CAHIER 21-2000

DEBATES AND DECISIONS : ON A RATIONALE OF ARGUMENTATION RULES

Jacob GLAZER and
Ariel RUBINSTEIN

Centre de recherche
et développement en économique

C.P. 6128, succursale Centre-ville
Montréal QC H3C 3J7

Téléphone : (514) 343-6557
Télécopieur : (514) 343-5831
crde@crde.umontreal.ca
http://www.crde.umontreal.ca/

Université de Montréal

CAHIER 21-2000

DEBATES AND DECISIONS : ON A RATIONALE OF ARGUMENTATION RULES*

Jacob GLAZER[1] and Ariel RUBINSTEIN[2]

1 The Faculty of Management, Tel Aviv University

2 The School of Economics, Tel Aviv University and Department of Economics, Princeton University

December 2000

* Most of this author's research was conducted while he was a visiting scholar at the Russell Sage Foundation, New York, during the 1996-1997 academic year. This paper was presented at a C.R.D.E. Special Lecture held at the Université de Montréal on November 17, 2000.

RÉSUMÉ

Nous voyons un débat comme un mécanisme par lequel un preneur de décision non informé (l'écouteur) extrait l'information de deux orateurs informés, qui maintiennent des positions contradictoires concernant la juste décision. Au cours du débat, les orateurs développent des arguments et, se basant sur ces arguments, l'écouteur arrive à une conclusion. En utilisant un simple exemple, nous examinons le problème de design de mécanisme de construire des règles de débat qui maximisent la probabilité que l'écouteur arrive à la juste conclusion, sous des contraintes sur la forme et la longueur du débat. Les règles de débat optimales ont la propriété que la conclusion atteinte par l'écouteur n'est pas nécessairement la même que la conclusion qu'il aurait atteinte s'il avait interprété l'information qui lui a été révélée durant le débat, littéralement. Il est démontré que l'interprétation optimale d'un contre argument dépend de l'argument qu'il contre argumente. Nous relions également notre discussion à la littérature des pragmatiques.

Mots clés : débats, pragmatique, théorie des jeux, design de mécanisme

ABSTRACT

We view a debate as a mechanism by which a uniformed decision maker (the listener) extracts information from two informed debaters, who hold contradicting positions regarding the right decision. During the debate, the debaters raise arguments and, based on these arguments, the listener reaches a conclusion. Using a simple example, we investigate the mechanism design problem of constructing rules of debate that maximize the probability that the listener reaches the right conclusion, subject to constraints on the form and length of the debate. The optimal debate rules have the property that the conclusion drawn by the listener is not necessarily the same as the conclusion he would have drawn, had he interpreted the information, revealed to him during the debate, literally. The optimal interpretation of a counterargument is shown to be dependent on the argument it counterargues. We also link our discussion with the pragmatics literature.

Key words : debates, pragmatics, game theory, mechanism design

1. Introduction

This paper is a part of our long-term research agenda for studying different aspects of debates using game theoretic tools. Debates are common phenomena in our daily life. In a debate, two or more parties (the debaters), who disagree regarding some issue, raise arguments to support their positions or to rebuff the other party's arguments. Sometimes the purpose of the debaters is to argue just for the sake of arguing, and sometimes their aim is to try to convince the other party to change his position. In this paper, however, the purpose of each debater is to persuade a third party (the listener) to support his position. Note that a debate is different from bargaining and war, which are also mechanisms for conflict resolutions, in that the outcome of those mechanisms heavily depend on the rivals' power. A debate is different from a conversation, which is also a mechanism in which interested parties make arguments, in that in a conversation, there is a common interest among the parties.

We view a debate as a mechanism by which an uninformed decision-maker (the listener) extracts information from two informed parties (the debaters). The debaters hold contradicting positions about the decision that should be made. The right conclusion depends on several outcomes. During the debate the debaters raise arguments to support their respective positions and on the basis of these arguments, the listener reaches a conclusion regarding the right decision. When we say that a debater raises the argument x, we mean that he reveals that aspect x supports his position. When the other debater responds to an argument x with an argument y, we refer to argument y as a counterargument. The realizations of the aspects are assumed to be independent. All aspects are assumed to be equally weighted, in the sense that all of them have the same value of information regarding the right decision.

In this paper, we address only the issue of the relative strength of arguments and counterarguments. Under the above assumptions one may expect the optimal debate conclusion to be a function only of the number of arguments made by each party. Our intuitions supported by some experimental evidence is that this is not correct: after one argument has been made by one party, the subjects, in the role of the other party, may find the seemingly equal counterarguments unequally persuasive. Normatively, we investigate the optimal debate rules within a simple example. We show, that the optimal debate rules have the property that the strength of a counterargument may depends on the argument it is countering, even when there is no informational dependency between the two arguments. In particular, we show the invalidity of the following principle, regarding the dependency of the outcome of a debate on two the argument raised by one debater and the counterargument raised by the other debater:

> **The Debate Consistency (DC) Principle:** It is impossible that"x wins the debate" if y is brought up as a counterargument to x, but "y wins the debate" if x is brought up as a counterargument to y.

We show that this principle is <u>not</u> necessarily a property of debate rules optimally designed to extract information from the debaters.

Let us emphasize that we do not intend to provide a general theory of debates. Our only aim is to point out that the logic of the optimal design of debating rules is subtle and contains some features which are not intuitive.

2. Motivating Examples

Before we proceed to the model we will discuss two examples to demonstrate our intuition regarding the subtle nature of what people perceive as a good counterargument.

<u>Question 1:</u> You are participating in a public debate about the level of education in the world's capitals. You are trying to convince the audience that in most capital cities, the level of education has risen recently. Someone is challenging you, bringing up indisputable evidence showing that the level of education in **Bangkok** has deteriorated. Now it is your turn to respond. You have similar, indisputable evidence to show that the level of education in **Mexico City, Manila, Cairo** and **Brussels** has gone up. However,

because of time constraints, you can argue and present evidence only about one of the four cities mentioned above. Which city would you choose for making the strongest counterargument against **Bangkok**?

Our intuition is that in this debate scenario a good counterargument will be "close" to the argument it is countering, even though the geographical proximity is irrelevant to the substance of the arguments. Therefore we expect that most people will find the evidence about Manila to be a better counterargument than that about Mexico City, Cairo or Brussels for rebuffing the evidence from Bangkok.

To support our intuition we presented the scenarios (described in Hebrew) to groups of subjects, students at Tel Aviv University. More experimental research is needed to establish our intuition as experimental observations, but we think that the results are of interest. A group of 38 subjects received the question and 50% of them agreed with our intuition and answered Manila. Each of the other alternatives was chosen by 21% of the students or less. A second group of 62 subjects was presented with Question 1 with the modification that Bangkok was replaced by Amsterdam. Here 78% of the subjects found Brussels the most persuasive counterargument against Amsterdam.

To prevent a possible claim that the above phenomenon is confined to cases where the subjects have some implicit beliefs about correlation between the arguments, we presented another group of subjects with the following question:

<u>Question 2:</u> Two TV channels provide fixed program schedules for the five weekdays. You and a friend are debating which is the better channel before a third party. Your opponent argues in favor of channel A while you argue in favor of channel B. Both of you have access to the same five reports prepared by a highly respected expert, each of which refers to a different day of the week and recommends the better channel for that day. Your opponent begins the debate by quoting **Tuesday**'s report, which found channel

A's programs to be better. The listener then stops him, asking you to reply. Both **Wednesday**'s and **Thursday**'s reports are in your favor, namely, they found Channel B to be superior. However, you have time to present only one of these expert reports as a counterargument, after which the third party will make up his mind. Will you choose Wednesday or Thursday as a better counterargument to Tuesday?

About 69% of the 58 subjects found Wednesday, rather than Thursday, a better counterargument to Tuesday.

A puzzling element emerges from considering these examples. If two arguments contain the same quality of information, why is one of them considered to be a stronger counterargument than the other? The fact that Manila is closer to Bangkok than it is to Mexico City seems irrelevant to the substance of the debate, and yet it appears to dramatically affect the choice of a counter argument. Similarly, Wednesday is not a more significant day than Thursday in regard to the TV schedule, and yet it is assessed as a better counterargument against Tuesday.

One may view this phenomenon as a rhetoric fallacy. We, however, suggest another view. The logic of debate mechanisms, optimally designed to extract information from the debaters, may be quite subtle. The strength of an argument made by one debater may depend on the argument made by the other debater even when no informational dependencies exist between the arguments.

3. A Model

A decision-maker, called **the listener**, has to choose between two **outcome**s, O_1 and O_2. The "correct" outcome, from the point of view of the listener, depends on the realization of five **aspects,** numbered 1,...,5. An aspect i may get one of two **values,** 1 or 2, with the interpretation that if it gets the value j, aspect i is evidence that supports the outcome O_j. A **state** is a five-tuple of 1's and 2's which describes the realizations of the five aspects. Let ω_j denote the value of aspect j at state ω. Let $i(\omega)=\{j| \omega_j=i\}$ be the set of aspects which support O_i, and denote by $n_i(\omega)$ the size of $i(\omega)$. The listener objective is to choose the outcome supported by the majority of arguments. For each state ω, denote by $C(\omega)$ the O_i for which $n_i(\omega) \geq 3$; we will refer to $C(\omega)$ as the **correct** outcome at ω.

The listener is ignorant of the state. Two agents, called **debater 1** and **debater 2**, have full information about the state. The preferences of the debaters are different from those of the listener. Each debater i prefers that outcome O_i will be chosen, regardless of the state.

We view a debate mechanism as a process in which each debater reveals pieces of information in order to persuade the listener to choose the outcome he (the debater) prefers. We model a debate as a combination of two elements:

<u>Procedural rules</u>, which specify which order and what sort of arguments each debater is allowed to raise.

<u>A persuasion rule</u>, which specifies the relationship between the arguments presented in the debate and the outcome chosen by the listener.

We impose some constraints on the feasible debate mechanisms. Our first constraint is that the only moves a debater can make are to raise arguments that support his preferred outcome; that is, the set of feasible moves of debater i at the state ω is $i(\omega)$, where the move "j" is interpreted as the **argument**: "aspect j is in my favor". There are, in fact, three assumptions implicit in this constraint: First, debaters cannot make any moves other than raising arguments. Second, a debater cannot lie, namely, debater i cannot claim that the value of aspect j is i unless it is indeed i. Third, a debater cannot raise arguments that support the outcome preferred by the other debater. When debater i makes an argument and then debater j makes an argument, in sequence, we will refer to debater j's argument as a **counterargument**.

The second constraint concerns the complexity of the debate. Of course, if a debate in which all five aspects could be revealed was feasible, the correct outcome would be obtained with certainty. The listener could simply ask one of the debaters to make three arguments, and if that debater was able to fulfill this task, his preferred outcome would be chosen. However, we take the view that debate rules are affected not only by the goal of obtaining a good outcome, but also by the existence of constraints on the length and complexity of the debate. We define a debate as follows:

We find it suitable to model a **debate** as an extensive game form of one of three types:

(1) **one-speaker debate:** one of the debaters has to choose two arguments;

(2) **simultaneous debate:** the two debaters move simultaneously, each one has to make one argument;

(3) **sequential debate:** a two-stage game, at the first stage, one of the debaters has to choose one argument and at the second stage the other debater has to make one argument.

An outcome, O_1 or O_2, is attached to each terminal history. The attachment of an outcome to a terminal history is the **persuasion rule**.

Note that a debate is not the game form that will be actually played, as at each state the debaters are not allowed to make false arguments. A debate Γ will induce at each state ω a distinct game $\Gamma(\omega)$, played by the two debaters. The game form of the game $\Gamma(\omega)$ is obtained from Γ by deleting, for each debater i, all moves that are not in $i(\omega)$. If player i has to move after a history h, and at ω none of the arguments he is allowed to make at h is in $i(\omega)$, then we make the history h in the game $\Gamma(\omega)$ a terminal history and attach to h the outcome O_j (debater i looses). As to the preferences of each debater i in $\Gamma(\omega)$, we assume that debater i strictly prefers O_i to O_j independently of ω.

To clarify our construction consider, for example, the simultaneous debate. In a simultaneous debate Γ each player has 5 choices. The game form Γ specifies for each pair of choices, one for player 1 and one for player 2, an outcome, O_1 or O_2. (Actually the assignment of outcomes for the five pairs (t, t) are redundant.) At the state ω=(1,1,2,2,1), for example, the game $\Gamma(\omega)$ is a 3×2 game, where player 1 has to choose between arguments 1,2 and 5 and player 2 has to choose between arguments 3 and 4. The outcome attached to each pair of moves in $\Gamma(\omega)$ is the outcome attached by the game form Γ.

Note the difference between our modeling and that of the standard implementation literature. In both cases the designer determines a game form and the state determines the game played at that state. In the standard implementation literature, the state is a profile of preference relations and the game form played at each state is fixed. In our framework, the preference relations are fixed and the game form varies with the state.

The game $\Gamma(\omega)$ is a zero-sum game and has a value, $v(\Gamma,\omega)$, which is a <u>lottery</u> over the set of outcomes. Let $m(\Gamma,\omega)$ be the probability that $v(\Gamma,\omega)$ assigns to the incorrect outcome. When $m(\Gamma,\omega)=1$ we say that the debate Γ induces a <u>mistake</u> in the state ω.

Note that in simultaneous debates, $m(\Gamma,\omega)$ may get a value that is neither 0 nor 1. Consider, for example, the simultaneous debate Γ with the persuasion rule according to which debater 1 wins if and only if he argues for some i, and debater 2 does not argue for either i+1(mod5) or i-1(mod5). For the state ω=(1,1,2,2,2), the game $\Gamma(\omega)$ does not have a pure Nash equilibrium; the value of this game is O_i with probability ½, for i=1,2 and $m(\Gamma,\omega)=1/2$.

All mistakes are weighted equally. An **optimal debate** is one that minimizes $m(\Gamma)=\sum_\omega m(\Gamma,\omega)$.

Note that we confined ourselves to the simplest example that could demonstrate our point. We needed two arguments in order to create at least the possibility for a debate. We found the case of three aspects uninteresting, as the optimal debate would be the one where only one speaker is required to raise two arguments in order to win. As we are not interested in actually designing optimal debates but only in investigating the logic of debates, we make do with investigating the case of five aspects.

3. Analysis

Main Claim: Every optimal debate procedure is sequential and violates the DC Principle.

The proof of the main claim is accomplished by analyzing the three possible persuasion rules:

(1) Only one debater is allowed to speak

We start with the procedural rule where only one debater, say, debater 1, is allowed to present two arguments, after which an outcome is chosen. A persuasion rule here can be presented as a set of pairs of arguments, where the presentation of one of these pairs is necessary and sufficient for debater 1 to win. For example, the listener may be persuaded by any two arguments that support debater 1's position: This persuasion rule induces 10 mistakes (all states ω where $n_1(\omega)=2$). A more interesting persuasion rule is one where the five aspects are ordered on a circle and the listener is persuaded whenever debater 1 presents any two arguments, referring to two successive aspects. Here, the number of mistakes is only five. (The five mistakes will take place in the five states ω where $n_1(\omega)=2$ and two aspects in favor of 1 are successive. Note that in any state ω in which $n_1(\omega)\geq 3$, two successive aspects must be in favor of debater 1). But, we can do a bit better:

Claim 1: The minimal $m(\Gamma)$ for debates in which only one debater speaks is four. The only debate (up to a permutation of the names of the arguments and the identity of the speaker) which induces 4 mistakes is the one where the speaker is persuasive if and only if he presents two arguments either from {1,2,3} or from {4,5}.

Proof: A debate in which only debater 1 speaks is characterized by a set E of sets of size 2, so that debater 1 wins in every ω in which his set of arguments, $1(\omega)$, contains a set in E. Any $e \in E$ produces one mistake in the state ω in which $1(\omega)=e$.

Consider the debate where debater 1 has to show two arguments regarding aspects which are either in {1,2,3} or in {4,5} (that is, E={{1,2}, {2,3}, {1,3}, {4,5}}). This debate induces four mistakes in favor of 1 and none in favor of 2 (if $n_1(\omega) \geq 3$, then there must be at least two aspects supporting 1, either in the set {1,2,3}or in the set {4,5}).

To see that there is no one-speaker debate with less than four mistakes, and that the one above is the unique optimal one speaker debate (up to a permutation of the names of the arguments and the identity of the speaker), consider a one-speaker debate that induces at most 4 mistakes. Each aspect must appear in one of the sets in E since otherwise there is at last 6 mistakes. It must be that the set E is such that one of the aspects, let us say 1, appears in at most one set of E, let us say {1,2}. For each $\{i,j\} \subseteq \{3,4,5\}$, either {i,j} is in E and there is a mistake in the state ω in which $1(\omega)=\{i,j\}$ and if not there is a mistake in the state ω in which $1(\omega)=\{1,i,j\}$. Neither {2,i} of {2,j} is in E since it would induces a fifth mistake. Thus, if {i,j} is not in E there will be one additional mistake in the state ω in which $1(\omega)=\{2,i,j\}$. We conclude that it must be that E={{1,2}, {3,4}, {3,5}, {4,5}}.

(2) Simultaneous debates

Claim 2: The minimal $m(\Gamma)$ for simultaneous debates is five.

Proof: A simultaneous debate specifies an outcome, O(x,y), to be chosen if debater 1 makes the argument x and debater 2 argues y. Consider the debate Γ, in which $O(x,y)=O_2$

if and only if y=x+1(mod5) or y=x-1(mod5). We will see that $m(\Gamma)=5$. Debater 1 rightly wins in any state ω where $1(\omega)$ contains three successive aspects (ordered on a circle), and he rightly loses in any state where $1(\omega)$ contains zero, one, or exactly two non-consecutive arguments. We are left with ten states, in five of which $1(\omega)$ contains three non-successive aspects (e.g., (1,1,2,1,2)). In the other five states, there are precisely two aspects in favor of debater 1 and they are successive (e.g., (1,1,2,2,2)). In each of these ten states, the value of the induced game is the lottery that selects the two outcomes equally. Thus, $m(\Gamma)=10(½.)=5$.

We will now show that $m(\Gamma)\geq 5$ for any simultaneous debate Γ. For a given aspect x, let v_x be the number of y's so that $O(x,y)=O_1$ and, similarly, for a given aspect y, let w_y be the number of x's so that $O(x,y)=O_2$. Of course, $\sum_x v_x + \sum_y w_y = 20$.

If, for some x, $v_x=4$, debater 1 wins at any state ω with $\omega_x=1$; hence, at least five mistakes are induced. Similarly, if for some y, $w_y=4$, there will be five mistakes induced. Thus, assume that $v_x\leq 3$ for all x and $w_y\leq 3$ for all y. Consider the set Ω` consisting of the 20 states, $\omega$, for which $n_1(\omega)$ is 2 or 3. We will see that $\sum_{\omega\in\Omega`} m(\Gamma,\omega)\geq 5$. For a state ω with $n_1(\omega)=2$, the induced game $\Gamma(\omega)$ is a 2X3 game and the correct outcome is O_2. The value of $\Gamma(\omega)$ will be O_2 with probability 1 only if there is a column (a winning strategy) that guarantees debater 2's victory. A similar argument can be made for a state ω with $n_1(\omega)=3$. Otherwise, $\omega\in\Omega$` contributes at least ½ to $m(\Gamma)$.

If $v_x=1$, then x will never be a winning strategy in any of the 20 games. If $v_x=2$, the action x can be used by debater 1 as a winning strategy in exactly one induced 3X2 game. If $v_x=3$, the strategy x can be used by debater 1 to win in three 3X2 games, but it will also allow him to win wrongly in one 2X3 game. Thus, given any vector $(v_1,..,v_5,w_1,..,w_5)$, any state ω with $\omega(1)$ equaling 2 or 3 will contribute ½ to $m(\Gamma)$, except in $\#\{x|\ v_x=2\}+3\#\{x|\ v_x=3\}+\#\{y|\ w_y=2\}+3\#\{y|w_y=3\}$ states, where it will contribute nothing to m(Γ), and in #{x| v_x=3}+#{y| w_y=3} states, where it will contribute 1 to m(Γ). Thus m(Γ) is at least

$[20-\#\{x|\ v_x=2\}-3\#\{x|\ v_x=3\}-\#\{y|\ w_y=2\}-3\#\{y|\ w_y=3\}]\ /2+[\#\{x|\ v_x=3\}+\#\{y|\ w_y=3\}]=$

$[20-\#\{x|\ v_x=2 \text{ or } v_x=3\}-\#\{y|\ w_y=2 \text{ or } w_y=3\}]/2.$

It is easy to see that the minimum is obtained when all v_x and w_y equal 2 and thus m(Γ)≥5.

(3) Sequential debates

In a sequential debate, one of the debaters, let us say debater 1, is asked to raise an argument and the other debater can respond by raising one argument. In cases where debater 1 raises the argument x, debater 2 raises the argument y, and debater 2 wins the debate, we will say that y counterargues x.

Claim 3: The minimal m(Γ) over all sequential debates is three.

Proof: Consider the sequential debate Γ_1 with the persuasion rule:

If debater 1 argues for....	....debater 2 wins if and only if he counterargues with...
1	{2}
2	{3,5}
3	{4}
4	{2,5}
5	{1,4}

This debate induces three mistakes two in favor of debater 1 (in the states (1,1,2,2,2), (2,2,1,1,2)) and one in favor of debater 2 (in state (1,2,1,2,1)). (Note that in this debate, aspect 3, for example, does not counetrargue aspect 1 and aspect 1 does not counterargue aspect 3). We shall now show that for any sequential debate Γ in which debater 1 starts the debate, $m(\Gamma)\geq 3$. Debater 1 has at most five possible moves. After each of these moves, there is a set of counterarguments, which will persuade the listener to select O_2. Thus, there is a set E of at most five sets of aspects, where debater 1 wins the debate at ω if and only if $1(\omega)$ contains one of the sets in E.

Assume that $m(\Gamma) \leq 2$. No $e\in E$ is a singleton since, had there been one, it would have induced, by itself, five mistakes in favor of debater 1. Any $e\in E$ that consists of two aspects induces one mistake. Thus, E can include at most two sets of size 2. Let $\Omega=\{\omega| n_1(\omega)=3\}$. This set contains ten states.

If no set in E contains two aspects, there is a mistake in each of the 5 states in Ω for which $1(\omega)$ is not a set in E.

If there is only one set in E that contains exactly two aspects, then there are at most 3+4 states in Ω in which debater 1 can win the induced game (three states in which $1(\omega)$ contains the set of two aspects in E and at most four states in which $n_1(\omega)=3$ and $1(\omega)$ is an element in E) and, thus, there are at least four mistakes (one in favor of debater 1 and three in favor of debater 2).

Suppose that E contains precisely two sets of two elements. There are at most 6 sets in Ω for which $1(\omega)$ contains one of these sets. Thus, there must be at least one element in Ω for which $1(\omega)$ does not include a set in E and the number of mistakes must be at least 3.

<u>Claim 4</u>: Any optimal persuasion rule violates the DC Principle.

Proof: By the proof of Claim 3, E does not contain any set of size greater than 3 and contains no more than three sets of size 3. Thus, the number of two-element sets, {x,y}, which are subsets of a set in E, cannot exceed 8 and, hence, there must be two elements, x and y, so that neither x counterargues y nor y counterargues x.

4. Comments

a) The sequential debate game

In the above analysis, the persuasion rule was a part of the design of the mechanism. The listener was not a player in the game. Alternatively, one could think of a sequential debate as a three-stage game in which the listener, after listening to the two arguments, chooses an outcome, with the aim of maximizing the probability that he will choose the correct one. The set of the listener's strategies is the set of all possible persuasion rules for sequential debates.

First, let us check whether the optimal persuasion rules are sequential equilibrium strategies of the extended game. Consider, for example, the persuasion rule described in Claim 3. This persuasion rule is supported by the following sequential equilibrium of the extended game:

- Debater 1's strategy is to raise the first argument, if there is one, for which debater 2 does not have a proper counterargument. Otherwise, debater 1 chooses the first argument which supports him.
- Debater 2's strategy is to respond with the first successful counterargument whenever he has one and, otherwise, to raise the first argument that supports his position.
- The listener chooses the outcome according to the persuasion rule described in the table in the proof of Claim 3.

The full proof that these three strategies indeed constitute a sequential equilibrium consists in dealing with many cases. We make do with demonstrating some of them:

If debater 1 raises the argument 1 and debater 2 responds with argument 3, the listener assigns a probability 0.75 that the correct outcome is O_1; if debater 2 responds with 4 or 5, the listener concludes that the correct outcome is O_1 with certainty. If debater 2 responds with argument 2, the listener concludes that aspect 1 is in favor of debater1 and, in addition to aspect 2, at least one aspect in {3,4} and one aspect in {4,5} are in favor of debater 2; thus the probability that the correct outcome is O_1 is 0.2.

If debater 1 raises the argument 2 and debater 2 responds with argument 1 or 4, the listener concludes that 2,3 and 5 are in favor of 1. On the other hand if debater 2 responds with 3, the listener concludes that aspects 1,3 and either 4 or 5 are in favor of debater 2. If debater 2 responds with 5, it means that aspect 3 is in favor of debater 1, but 1, 4 and 5 are in favor of debater 2.

The case where debater 1 raises the argument 3 and debater 2 responds with 5 is an "out of equilibrium event". From the fact that debater 1 did not raise argument 1, the listener should conclude that either aspect 1 or aspect 2 is in favor of debater 2, and from the fact that debater 2 responded with 5, the listener should conclude that aspects 1 and 2 are in favor of debater 1. We assign to the listener at that history, the belief that aspects 1, 2 and 3 are in favor of debater 1.

Note that the three-stage debate game has other sequential equilibria as well. One of them is particularly natural. In any state ω, debater 1 raises the first argument i which is in his favor. Debater 2 responds with the argument j, which is the smallest j>i in his favor, if such an argument exists; otherwise, he responds with the smallest argument which is in his favor. The listener's strategy will be guided by the following logic: Debater 1, in equilibrium, is supposed to raise the first argument in his favor. If he raises argument i, the listener believes that arguments 1,2,..,i-1 are in favor of debater 2. Debater 2, in equilibrium, is supposed to raise the first argument in his favor following argument i. Hence, if debater 2 raises argument j, the listener believes that arguments i+1,..,j-1 are in favor of debater 1. The listener chooses O_1 if the number of aspects he (the listener) concludes to support 1 is at least the number of those he concludes to support 2. This equilibrium induces seven mistakes.

b) Debates which are mixtures between the one speaker debate and the sequential debate

We restricted the number of arguments which can be raised during a debate to two. This restriction does not preclude another type of debate, a type which was not analyzed in the previous section: Debater 1 makes the first argument and then, depending on which argument was brought up, either debater 1 or debater 2 is required to raise the second argument. One can verify that the minimal $m(\Gamma)$ over these type of debates is also three.

c) On the assumption that a debater can only raise arguments

The assumption that the only actions a debater can take are raising arguments is of course restrictive. For example, consider a mechanism where debater 1 is required to list three aspects and debater 2 wins the debate if and only if he shows that one of these aspects have been realized in his favor. This "mistake-free" debate consists of only two moves and requires "proving" only one argument. However, in this paper, we wish to focus on the relationship between arguments and counterarguments and thus we limit the scope of the discussion to debates where a debater can only raise arguments.

By the way, as a curiosity let us also mention that if we retain the assumption that a debater cannot lie, but we allow a debater to raise arguments against himself, there is a mistake-free debate. Consider a debate where, if debater 1 argues that aspect i is in his favor, debater 2 has to counterargue by showing that either aspect i+1(mod5) or aspect i+2(mod5) is in his (debater 2's) favor in order to win, and if debater 1 concedes that aspect i is in favor of debater 2, debater 2 has to counterargue by showing that one of the aspects, i-1(mod5), i+1(mod5), or i+2(mod5), is in his (debater 2's) favor in order to win. If $n_1(\omega)\geq 4$, then for some i, the 3 aspects, i(mod5), i+1(mod5), and i+2(mod5), are in $1(\omega)$, thus, debater 1 can win by raising argument i. If $n_1(\omega)=3$, then either there is an aspect i so that the three aspects, i(mod5), i+1(mod5), and i+2(mod5) are in $1(\omega)$, and he can win by raising the argument i, or there is an aspect i such that aspects i-1(mod5), i+1(mod5) or i+2(mod5) are in $1(\omega)$, and debater 1 can win by conceding on aspect i. If $n_1(\omega)\leq 2$, debater 2 can rebuff all debater 1's arguments.

5. Discussion

The reader may wonder why it is that in practice we do not observe persuasion rules of the kind described in claim 3. Our view is that an important feature of real life persuasion rules is that they are stated in natural language. The persuasion rules we expect to observe are those that are easily definable using a binary relation over the set of aspects, and which the parties naturally associate with the problem. Let us go back to our "five cities" example. The most salient binary relation on the set of these five cities is the one which partitions the set into the "Far East cities" and the "non-Far East". The sequential debate with the persuasion rule described in Claim 3 cannot be described using these terms alone without referring to the names of the cities. On the other hand, consider the sequential debate with the following persuasion rule: The second speaker is required to counterargue an argument referring to a Far East city with another Far East city and he is required to counterargue a non-Far East city with a non-Far-East city. The sequential debate with this persuasion rule induces seven mistakes. This is actually the best persuasion rule from among those which can be described just by using the terms "Far East city" and "non-Far East city" with no references to the names of the cities. Note that the latter sequential debate is worse than the one-speaker debate in which the speaker is required to present two arguments, either from the set of the Far East cities or from the set of the non-Far East cities, in order to win. This one-speaker debate induces only four mistakes.

Our purpose in this paper was not to provide a general theory of debates. The implications of the constraints imposed by the natural language on optimal debates were not studied. Our only purpose was to demonstrate that a phenomenon we often observe in debates can be explained as an outcome of an optimization which takes into account another "real life" constraint: the limit on the amount of information that the listener can process. Thus, during a debate, the relative "strength" of an argument may not derive from the relative "quality" of the information embodied in that argument. Or, put differently, there may be two arguments such that neither is a persuasive counterargument to the other. We show that this phenomenon is not necessarily a rhetorical fallacy; instead it may be consistent with the outcome of a constrained optimization of the debate rules.

We believe that this phenomenon is connected to considerations of pragmatics. Within a debate, a responder counterarguing to “Bangkok” with anything other than Manila, is interpreted as an admission that Manila is also an argument in favor of the opponent’s position. Or, in the context of question 3, if a debater responds to Tuesday by Thursday, skipping Wednesday, it is considered as an admission that Wednesday does not go in his favor. The fact that the sentence “On Thursday Channel B is superior” is uttered as a counterargument to “On Tuesday Channel A is superior”, gives the sentence a meaning different than what it would have received, had it been stated in isolation.

The fact that, during a conversation, an utterance may acquire meanings beyond those it would have received had it been stated in isolation is related to ideas discussed in philosophy of language. Grice (1989), in particular, argues that the natural interpretation of utterances in natural language contains more information than the meaning given by the utterances in isolation. We apply an economic approach in the sense that we try to provide a rationale for such phenomena by showing that they are outcomes of the optimization of a certain objective function, subject to constraints (see also Rubinstein (1996)).

6. Related literature

We find the spirit of Fishman and Hagerty (1990) the closest to ours. The following is one interpretation of their model. A listener wishes to obtain information from one speaker about the state of nature. The state of nature may receive one of two equally likely values, H or L. The speaker observes K signals about the state, each receives the value of the state with probability $p>1/2$. The speaker aims to increase the probability that the listener assigns to the state H. The constraint on the complexity of the mechanism is that the speaker can present only one signal (he cannot cheat). The "social designer" looks for a persuasion rule that decreases the expected probability that the listener assigns to the false state. The best "persuasion rule" is to order the signals, $s_1,..,s_K$, and to have the listener interpret the signal s_k as an admission by the speaker that the signals $s_1,..,s_{k-1}$ receive the value L. Thus, though all signals are equally informative, the optimal persuasion rule treats them unequally.

Despite its common use in supporting decision making, little has been said about debates in economics and game theory. Several exceptions are reviewed below.

Austen-Smith (1993) studies cheap talk debates where two parties try to influence the action taken by a third party. Initially, the ideal points of the two players are +1 and -1, whereas the third party's ideal point is 0. Each debater gets a signal about a random variable d, which represents for him the shift in his ideal point. The paper investigates the existence of informative equilibria for two types of a cheap talk game, one where the messages are sent simultaneously and the other where the debaters send the messages sequentially. Krishna and Morgan (2000) studies a similar game where the experts have full information about a non binary state.

Shin (1994) analyzes the sequential equilibria of what he calls a "simultaneous moves game of persuasion" (see also Milgrom and Roberts (1986)). In his work, a state of nature, $x \in \{y_1,...,y_K\}$, is the "true" amount of money that party 2 should pay to party 1. For each k, each of the two parties receives (with some probability) a signal that tells him whether $x<y_k$ or $x \geq y_k$. A party cannot present a wrong signal but does not have to disclose all the signals he has received. The two debaters move simultaneously, disclosing whichever signals they decide to reveal to a third party who then determines the amount of money that party 2 will pay party 1. The third party's goal is to decrease the expected distance between his own ruling and the "right" amount (the state). Shin (1994) shows that there is a sequential equilibrium with a simple, attractive structure, according to which each of the two parties reports only good news, namely, signals which confirm that the state is above (or below) a certain cutoff point.

Lipman and Seppi (1995) study another phenomenon, often observed in real life: debaters who convey wrong information are severely punished. They study a model in which a debater can present messages as well as bring evidence to refute an opponent's claims. They argue that “little provability” is sufficient for the existence of an equilibrium in which the listener believes a certain party, unless that party's claim is refuted.

Spector (2000) consider a debate between two parties. Each party tries to “move” the position held by the other party (a point in an Euclidean space) closer to his (the first party's) position. The situation is analyzed as a multi-stage game. At each period, one debater gets information about the true state in the form of a signal and he has to choose whether to disclose the signal to the other debater or not. If a debater presents evidence, it is evaluated by his opponent by taking into account strategic considerations. The authors show that the debaters' positions will "converge" to "a stable configuration of positions", in which no evidence presented by one debater can change the position held by the other.

References

Austin-Smith, D. (1993), "Interested Experts and Policy Advice: Multiple Referrals under Open Rule", Games and Economic Behavior, 5, 3-43.

Fishman, M.J. and K.H. Hagerty (1990), "The Optimal Amount of discretion to Allow in Disclosures", Quarterly Journal of Economics, 105, 427-444.

Grice, P. (1989), Studies in the Way of Words, Cambridge: Harvard University Press.

Krishna, V. and J. Morgan (2000), "A Model of Expertise", mimeo.

Lipman, B.L. and D.J.Seppi (1995), "Robust Inference in Communication Games with Partial Provability", Journal of Economic Theory, 66, 370-405.

Milgrom, P. and J. Roberts (1994), "Relying on the Information of Interested Parties", Rand Journal of Economics, 17, 18-32.

Rubinstein, A. (1996), "Why are Certain Properties of Binary Relations Relatively More Common in Natural Language?" Econometrica, 64, 343-356.

Shin, H.S. (1994), "The Burden of Proof in a Game of Persuasion", Journal of Economic Theory, 64, 253-264.

Spector, D. (2000), “"Rational debate and one-dimensional conflict" Quarterly Journal of Economics (forthcoming).

Proof of comment b

Suppose that debater 1, after making some argument i has to make another argument and he wins if the second argument he raises is in b(i). Note that for every $j \in b(i)$, there is one mistake in the state ω in which $1(\omega)=\{i,j\}$. If b(i) is a singleton, the debate is equivalent to the sequential one in which after debater 1 argues i, debater 2 has to respond with the element which consists of b(i) in order to win. Thus, if $m(\Gamma)\leq 2$, there could be at most one aspect, let us say 1, after which debater 1 is allowed to speak again and he wins if he raises an argument in a set b(i) that consists of exactly 2 aspects, let us say 2 or 3. If debater 1 starts with another argument, he wins unless debater 2 counterargues with an appropriate argument. Thus, there is a set E, of six sets of arguments, sufficient for debater 1 to win, two of which are {1,2} and {1,3}. Out of the 10 sets in $\Omega=\{\omega| n_1(\omega)=3\}$, debater 1 can use the two two-argument sets to win in five states and the other four sets to win in at most four states, leaving one state in which he will lose. Thus, the number of mistakes must be at least three.