Correspondingly, in Figure 1, the IC's threat gradient contains the column **default,** preferred by the IC to the threatened future (column **threat).** The IC is blustering.

2. The *deterrence* dilemma is a dilemma of communication. The *deterrence gradient* for c contains all characters incompatible with c that prefer the threatened future to any of c's proposals. Thus it is

$$\mathrm{DeGrad}(c) = \{b \in C \mid p^c \cap p^b = \emptyset;\ \forall s \in p^c: f >_b s\}. \tag{13}$$

If c has a deterrence dilemma with respect to b (i.e., if $b \in \mathrm{DeGrad}(c)$), c is said to be *unrealistic* toward b; the threatened future places no pressure on b to accept c's position. Hence if c is to be taken seriously it must change either its position, the preferences of those in its gradient, or the threatened future.

<u>Example:</u> In Figure 3(a), 2 belongs to 1's deterrence gradient, since 1's position is ranked 5 by 2, whereas f is ranked 4; hence $\forall s \in p^1: f >_5 s$. Correspondingly, in Figure 1, the Mayor has a deterrence dilemma, and is unrealistic.

The definition differs from that in reference 10 in its focus on the preference of b (a single character) for the fallback f, compared to c's position. In reference 10, the condition for a dilemma is more complicated (a deterrence dilemma is caused when the characters incompatible with c have a joint strategy that – starting from f – makes them all better off than they are at c's position.). The simpler definition seems to capture rather better the idea of a deterrence dilemma and works better in applications; it is the case, moreover, that if a character is realistic toward another in the more complex sense it is so in the simpler sense.

3. The *inducement* dilemma is again a dilemma of communication. The *inducement gradient* for c contains all characters incompatible with c that offer proposals as good for c as the threatened future f. Thus it is

$$\mathrm{InGrad}(c) = \{b \in C \mid p^c \cap p^b = \emptyset;\ \exists s \in p^b: s \geq_c f\}. \tag{14}$$

If this were non-empty, it would be rational for c to accept one of b's proposals, rather than reject it and suffer f ; yet c is insisting it won't do that. We say that c is *obdurate* toward b.

We now have:

Theorem 9: If b and c are incompatible, b is obdurate toward c if and only if c is realistic toward b. (Symbolically, $b \in \mathrm{InGrad}(c) \Leftrightarrow c \notin \mathrm{DeGrad}(b)$).

Proof: The first defining condition in definitions (13) and (15) is that c and b should be incompatible. Now exchange the letters c and b in (13). The second defining condition is transformed into the negation of that in (15).

This theorem allows us to use the following terminology. If $b \in \mathrm{InGrad}(c)$ *or* $c \notin \mathrm{DeGrad}(b)$, we say that b (and b's position p^b) *induces c*. If $b \notin \mathrm{InGrad}(c)$ *or* $c \in \mathrm{DeGrad}(b)$, we say that c is *uninduced by b* (and by b's position p^b).

Theorem 9 tells us that a single fact – i.e., the truth of a single statement ($\exists s \in p^b: s \geq_c f$) – has two different effects. It creates a dilemma for c (putting c under pressure to give in to a proposal of b's) and eliminates a dilemma for b (b no longer has the problem that c is under no pressure to accept any of its proposals).

We now have

Theorem 10: If positions are co-ordinated, a cast is realistic and non-obdurate if and only if it is compatible.

Proof: A compatible cast is realistic and non-obdurate by definition – see (13) and (14). On the other hand, consider a co-ordinated cast that is realistic and non-obdurate. We show that it cannot be incompatible. For suppose (if possible) that it is. From Theorem 7, it must contain an incompatible pair. Each member of this pair is, by assumption, realistic toward the other; hence from Theorem 9, each is obdurate, contradicting our initial assumption about the cast.

4. The *positioning* dilemma is another dilemma of communication. The *positioning gradient* for c contains all characters whose positions contain proposals better for c than some future belonging to its own position p^c. Thus it is

$$\mathrm{PoGrad}(c) = \{b \in C \mid p^c \cap p^b = \varnothing;\ \exists s \in p^b, t \in p^c: s >_c t\}. \tag{15}$$

If c has a positioning dilemma in relation to b, c rejects b's position, yet prefers a proposal of b's to some proposal of its own. This makes it hard for c to sustain its rejection, or to argue that b should give in and accept c's own position. We say that c is *inconsistent toward b.* We have:

Theorem 11: Compatible characters are consistent toward each other.

Proof: Immediate from (15).

The positioning dilemma typically occurs when a character has adopted a position incompatible with another's, not because it prefers (all) its own proposals to the other's, but because it considers the other's position to be irremediably unrealistic in relation to a third party – i.e., the other's position faces a deterrence dilemma it considers to be insurmountable. To overcome a positioning dilemma, the character can change its preferences, abandon or modify its own position or persuade the other to abandon or modify its position.

Two dilemmas of collaboration

The remaining two dilemmas, unlike the previous four, are relevant when the cast is compatible.

5. The *co-operation dilemma,* like the threat dilemma, is a dilemma of implementation. A character faces this dilemma when others might not be able to trust it to implement its part of it own position, should they agree to it. The *co-operation gradient* for character c contains all potential improvements from c's position for groups containing c. Thus it is

$$\begin{aligned} \mathrm{CoGrad}(c) &= \cup\{\mathrm{Imp}_G(p^c) \mid c \in G \subseteq C\} \\ &= \{s \notin p^c \mid \exists t \in p^c: \exists G \ni c: s_{-G} = t_{-G};\ s \geq_G t\}. \end{aligned} \tag{16}$$

If this is empty, c is said to be *trustworthy*. Otherwise, c is untrustworthy in relation to at least one of its proposals, t, in that c is open to persuasion by the group G (which may contain just c itself) to 'defect' from t to a point outside p^c.

Note that the co-operation gradient (unlike the threat gradient) may include *group* potential improvements that no individual character can carry out. This is because a group of characters that take the same position may well consider pursuing a joint strategy (such as the strategy of 'defecting' from that position), since they have already agreed in regard to the positions they are taking. They may communicate and co-operate with each other, unlike characters who, after the resolution process has broken down, have implied that they will go their separate ways.

<u>Example:</u> The selection $\varnothing$ in Figure 3(a) (i.e., the fifth column, 0000) belongs to CoGrad(1), since it belongs to $\mathrm{Imp}_{\{1,2\}}(p^1)$. It does so because 1110 (the sole selection belonging to p^1) is ranked second by 1, whereas 0000 is ranked first, and is ranked fifth by 2, whereas 0000 is ranked third. At the same time, $1110_3 = 0000_3 = 0000$.

Correspondingly, the Serb mayor in Figure 1 has a co-operation dilemma; he is untrustworthy, since if the others agreed to his position, he would be tempted to collude with the Croats in moving to the default column, which is preferred by both him and them. He and the Croats could make this 'move' by simply not doing what the agreement requires – i.e., he would not connect utilities nor support returns while they would not delay returns. Meanwhile, the IC would continue to give aid. Why is this a dilemma for the mayor? Because the existence of this improvement from his position means that the IC has reason not to trust him, hence it makes it hard for him to convince them to accept his position.

If we looked at the whole of the mayor's preference relation $>_{\mathrm{Mayor}}$, we would undoubtedly find that he has other, individual improvements from his position – i.e., he would prefer to not connect utilities and/or not support returns whether or not the Croats agree to delay returns. These are additional reasons not to trust him; in particular, they are reasons for the Croats not to trust him. (These other improvements are not shown in Figure 1 because a co-operation dilemma does not necessarily show up on a card table that shows just n+2 futures – n characters' positions, the fallback future and the default future –, since it involves futures (temptations to 'defect' from a position) that may not be among these n+1. In this it is like the other two dilemmas of implementation, which also represent temptations to 'defect' from either a position or the fallback future.)

6. The *trust dilemma* is again a dilemma of implementation. A character faces this dilemma when it might not be able to trust others to implement their part of its position, even if they agreed to it. The *trust gradient* for c is

$$\mathrm{TrGrad}(c) = \cup(\mathrm{Imp}_G(p^c) \mid c \notin G) \tag{17}$$

$$= \{s \notin p^c \mid \exists t \in p^c \colon \exists G \colon c \notin G;\ s_{-G} = t_{-G};\ s \geq_G t\}.$$

If this set is empty, c is said to be *trusting*. Otherwise, it has to be *mistrusting* in relation to at least one of its own proposals, t, as a group G not containing c would be tempted to defect from t to a selection not in p^c.

Example: In Figure 3(a), 0000 belongs to both TrGrad(2) and TrGrad(3), since

$$0000 >_1 1100;\ \ 1100 \in p^1, p^2;\ \ 0000_{\{1,2\}} = 1100_{\{1,2\}}.$$

Correspondingly, in Figure 1 the Croats and the IC must be mistrusting, since the Serb mayor could not be trusted to implement the future{connect, support} (which belongs to both their positions) even if he agreed to it, but would be likely instead to continue to implement $\varnothing$ – the default future.

We now have

Theorem 12: A character is trusting and trustworthy if and only if its position is a strict, strong equilibrium.

Proof: A character is trusting and trustworthy just when there are no potential improvements from its position for any group (containing or not containing itself). This is the condition for its position to be a strict, strong equilibrium.

Theorem 13: The joint position of a group that is trusting and trustworthy is a strict, strong equilibrium.

Proof: From Theorems 8 and 11.

Theorem 14 (theorem of the final state): If positions are co-ordinated, a cast that is realistic, non-obdurate, trusting and trustworthy is compatible at a co-ordinated, strict, strong equilibrium. Furthermore,

- no character has a positioning dilemma;
- if the deck is minimal, the cast's joint position is singular.

Proof: Assume co-ordinated positions. From Theorem 10, a realistic, non-obdurate cast is compatible. From Theorem 13, it is compatible at a strict, strong equilibrium which, from Theorem 8, is co-ordinated. The bulleted points follow from Theorems 11 and 6.

This theorem tells us that if the dilemmas of deterrence, inducement, co-operation and trust are eliminated, then the characters' joint decision problem is satisfactorily resolved. Thus it re-proves, in the case of co-ordinated general positions, the basic theorems proved in reference 4 for the case of singular positions.

Implications for command and control of PO

As said, this research aims to establish the mathematical foundations for a decision support/C2 system for PO. The discoveries made this year have helped in the following ways.

General positions

The fact that often it is realistic to model a situation in terms of general positions is illustrated in Figure 1, where for different reasons both the Croat ethnic leaders and the International Community clearly took general positions. This is one reason why the system must be able to handle general positions.

In addition, this is necessary for staff work. We have noted that the proposed system links commanders at each level, each of whom will, in general, be making decisions in coalition with civilian agencies belonging to the IC. Thus the 'player' that the system supports is, at each level, an IC coalition. The commander's staff operates, at each decision node, to screen military intelligence before inputting it into this system, which is a system of models containing information shared by all members of the IC coalition. See reference 2 for a more detailed description of the system.

In order to support co-ordination between command levels, the commander's staff need to be able to take a simplified model such as that in Figure 1 and add many cards to each column to model the detailed implementation of the future represented by that column. They need to do this for two reasons: (a) to estimate how these cards, representing details, are likely to be played as that future is implemented; (b) to set out cards that are relevant to that future but are controlled by subordinate IC coalitions, or IC coalitions on the same level, or by the NCPs that they interact with. Since the playing, or not, of such cards is not yet decided, their insertion makes the column represent a 'general' future – i.e., a set of possible futures – rather than a specific one.

Insertion of such cards into a column representing our position is a way of delegating missions to subordinate IC coalitions. Writing tildes against these cards, or requiring subordinate coalitions to respond by editing and amending the model of their confrontation that we have suggested to them, is a way of implementing the principle of Mission Command. It ensures that details of implementation are left for subordinates to decide, and allows them to suggest modifications. At the same time, they are able clearly and succinctly to inform their superiors of the modifications they wish to make, enabling their superiors to support their mission as and when necessary, so that mission support is reciprocal between command levels.

Thus the ability to use general positions in the system is necessary, not only in order to realistically model particular situations, but also in order to model the proper relations between different command levels.

Co-ordinated general positions

The theorems we have proved about co-ordinated positions show, in the first place, why it is useful to be able to model general positions as co-ordinated. They also give reasons to suppose that positions encountered in application to PO will be co-ordinated. Finally, they show how, if we have formulated a model in which a position is represented as uncoordinated, we can (using the procedure COORD), re-model it as co-ordinated.

How would this be used in practice? How would we proceed if a position is found (in the form in which it is being presented) to be uncoordinated? The position might be our own position or that of an NCP or other party to the interaction.

(a) If *our own* position is uncoordinated, we ought to think how to make it co-ordinated, since this is required to make it implementable. The procedure COORD gives us a broad definition of co-ordination, which can be refined and made more specific by thinking of concrete procedures.

(b) If *another's* position is uncoordinated, we need to think how it would in fact be implemented, if genuinely accepted by the characters – and how others would expect it to be implemented, etc. This will influence our and their estimation of its consequences and preferences for it. But thinking through the implementation of a position requires thinking how it is likely to be co-ordinated. Hence, the procedure COORD will again be helpful in arriving at these judgements.

Use of dilemma analysis

Reference 2 sets out the method by which a commander whose Main Effort lies in confronting NCPs can use dilemma analysis to formulate an effective strategy in coalition with civilian members of the International Community. Dilemma analysis has a mathematical basis that allows us, once a situation has been specified in card-table terms, to employ automatic generation of the dilemmas faced by all parties, thereby helping the commander and the IC coalition as they formulate and implement their strategy.

Once 'general' positions are admitted, this mathematical basis is in danger of disappearing insofar as it depends upon single-future positions. Hence, one of the main tasks, achieved in this paper, has been to re-prove the theorems for the more general case. This again makes possible automatic generation of dilemmas, given a card-table model of the situation. Thus, it gives a sound mathematical basis to the proposed decision support/C2 system for PO.

---

1 Howard N (1999) *Confrontation Analysis: how to win Operations Other Than War.* CCRP Publications, OASD(C3I), Pentagon, Washington DC. Available from www.dodccrp.org.

2 ISCO Ltd (2000), *A Decision Support System for Peace Operations Commanders: a mission capability package for a system based on confrontation analysis,* Report to C4ISR Cooperative Research Program, OASD(C3I), Pentagon, Washington DC. Available from www.dodccrp.org.

3 Howard, N (1971). *Paradoxes of Rationality: Theory of Metagames and Political Behavior.* MIT Press.

4 Howard N (1998). n-person 'soft' games. *Jour of Opl Res Soc.* **49** 144-150