9

A Game Theoretic Approach to the Pragmatics of Debate: An Expository Note

Jacob Glazer and Ariel Rubinstein

1 The Logic of Debates

In this paper, the term 'debate' refers to a situation in which two parties disagree over some issue and each of them tries to persuade a third party, the listener, to adopt his position by raising arguments in his favor. We are interested in the logic behind the relative strength of the arguments and counterarguments; we therefore limit our discussion to debates in which one side is asked to argue first, with the other party having the right to respond before the listener makes up his mind.

Casual observation tells us that when arguing in a debate, people often mean more than the literal content of their statements and that the meaning we attach to arguments in debates is dictated by some implicit rules particular to debates. Grice initiated the analysis of the logic by which we interpret utterances in a conversation. His fundamental assumption is that the maxims which guide people in interpreting an utterance in a conversation are derived from the 'cooperative principle'. The cooperative principle is probably valid for interpreting utterances in a conversation but it makes no sense in the context of a debate. In a debate, the listener does not know whether he and a debater who raises an argument have the same or opposing interests. The listener must consider the possibility that the debater's intention is to manipulate him.

Consider, for example, the following scenario presented to a group of students at Tel Aviv University:

> Alice, Michael and you are good friends. You and Michael tend to disagree over most things and you always find yourself trying to persuade Alice that your opinion is the right one. Imagine that the three of you attend the performance of a new band. You and Michael disagree as to whether it is a good band.
>
> Michael says: 'This is a good band. I am especially impressed by the

guy standing on the very left'.

Now is the time for you to make a quick reply. The following two arguments come to mind but you can only choose one of them:

A: 'Yes, but the guy standing right next to him was terrible.'

B: 'Yes, but the one standing fourth from the left was terrible.'

Which of the two arguments would you choose in order to persuade Alice?

About 66% of the subjects thought that A is more persuasive (the results were not sensitive to the order in which the two counterarguments have appeared).

This example appears to confirm that the power of a counterargument goes beyond its literal content. Raising a counterargument that is 'distant', in some natural sense, from the original argument is interpreted as an admission that the debater could not raise a 'closer' counterargument to support his view.

Thus, for example, countering an argument about the musician standing on the very left by an argument about the musician standing fourth from the left will likely be interpreted by Alice as an admission that the musicians standing between these two are good, thus providing an additional evidence in favor of Michael.

We observed similar phenomena in two other experiments discussed in Glazer and Rubinstein (2001). Having to respond to an argument regarding Bangkok, an argument regarding Manila was considered more persuasive than an argument regarding Brussels. Having to counter an argument regarding Monday, an argument regarding Tuesday was considered more persuasive than an argument regarding Thursday.

In other words, a statement made in a debate will have a different meaning than it would have had in a conversation. Though the fact that Manila is closer to Bangkok than Brussels is irrelevant to the substance of the debate, it is relevant to the evaluation of the evidence regarding Manila and Brussels as potential counterarguments to evidence about Bangkok.

Our purpose is not to provide a general theory of the pragmatics of debating but to provide a possible rationale for the phenomenon that the strength of a counterargument depends on the argument it counters. Specifically, we study the listener's optimization problem when his aim is to adopt persuasion rules that minimize the probability that he will reach the wrong conclusion from the arguments. We study this optimization under a constraint on the amount of information the listener can process. We show that the phenomenon in which the 'strength' of the evidence which is brought as a counterargument depends on the initial argument is not necessarily a rhetorical fallacy but may be consistent with optimal persuasion rules.

2 Our approach

Our approach to investigating the logic of debating has several basic components:

1 We view a debate as a mechanism by which a decision maker (the listener) extracts information from two parties (the debaters). The right decision, from the point of view of the listener, depends on the realization of several bits of information. The relevant information is fully known to both debaters but not to the listener. The debaters have opposing interests regarding the decision to be made. During the debate the debaters raise arguments to support their respective positions in the form of providing hard evidence about the realization of some of the bits of information. On the basis of these arguments, the listener reaches a conclusion.

2 The listener bases his decision on what he infers from the arguments made in the course of the debate. The listener might make inferences beyond what follows directly from the information contained in the hard evidence presented by the parties. The pragmatics of debating is the set of rules which guide the listener in interpreting the arguments made by the debaters beyond their literal content.

3 The interpretation of the utterances and evidence is something embedded in the mind of the listener and not a matter of decision.

4 The two debaters view the debate as a two player game. The listener is not a player in the game. The actions in the game are the arguments the debaters can raise. The outcome of the game reflects the debaters' common understanding of how the listener will interpret their arguments in reaching a conclusion.

5 The pragmatics of debating are viewed as if chosen by a fictitious designer. The designer is aware of the fact that the rules he chooses will determine the game that will be played by the debaters and therefore that the listener's conclusion will depend indirectly on the rules the designer chooses. The designer aims to maximize the probability that the listener will reach the same conclusion he would have, had he known all the information known to the debaters.

6 The designer is constrained by physical limitations such as the difficulty in processing information, the length of the debate and the cost of bringing hard evidence.

In sum, we apply an economic approach to the pragmatic of debating. Our study is a part of a research program associating phenomena in language with a rational choice of a mechanism, given physical constraints (see also Rubinstein (1996) discussing a different issue within linguistics).

3 The model

Following economic tradition, we demonstrate our approach using a simple model that is not meant to be realistic.

3.1 The basic scenario

In our model there are three participants, debater 1, debater 2 and a listener. In the final stage of the debate, the listener chooses between two actions, O_1 and O_2. The 'correct' action, from the listener's point of view, depends on the state of the world, thereafter referred to as a 'state'. A *state* is characterized by the realization of five *aspects*, numbered $1, \ldots, 5$. Each aspect i receives one of two values, 1 or 2, with the interpretation that if its value is 1, aspect i supports the action O_1 and if its value is 2 it supports O_2. The profile of realizations of all five aspects is a state. For example, $(1, 1, 2, 2, 1)$ is the state in which aspects $1, 2$ and 5 support O_1 and aspects 3 and 4 support O_2. Each aspect receives the value 1 or 2 with probability 0.5. The realizations of the five aspects are assumed to be independent.

The listener does not know the state. Had he known the state he would have chosen the action supported by a majority of aspects (e.g., O_1 at state $(1, 1, 2, 2, 1)$). The debaters, on the other hand, have full information about the state. The listener, therefore, needs to elicit information from the debaters. However, the listeners do not necessarily wish to give him the missing information since debater 1 prefers the action O_1 and debater 2 prefers O_2 independently of the state. A debate is a process (mechanism) by which the listener tries to elicit information from the two debaters about the true state.

3.2 Debate

A debate procedure specifies the order in which the debaters speak and the sort of arguments they are allowed to raise. We restrict our attention to sequential debates in which the two parties raise arguments one after the other. Thus, debater 1 moves first and debater 2 moves second after listening to debater 1's arguments. We do not allow debater 1 to counter the counterargument of debater 2 and we do not allow simultaneous argumentations (see Glazer and Rubinstein (2001) for a discussion of other debate procedures).

We assume that a debater can only raise arguments regarding aspects that support his position. Denote by $\mathrm{arg}(i)$ the argument: 'the realization of aspect i supports my position'. For example, in state $(1, 1, 2, 2, 1)$, debater 1 can raise one of the three arguments $\mathrm{arg}(1)$, $\mathrm{arg}(2)$, and $\mathrm{arg}(5)$ while debater 2 can counterargue with $\mathrm{arg}(3)$ and $\mathrm{arg}(4)$.

A few assumptions are implicit in this debate model: Debaters cannot make any move other than raising arguments; they cannot shout, curse or whatever. More importantly, a debater cannot lie. Thus, for example, a debater cannot claim that the value of an aspect is 1 unless that is in fact the case. The motivation of this assumption can either be that debaters simply hate to lie (or are afraid to get caught lying) or that making an argument that aspect 4 has the value 1 requires more than just making the utterance 'aspect 4 has the value 1' and requires providing hard evidence that proves this beyond any doubt. Third, a debater cannot raise arguments that support the outcome preferred by the other debater.

We now come to another key feature of our approach which concerns the complexity of the debate. We assume that each debater is allowed to raise at most one argument, i.e., to present the realization of at most one aspect. This assumption is especially realistic for debates in which an argument requires more than simply making a short statement. In a court case, for example, parties have actually to bring witnesses. In a scientific debate, debaters may have to provide an explanation as to why their findings support a particular hypothesis. Thus, arguments in a real life debate may require scarce resources such as time or the listener's attention.

Of course, if there were no restrictions on the complexity of the debate, the listener could simply ask one of the debaters to present three arguments, ruling in his favor if he is able to fulfill this task. This would allow the listener to elicit enough information so that he would always make the correct decision. However, we take the view that debate rules are influenced by the existence of constraints on the complexity of the debate. This approach is in line with our understanding that Grice's logic of conversations is dictated, among other things, by the fact that there are limits on the ability of the conversers to communicate and process information.

3.3 Persuasion rules and the debate game

The focus of our analysis is the concept of *persuasion rules*. A persuasion rule determines whether the action O_1 or O_2 is taken, given the events that took place in the debate. It reflects the way that the listener interprets debater 1's argument and debater 2's counterargument. We assume that if a debater does not raise any argument in his turn, he loses the debate. Formally, denote by Λ the set of feasible arguments. A *persuasion rule* F is a function which assigns to every element in Λ a subset (possibly empty) of arguments in Λ. The meaning of $\lambda_2 \in F(\lambda_1)$ is that if debater 1 makes the argument λ_1 then debater 2's counterargument λ_2 persuades the listener to take the action O_2.

A persuasion rule induces, for every state, a two-stage extensive game with perfect information: In the first stage, debater 1 chooses one of the arguments that support his claim (if such an argument exists, otherwise the game degenerates to be the outcome O_2). In the second stage debater 2 chooses an argument from those that support his claim (if such an argument exists; otherwise the game degenerates to debater 1's choice of an argument followed by the outcome O_1). Following any possible pair of arguments, the action to be taken (O_1 or O_2) is determined by the persuasion rule. With respect to preferences, debater 1 always prefers O_1 and debater 2 always prefers O_2. Note that the listener is not modeled as a player. At this stage we assume that the listener simply follows his 'instincts' about what to infer from the players' moves. The rules of pragmatics are embedded in those instincts.

Example: To clarify our construction, consider the following persuasion rule: the listener is persuaded by debater 2, if and only if (i) debater 1 is silent or (ii) debater 1 presents $\arg(5)$ or (iii) debater 1 presents $\arg(i)$ ($i = 1, 2, 3, 4$) and debater 2 counterargues with $\arg(i + 1)$. At the state $(1, 2, 1, 2, 2)$, for example, this persuasion rule induces the game in Figure 9.1.

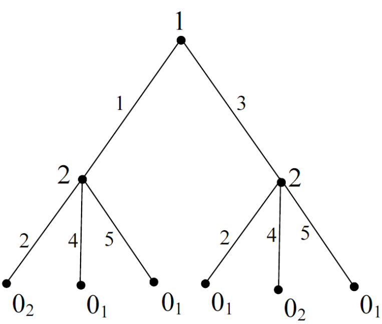

Figure 9.1

3.4 The designer's problem

For any state s, and for every persuasion rule, the induced game between the two debaters is a zero-sum game $\Gamma(s)$ with perfect information. Adopting the standard game theoretic view that players play rationally and that player 1 anticipates player 2's move, we can use the concept of Nash equilibrium. Such games have a unique Nash equilibrium outcome and thus, we can talk about 'the outcome of the game $\Gamma(s)$'. If the listener's correct action at state s is not the outcome of the game $\Gamma(s)$, we say that the persuasion rule induces a mistake at state s.

To clarify our terminology, consider the persuasion rule described in the above example. Debater 1 can persuade the listener to take the action O_1 only in a state with two aspects i and $i+1$ supporting him. In such states debater 1 is able to raise the argument $\arg(i)$ without player 2 being able to rebuff him with $\arg(i+1)$. Therefore, in any of the 4 states $(1,1,2,2,2)$, $(2,1,1,2,2)$, $(2,2,1,1,2)$, $(2,2,2,1,1)$ debater 1 will win the debate although he should have lost, whereas in the state $(1,2,1,2,1)$ he will lose the debate although he should have won. In any other state the outcome of the induced game is the correct one. Thus, the number of mistakes is 5 and the probability of a mistake, induced by this persuasion rule, is $5/32$.

We finally come to the designer's problem. The designer seeks a persuasion rule that minimizes the probability of a mistake. Given our assumptions this is equivalent to finding a persuasion rule with the smallest number of states in which the outcome of the game induced by the rule is the wrong one. Note that this problem reflects the assumption that all mistakes are of equal significance. In addition, it makes no differentiation between the weight put on ruling O_2 in the state $(1,1,1,1,1)$ where all aspects support the decision O_1 and the weight put on ruling O_2 in the state $(1,1,1,2,2)$ where the realization of the aspects only marginally support O_1. Intuitively we often feel that the former mistake is 'bigger' than the latter.

4 Analysis

We will now investigate optimal persuasion rules, namely those that minimize the number of mistakes.

Claim 1: The minimal number of mistakes induced by a persuasion rule is three.

Proof: Consider the following persuasion rule:

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

This persuasion rule induces three mistakes. Two mistakes are in favor of debater 1: in the state $(1, 1, 2, 2, 2)$, if debater 1 raises $\arg(1)$ debater 2 does not have a good counterargument and in the state $(2, 2, 1, 1, 2)$, debater 2 does not have a counterargument to $\arg(3)$. There is also one mistake in favor of debater 2 in the state $(1, 2, 1, 2, 1)$ since whatever debater 1 argues debater 2 has an appropriate counterargument.

We shall show now that any persuasion rule induces at least three mistakes. Fix a persuasion rule. Following any move by debater 1, there is a set of counterarguments, available to debater 2 that will persuade the listener to select O_2. Debater 1 can win the debate by raising $\arg(i)$ only if all those aspects that, according to the persuasion rule, debater 2 could possibly counterargue with and win, are in favor of debater 1. Thus, a persuasion rule is characterized by a set E of at most five sets of aspects such that debater 1 can win the debate in state s if and only if the set of aspects that support him in state s contains a set in E.

If any of the sets in E is a singleton $\{i\}$, raising $\arg(i)$ would be sufficient to persuade the listener to decide in favor of debater 1, inducing at least 5 mistakes (in the states were the arguments supporting debater 1 are only $\{i\}$ or $\{i, j\}$ for some $j \neq i$).

Denote by S_3 the set of 10 states with exactly 3 aspects supporting debater 1 (such as $(1, 1, 1, 2, 2)$). Any set in E that consists of two aspects only, induces one mistake in the state where only these two aspects support debater 1.

If there is only one set in E that contains exactly two aspects, then there are at most 7 states in S_3 in which debater 1 can win the induced game (in three states in S_3 these two aspects support 1; and there are at most four states in which the set of aspects supporting 1 contains an element in E) and thus there are at least three mistakes.

Suppose that E contains precisely two sets of two aspects. There are at most 6 sets in S_3 that contain one of these two sets of aspects. Thus, there must be at least one element in S_3 for which the set of aspects supporting 1 does not contain any set in E and, in that state, debater 1 cannot make his case. Thus, the number of mistakes must be at least 3.

Comment: The above persuasion rule is not the only optimal persuasion rule. See Glazer and Rubinstein (2001) for another persuasion rule which is not isomorphic to this one.

At this point, it is of interest to discuss what we call the *Debate Consistency (DC) Principle*: for any pair of arguments x and y, either y is a persuasive counterargument to x or x is a persuasive counterargument to y but not both.

This principle is violated by the optimal persuasion rule described in the proof of Claim 1. If debater 1 argues $\arg(1)$ and debater 2 counterargues $\arg(3)$, debater 1 wins and if debater 1 argues $\arg(3)$ and debater 2 counterargues $\arg(1)$, debater 1 also wins. This violation is not a coincidence:

Claim 2: Any optimal persuasion rule violates the DC Principle.

Proof: By the proof of Claim 1, the set E associated with an optimal persuasion rule 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 aspects, x and y, such that neither $\arg(x)$ counterargues $\arg(y)$ nor $\arg(y)$ counterargues $\arg(x)$.

Thus, the logic of debate mechanisms may be quite subtle and the relative strength of arguments may depend on the order in which they are raised even in the absence of informational dependencies between these arguments.

5 Ex-post optimality

An important feature of our discussion so far has been that the listener's optimization is carried out *ex-ante*, that is he chooses the persuasion rule before the debate starts. The listener is committed to follow the persuasion rule. Given an argument and a counterargument, the persuasion rule 'dictates' to the listener which action he should take without leaving him the liberty to deviate from the action he is supposed to take according to the persuasion rule. This kind of commitment induces the debaters to argue in a way that minimizes the probability that the listener makes the wrong decision.

It is possible, however, that a persuasion rule is ex-ante optimal but is not *ex-post* optimal: given a course of events in the debate, the listener might make inferences which would lead him not to follow the persuasion rule he has originally announced. If the listener figures out the debaters' strategies in the induced game at every state, it may happen that a certain combination of an argument and a counterargument is more likely to take place in states where the correct outcome is O_1 but the persuasion rule assigns the outcome O_2 for that combination. It could happen that even though a certain persuasion rule was ex-ante optimal, the listener would find the action that he is supposed to take according to the persuasion rule ex post (i.e., after the debaters have spoken and after he updated his beliefs given his knowledge of the debaters' strategies in the different states) unoptimal and, hence, he would want to deviate from the persuasion rule.

In this section we examine the question whether the (ex-ante) optimal persuasion rule described in the proof of Claim 1 above, is optimal also ex-post.

Think about the debate as a four-stage (Bayesian) game. In the first stage, nature chooses the state and the debaters, though not the listener, are informed of nature's choice. In the second and third stages the debaters sequentially make arguments (within the above constraints) while in the fourth stage the new player, namely the listener, chooses an outcome. The listener prefers the correct outcome over the incorrect one. In other words, the listener's strategy is equivalent to the choice of a persuasion rule and his objective in the game is to maximize the probability of making the correct decision. Does this game have a sequential equilibrium in which the listener follows the optimal persuasion rule?

One can show that the persuasion rule specified in Claim 1 is indeed a part of the following sequential equilibrium:

- Debater 1's strategy is to raise the first argument, if such an argument exists, for which debater 2 does not have a persuasive counterargument. Otherwise, debater 1 chooses the first argument in his favor.

- Debater 2's strategy is to respond with the first successful counterargument, whenever such an argument exists. Otherwise, he raises the first argument in his favor.

- The listener chooses the outcome according to the persuasion rule described in the proof of Claim 1.

The full proof that these three strategies are part of a sequential equilibrium involves dealing with a large number of cases. We will make do with demonstrating the main idea.

Assume that debater 1 raises $\arg(1)$ and debater 2 responds with $\arg(3)$. Given the above strategies, the play of the game described above is consistent with the four states $(1, 1, 2, x, y)$, where $x = 1, 2,\ y = 1, 2$. Given that in three of these states the majority of the aspects support debater 1, the listener's decision O_1 is ex post optimal.

If debater 2 responds to $\arg(1)$ with either $\arg(4)$ or $\arg(5)$, given the above strategies, the listener should conclude that aspects 2 and 3 are in favor of debater 1 and, therefore, it is ex post optimal for the listener to choose O_1.

If debater 2 responds with $\arg(2)$, the listener must conclude that, in addition to aspect 2, at least one aspect in $\{3,4\}$ and one aspect in $\{4,5\}$ are in favor of debater 2. There are five states that are consistent with the above conclusion. In only one of these, $(1,2,1,2,1)$, debater 1 should win. Thus, the probability that the correct outcome is O_1 is 0.2 and the listener's plan to choose O_2 is also ex-post optimal.

Note that the four-stage debate game has some other sequential equilibria as well, one of which is particularly natural: Debater 1 raises $\arg(i)$ where i is the first aspect whose realization supports debater 1. Debater 2 responds with $\arg(j)$, where j is the next aspect after i whose realization supports debater 2, if such an argument exists and otherwise, he responds with the first argument which is in his favor.

The listener's strategy will be guided by the following logic: in equilibrium, debater 1 is supposed to raise the first argument in his favor. If he raises $\arg(i)$, then the listener believes that aspects $1,2,..,i-1$ are in favor of debater 2. If debater 2 raises $\arg(j)$, the listener believes that the realization of aspects $i+1,..,j-1$ are in favor of debater 1. The listener chooses O_1 if the number of aspects that he believes to support debater 1 is at least the same as the number of those that he believes support debater 2. This equilibrium induces seven mistakes. For example, in the state $(1,1,2,2,2)$ agent 1 will start with 1 and debater 2 will respond with 3, inducing the listener to correctly believe that aspect 2 supports debater 1 and to wrongly choose O_1.

6 Language

Even though the persuasion rule described in the proof of Claim 1 is optimal, it does not have an intuitive interpretation and cannot be described using natural language. We expect real life debate rules to be easy to understand and, in particular, that they can be described in terms available within their context. In this section we wish to demonstrate, using a number of examples, how the vocabulary available to the fictitious designer affects the optimal persuasion rules. In all the examples, the group of aspects is a set of five individuals $I = \{1,2,3,4,5\}$.

1 A partition with two cells

Assume that the set I divides naturally into two sets M (males) and F (females). Assume that a debater cannot state an argument of the type $\arg(i)$, but can only make one of the following two arguments: $\arg(M) =$

'the facts regarding a male support my position' and $\arg(F)$ ='the facts regarding a female support my position'. Recall that a persuasion rule assigns to each of the arguments a subset of arguments each of which persuades the listener to take the action O_2. Thus, a persuasion rule is a function which assigns to every element in $\{\arg(M), \arg(F)\}$ one of the four sets of arguments $\emptyset$, $\{\arg(M)\}$, $\{\arg(F)\}$ or $\{\arg(M), \arg(F)\}$). There are exactly 16 persuasion rules.

Consider, for example, the case $M = \{1, 2, 3\}$ and $F = \{4, 5\}$. It is easy to see that within this vocabulary, the minimal number of mistakes is 7 which is attained by the following persuasion rule: Debater 2 wins the debate if and only if he counters $\arg(M)$ with $\arg(M)$ and $\arg(F)$ with $\arg(F)$. This persuasion rule induces one mistake in favor of debater 1 in the state $(2, 2, 2, 1, 1)$ in which debater 2 does not have a counterargument to $\arg(F)$. In the 6 states in which two aspects in M and one in F support debater 1, debater 2 is always able to counterargue successfully although he should lose.

Note that the persuasion rule which requires debater 2 to counter $\arg(M)$ with $\arg(F)$ and $\arg(F)$ with $\arg(M)$ also yields 7 mistakes (in the states in which 3 aspects, two of which are males, support debater 1).

Now consider the case $M = \{1\}$ and $F = \{2, 3, 4, 5\}$. One optimal persuasion rule is such that debater 1 wins the debate if and only if he argues $\arg(F)$ and debater 2 does not counterargue with $\arg(M)$. Debater 2 will win, even though he should lose, in the 5 states where at least 3 aspects in F support debater 1 and aspect 1 supports debater 2. Debater 2 will lose the debate, even though he should win, in the four states in which agent 1 is supported by aspect 1 in addition to exactly one of the aspects in F. Another optimal persuasion rule is the one according to which debater 1 wins if and only if he argues $\arg(M)$.

2 One single and two couples

Individual 3 is single, whereas '1 and 2' and '4 and 5' are couples. Assume that the debaters cannot refer to any particular couple. Once debater 1 refers to a particular married individual debater 2 can refer to 'his partner' or to 'an individual from the other couple'. In other words, the persuasion rule cannot distinguish between 1 and 2 and between 4 and 5, or between the couple $\{1, 2\}$ and the couple $\{4, 5\}$.

Thus, the set of arguments for debater 1 includes arg(single)= 'the facts regarding the single individual support my position' and arg(married) = 'the facts regarding a married individual support my position'.

The only possible counterargument to arg(single) is arg(married). There are three possible counterarguments to arg(married) : arg(single), arg (married to the individual debater 1 referred to) and arg(married but not to the individual to whom debater 1 referred to).

One can show that any persuasion rule induces at list 6 mistakes. One optimal persuasion rule is the following: if debater 1 makes the argument arg(single) and debater 2 replies with arg(married) he loses the debate. If debater 1 makes the argument arg(married) the only successful counterargument is arg(married to the individual debater 1 referred to).

Given this rule, debater 1 will win the debate if and only if there is a couple $\{i, j\}$ such that both i and j support debater 1′s position. Thus, debater 1 will win the debate, although he should lose, in the two states $(1, 1, 2, 2, 2)$ and $(2, 2, 2, 1, 1)$ and will lose the debate, when he should win, in the 4 states where the single individual and two married individuals from distinct couples support his case.

Another optimal persuasion rule is to require debater 1 to argue arg(married) and to require debater 2 to counterargue with arg(married but not to the individual to whom debater 1 referred to) In this case all 6 mistakes will be in favor of debater 2.

3 Neighborhoods

Assume that the debaters have in mind a notion of neighborhood but cannot refer to any particular individual. Debater 1 can only make the argument 'here is an individual who supports my case' and debater 2 has two possible counterarguments: arg(neighbor) = 'the facts about a neighbor of the person debater 1′s referred to, support O_2' and arg(non-neighbor) = 'the facts about a non-neighbor of the person debater 1′s referred to, support O_2'. Thus, there are four possible persuasion rules.

If the five aspects are located on a line in the order $1, 2, 3, 4, 5$, the optimal persuasion rule dictates that in order to win, debater 2 has to counterargue with arg(neighbor). This persuasion rule induces 5 mistakes. Two mistakes are in the states $(1, 1, 2, 2, 2)$ and $(2, 2, 2, 1, 1)$, where debater 1 can win the debate by referring to individuals 1 or 5 respectively although he should lose. In the three states $(1, 2, 1, 2, 1)$, $(1, 2, 1, 1, 2)$ and $(2, 1, 1, 2, 1)$, debater 1 cannot win the debate even though he should, since arg(neighbor) is always available to debater 2.

4 An Ordering

Assume that the debaters have in mind an of ordering of the aspects. Debater 1 can only make the argument 'here is an individual who supports my case' and debater 2 has only two possible counterarguments: $\arg$(superior)='the facts about an individual who is superior to the individual debater 1 referred to, support O_2' and $\arg$(inferior)='the facts about an individual who is inferior to the person debater 1 referred to, support O_2'. Thus, there are four possible persuasion rules. The minimal number of mistakes in this example is 10 and is obtained by the persuasion rule which requires debater 2 to counterargue with $\arg$(superior).

7 Related literature

This chapter is based on Glazer and Rubinstein (2001) which presented a game theoretic approach to the logic of debates.

Several papers have studied games of persuasion in which one agent tries to persuade another to take a certain action or to accept his position. In particular, see Milgrom and Roberts (1986), Fishman and Hagerty (1990), Shin (1994), Lipman and Seppi (1995) and Glazer and Rubinstein (2004).

Several other papers studied cheap talk debates in which two parties attempt to influence the action taken by a third party. In particular, see Austen-Smith (1993), Spector (2000) and Krishna and Morgan (2001).

Those readers who are familiar with the implementation literature may wonder about the relation between that approach and ours. In both cases the designer determines a game form and the state determines the particular game played in that state. However, in the standard implementation literature, the state is a profile of preference relations and the game form played in each state is fixed. In our framework, the preference relations are fixed and the game form varies with the state.

Bibliography

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**, 427–44.

Glazer, J. and A. Rubinstein (2001). Debates and decisions, on a rationale of argumentation rules. *Games and Economic Behavior*, **36**, 158–73.

Glazer, J. and A. Rubinstein (2004). On optimal rules of persuasions. *Econometrica*, **72**(6), 1715–36.

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

Krishna, V. and J. Morgan (2001). A model of expertise. *The Quarterly Journal of Economics*, **116**(2), 747–75.

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 (1986). 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**, 343–56.

Shin, H. S. (1994). The burden of proof in a game of persuasion. *Journal of Economic Theory*, **64**, 253–64.

Spector, D. (2000). Rational debate and one-dimensional conflict. *Quarterly Journal of Economics*, **115**, 181–200.