NORTHWESTERN UNIVERSITY

Models of Persuasion

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Economics

By

Itai Sher

EVANSTON, ILLINOIS

June 2007

ABSTRACT

Models of Persuasion

Itai Sher

I study several aspects of a game-theoretic model of persuasion. A speaker attempts to persuade a listener to take an action which is highly ranked by the speaker. The listener knows the speaker's preference but is uncertain about what the speaker can say. The listener can commit to a persuasion rule, which is a response to the speaker's messages.

Chapter 1 provides an introduction.

Chapter 2 studies conditions under which optimal persuasion rules are deterministic and credible, extending results of Glazer and Rubinstein (2006) from two to many actions.

Chapter 3 studies the lattice theoretic structure underlying the persuasion problem. I study implementable outcome functions (i.e., mappings from types to actions induced by some persuasion rule). Families of implementable outcome functions which can arise in some persuasion problem correspond to interior systems on the set of types, a notion from lattice theory. This leads to a characterization of messages as being essential or redundant.

Chapter 4 studies the additional structure which imposed by the assumption that the speaker does not face time, attention, or other similar communication constraints. The absence of such constraints is captured by the notion of normality of Bull and Watson (2007), and related to the nested range condition of Green and Laffont (1986). Under normality, the representation in terms of interior systems reduces to one in terms of quasi-orders.

The main result of Chapter 5 is that in the finite case, the listener's utility function is guaranteed to be a modular function of the set of implementable outcome functions exactly when normality holds; otherwise, the listener's utility function may not be quasi-supermodular. It follows that under normality, all messages become more persuasive as the interests of the speaker and listener become more aligned, and when normality fails, one can always find a counter-example. Likewise, under normality, there always exists a symmetric optimal rule, whereas when it fails, examples in which all optimal rules are asymmetric are found.

Chapter 6 studies an integer programming formulation of the problem.

Acknowledgements

I am very grateful to my dissertation committee: Eddie Dekel, Asher Wolinsky, Marciano Siniscalchi, and David Austen-Smith. Eddie Dekel, my advisor, gave me valuable advice and guidance throughout the course of this project. I feel that the project gained enormously in response to the various constructive criticisms and observations made by Eddie. Eddie often spots key rhetorical and substantive weaknesses in a model or argument, and thinking about and responding to these comments over a long period of time really helped to improve the quality of my work. In addition, Eddie's practical advice was very important to me. My discussions with Asher were always very valuable. Asher made several comments which stayed in my mind for months after he made them. Marciano's advice helped me tremendously throughout my graduate career. I always felt that Marciano devoted a lot of thought to what I was doing. There was at least one criticism which I had a very hard time responding to, and Marciano would bring it up almost every time we talked. I want to thank him for pushing me in this way. It was always very helpful to talk to David. David really encouraged me to try to make my ideas more accessible, and this was very useful advice.

There were many other people at Northwestern who made my graduate career very memorable and fruitful. Among them were Nabil Al-Najjar, Michael Coates, Allan Collard-Wexler, Jeff Ely, Kripa Freitas, Maria Goltsman, Hanna Halaburda, Jakub Kastl, Ron Laschever, David Lucca, Charles Manski, Alessandro Pavan, Marcin Peski, Viswanath

Pingali, Adam Rosen, Maria Salgado, Joerg Stoye, Tomasz Strzalecki, Mike Whinston, Fan Zhang, and Charles Zheng. Northwestern is a great place to be a graduate student in economics, both because of the intellectual atmosphere, and because of the people.

Finally, I would like to thank my family. Among many other things, my brother Shlomi provided me with very valuable practical advice while I was on the academic job market, which I think was critical to the outcome. I don't understand how he had so much insight into a process which he has never experienced. I am very grateful to my mother Gila for many things. She set an intellectual example for me to follow from an early age. My father Peter has helped me in so many ways. He is an endlessly generous person, and discussions that I have had with him on so many topics have done much to shape me into who I am today. My wife Melissa has been a true friend throughout this process. On occasions when I was discouraged, Melissa helped me, and when I was happy, Melissa made me happier. Going through this process was essentially a joint experience. My son Jonah was born about a year and a half before the completion of this dissertation. Jonah has brought so much happiness into our lives. Already, Jonah has many insightful things to say, and I am sure that I have a lot more to learn from him in the future. I dedicate this dissertation to Melissa and Jonah.

Table of Contents

List of Figures

CHAPTER 1

Introduction

This dissertation studies several aspects of a game-theoretic model of persuasion, involving a speaker and a listener. The listener can choose one of several actions. The speaker knows the true state, prefers higher actions (according to some ranking) and may make a statement in order to persuade the listener to take a high action. The listener's preference depends on the state and he must commit to a decision rule–called a *persuasion rule*–in response to the speaker's messages before the interaction.

An important property of the model studied here is that while the speaker's preferences are known to the listener, the set of messages available to the speaker is not known. Thus, in selecting a persuasion rule, the listener must only reason about what the speaker could say in various circumstances. The main contribution of this dissertation is a characterization of the underlying lattice structure of the persuasion problem, which provides an approach to analyzing the diverse possibilities which can arise in this setting. However, I also deal with aspects of the persuasion problem which do not directly invoke the lattice structure. In the remainder of this chapter, I will outline the contributions of the chapters of this dissertation, and then I will present the basic model which is developed in different ways throughout.

Chapter 2 studies the determinism and credibility of optimal persuasion rules. A persuasion rule is *deterministic* if in response to each message, it selects a single action, rather than a probability distribution over many actions. A persuasion rule is *credible* if the commitment assumption is unnecessary. In other words, a persuasion rule f is credible if the timing of the game can be reversed so that first–anticipating the listener's response–the speaker sends a message m using a (possibly mixed) strategy which is a best response to f, and then, for any message m which the speaker uses with positive probability, the listener actually finds it in his interest to play $f(m)$ given the speaker's strategy. Glazer and Rubinstein (2006) showed that when there are two actions, there exists an optimal persuasion rule which is both credible and deterministic. In a related model, Glazer and Rubinstein (2004) found that the optimal persuasion rule is credible but not necessarily deterministic.

In Chapter 2, I extend the analysis of Glazer and Rubinstein (2006) to the case where the listener can choose from several actions and not just two. This extension is interesting for several reasons. To begin with, with several actions it is no longer necessarily true that there exists an optimal persuasion rule which is *either* credible *or* deterministic. Thus, we can ask, what properties of the problem would lead to the existence of such optimal rules. Secondly, in the case of multiple actions, a concavity property is relevant to both determinism and credibility, and this is hidden in the case of two actions.

With respect to both determinism and credibility, it is possible to get partial results without any special assumptions on the utility functions aside from the assumption that the speaker's preferences are known and strict, where the latter assumption was also made by Glazer and Rubinstein (2006). These partial results fall short of ensuring the existence of either an optimal deterministic or optimal credible persuasion rule, but their restriction to the case of two actions ensures the existence of an optimal deterministic credible rule.

Moreover, it is possible to establish existence of an optimal credible deterministic persuasion rule under stronger conditions on the utility functions. However, these conditions are always satisfied in the case of two actions. In a sense, looking from the perspective of multiple actions, the results of Glazer and Rubinstein are over-determined. One can look at them as special cases of either the partial results with many actions, or of the results guaranteeing the existence of an optimal deterministic credible rule under stronger assumptions.

The partial results are as follows. First, there always exists an optimal persuasion rule f such that (i) for every message m, $f(m)$ either puts probability one on a single action or $f(m)$ is a probability distribution over two non-adjacent actions (according to the ranking representing the speaker's ordinal preference), and (ii) any best response to f gives each type of speaker a utility equal to the utility to some pure action. This falls short of determinism. However, with two actions (i) implies the existence of an optimal deterministic persuasion rule, since there are no non-adjacent actions. (ii) also implies the existence of an optimal deterministic persuasion rule, because when there are only two actions, over which the speaker is not indifferent, the only way to give the every type of speaker the utility to some pure action is by actually giving him that pure action with probability one. One can also show that for every persuasion rule f which is optimal *among deterministic rules*, and any pair of adjacent actions a and a', there is a speaker strategy σ which is a best response to f and such that for any message m with $f(m) = a$, conditional on the distribution of types and σ, the listener would not be better off playing a' than a after seeing m, and such that the same condition also holds when the roles of a and a' are reversed. This falls short of implying credibility because the strategy σ depends on the pair of actions a and a', and nothing is said about deviations other then a and a'. Nevertheless, if there are only two actions, this clearly implies credibility. In fact, as one might expect the proof of this fact is essentially the same as Glazer and Rubinstein's proof for two actions. Notice however that the logic here depends on the result that with two actions there exists an optimal deterministic persuasion rule, because we are not assuming that f is optimal, only that it is optimal among deterministic rules.

A sufficient condition for the existence of an optimal deterministic persuasion rule with many actions is that at every state, the listener's state-dependent utility function is a concave transformation of the speaker's state-independent utility function. A sufficient condition for the ensuring that any persuasion rule which is optimal among deterministic persuasion rules is credible is that that there exists some state-independent real valued function r representing the speaker's ordinal preferences, such that the listener's utility function is a concave transformation of r at every state of the world. It is not necessary that r represent the speaker's cardinal preferences, but if it does, then it follows from the above that there exists an optimal deterministic credible persuasion rule.

Chapter 3 examines the lattice structure underlying the persuasion problem. Here and throughout the remainder of the dissertation, I restrict attention to deterministic persuasion rules. In a persuasion problem, the choice set for the listener is implicitly the set of outcome functions–i.e. functions from states to actions–which are consistent with the speaker's incentives. I characterize the choice sets which can arise in some persuasion problem. The characterization is in terms of a notion from lattice theory known as an *interior system*. This is a family of sets of types of speaker which contains the empty set, the set of all types, and is closed under union. I show that for a fixed number of actions, interior systems correspond one-to-one with the possible choice sets for the listener.

The characterization has several consequences for the structure of the listener's decision problem. For instance, if one knows which outcome functions are feasible when there are two actions, one can infer which outcome functions would be feasible with k actions. Moreover, for a fixed number of actions, I define a closure operator τ which allows one to infer the feasibility of some outcome functions from the feasibility of others. A family of outcome functions corresponds to a choice set in some persuasion problem if and only if it is a fixed point of τ. This means that in a persuasion problem, for the listener, having some choices entails having other choices, and reflects the listener's ability to make decisions based on arbitrary properties of the speaker's messages.

The representation of the listener's possible choice sets also allows me to undertake a detailed analysis of the structure of messages. I provide a method for inferring all important information about messages from the listener's choice set. Moreover, I identify a set of *essential* messages and show that all other messages are redundant. In deciding which arguments to find persuasive, the listener must only consider essential messages. Essential messages are characterized both from a global perspective (i.e., relative to the set of all messages), and from a local perspective (i.e., relative to the messages available to a given type). The maximum possible number of essential messages when there are n types is shown to be approximately $\binom{n}{n/2}$.

Chapter 4 studies the additional structure which is imposed on the persuasion problem when it is assumed that the speaker can summarize his information. An important distinction in persuasion situations concerns the question of whether the speaker has the ability to present all of his information, or on the other hand, faces time or attention constraints which limit his ability to present all of his information. The assumption that the messages available to the speaker depend on the state can be used to model the speaker's inability to provide all his information. For example, consider a job applicant, who may or may not have time to present all of his qualifications. If the job applicant faces a time constraint, the applicant may have a message corresponding to each of his qualifications, but no message corresponding to all of his qualifications. The absence of time constraints can be modeled by assuming that the applicant has a single message which corresponds to all of his qualifications. Presentation of this message is tantamount to the transmission of all of his information.

In Chapter 4, I use the lattice theoretic characterization derived in Chapter 3 to study what additional structure is imposed on the speaker's choice set when it is assumed that the speaker can summarize his information. In order to define formally what it means to be able to summarize, I consider a condition on the message space known as *normality*. This notion was introduced by Bull and Watson (2007), and is similar to an earlier concept known as the *nested range condition*, which was introduced by Green and Laffont (1986). These authors studied mechanism design environments with provability, and their focus was primarily to find conditions, such as normality, which validate a version of the revelation principle in such environments. Related papers include Deneckere and Severinov (2001), Forges and Koessler (2005) and Singh and Wittman (2001).[1] In contrast

[1]Singh and Wittman (2001) study implementation with and without the nested range condition. Unlike these authors, I provide a method of identifying whether a family of outcome functions could be the choice set in any persuasion problem, use a lattice-theoretic characterization to study properties of optimal persuasion rules, and show how these properties depend on conditions such as normality. Other more distantly related papers which deal with implementation and related issues under provability include Bull and Watson (2004) and Lipman and Seppi (1995).

to these papers, my primary focus is not the revelation principle but rather the effect of the ability to summarize on other properties in a persuasion setting.

I show that under normality, the representation of the listener's choice set in terms of interior systems reduces to a representation in terms of quasi-orderings. In particular, under normality, a family of outcome functions corresponds to the listener's choice set for some specification of the persuasion problem if and only if it is the family of outcome functions which are monotone with respect to some quasi-ordering. Thus the notion of an interior system generalizes that of a quasi-ordering precisely by providing a means of representing gaps in the speaker's ability to summarize information. In fact, one can use this analysis to identify precisely which choices are not available to the listener because of the speaker's inability to summarize: these choices correspond to outcome functions which are monotone but which do not satisfy a stronger condition. In particular, there is some action such that the listener cannot separate the types who receive at least that action from those who do not.

I also introduce the related notion of *weak normality*. Intuitively, weak normality is intended to capture the situation in which the speaker can summarize small, but not large pieces of information. For example, suppose that if the speaker has time to relate a piece of information A, as well as a piece of information B, he has time to relate both, but he does not have time to relate *all* of his information. This can be made rigorous by an assumption which allows the speaker to summarize finite but not infinite collections of information. As discussed above, the listener's choice set is always a lattice. However, I prove that it is a *sub*lattice of the set $\mathbf{A}^T$ of all functions from the set of types to the set of actions (ordered by the componentwise order) if and only if the message structure satisfies weak normality. The listener's choice set is a subcomplete sublattice of $\mathbf{A}^T$ if and only if the message structure is normal. With finitely many types, the notions of normality and weak normality coincide, as do the notions of sublattice and subcomplete sublattice.

Chapter 5 uses the characterizations derived in the previous chapters to analyze some qualitative properties of the persuasion situation. I find that the property of normality is critical, assuming finitely many types so that normality and weak normality are equivalent. In particular I find that the comparative statics of the persuasiveness of messages depends on normality. Secondly, I find that the existence of symmetric optimal persuasion rules depends on normality. Thus optimal persuasion rules may be necessarily asymmetric which, I interpret in terms of the requirement that certain messages must be treated nonliterally.

The first step, leading to these results is a result showing that the speaker's ability–or lack thereof–to summarize information can be represented as a property of the listener's utility function, when considered directly as a function of the lattice of implementable outcome functions. I show that the listener's objective function can *always* be represented as a modular function of this lattice if and only if normality holds. Moreover, whenever it does not hold, one can specify the listener's objective in such a way that it is not even quasi-supermodular. This is a consequence of the fact that with finitely many types, normality is necessary and sufficient for the listener's choice set to be a sublattice of $\mathbf{A}^T$. The listener's utility function is easily seen to be a modular function of $\mathbf{A}^T$. Thus, only under normality does it inherit this property when considered as a function of the set of implementable outcome functions.

This has consequences in terms of comparative statics. In particular, it is intuitive that if the interests of the speaker and the listener become more aligned *in every state*, then the listener will grant the speaker a weakly higher (i.e., more preferred) action in every state, but this conclusion is valid for all specifications of the listener's objective only if the speaker can summarize his information. Otherwise, the listener may choose a persuasion rule which makes the speaker better off in some states and worse off in others. This conclusion can be translated into a result about messages: under normality as interests become more aligned, all messages become more persuasive, but when normality fails, some message may become less persuasive.

Secondly, I show that failure of normality is critical for the ability for the solution to the persuasion problem to represent certain pragmatic phenomena. In a series of papers, Glazer and Rubinstein (2003, 2004, 2006) used models of persuasion similar to the model studied in this dissertation to provide a model which could represent certain pragmatic phenomena in a strategic setting. Pragmatics is the subfield of linguistics which studies conversational meaning, which is meaning that arises in conversation over and above the literal meaning of the words used. In a classical account, Grice (1989) proposed certain cooperative principles of conversation to account for such phenomena. However, as Glazer and Rubinstein point out, such phenomena can arise in strategic situations in which agents would not be expected to obey Grice's cooperative principles. They present a series of examples in which structurally identical messages are treated differently by optimal persuasion rules. This can be interpreted as representing conversational meaning, because under the optimal persuasion rule, messages which do not differ structurally, and hence in their "literal meaning" are treated differently. While Glazer and Rubinstein never explicitly mention the notion of normality, it so happens that all the examples that they present violate normality. In other words, all these examples feature something like time constraints. I show that this is not an accident. In particular, under normality, the listener's problem corresponds to maximization of a modular–hence supermodular–function on a lattice. By a well known result, this implies that the set of maximizers is a lattice, hence has a greatest element. I define a formal notion of symmetry for persuasion problems, and show that under normality the greatest element of the lattice of maximizers can always be implemented by a symmetric persuasion rule, that is, a persuasion rule which treats messages which are structurally the same similarly. Thus while there may be optimal rules which treat structurally identical messages differently, this does not happen essentially in the sense that such different treatment is never a necessary condition for a solution of the persuasion problem.

I also identify a class of non-normal problems where no optimal symmetric rule exists. Ideally, one would like to show that whenever normality fails one could specify the listener's utility function in some way so that there would be no optimal symmetric rule. Unfortunately this is not possible for a rather trivial reason; the message structure may be such that every pair of messages are structurally different from one another. However another theorem is possible which emphasizes the same point. In the counter-examples to the existence of symmetric optimal rules when the message structure is not normal, there are collections of messages $\{m_1, m_2, \ldots, m_n\}$ which are structurally identical but such that no optimal rule treats them all in the same way. Suppose for example that every optimal rule treats m_1 and m_2 differently, in that $f(m_1) = a_1$ and $f(m_2) = a_2$. Of course, it is possible to find another optimal persuasion rule f' such that $f'(m_1) = a_2$

and $f'(m_2) = a_1$. But there may be no optimal rule which assigns either a_1 or a_2 to both m_1 and m_2. Thus the treatment of m_1 and m_2 is interdependent at the optimal rule. Such interdependence is impossible for any pair of messages, regardless of whether they are structurally the same or not, under normality. However, one can always specify the listener's utility function so that such interdependence emerges for some pair of messages whenever normality fails.

Chapter 6 relates the approach of the research of this dissertation to the "L-principle" of Glazer and Rubinstein (2006), which is an integer program whose solution is equivalent to that of the persuasion problem. I use the lattice theoretic approach to extend the "L-principle" to multiple actions, and show that it reduces to the well-known maximal closure problem under normality with two actions.

1.1. The Model

Assume that there is a speaker and a listener. The listener may be one of several types in a set T. The speaker is of type t with probability $\pi(t) > 0$. For each type $t \in T$, there is a *nonempty* set $M(t)$ of messages which are available to t. In other words, different types have different messages at their disposal. I assume that for all $t \in T$, $M(t) \neq \emptyset$, which means that each type can say something. Let $\mathbf{M}$ be the set of all possible messages, so that $\bigcup_{t \in T} M(t) \subseteq \mathbf{M}$. I refer to the set-valued map $M : T \rightrightarrows \mathbf{M}$, as the *message correspondence.* The tuple $(T, \mathbf{M}, M(\cdot))$ is called a *message structure.* Except for in Chapters 3 and 4, T and $\mathbf{M}$ are assumed to be finite.

Most models in which the set of available messages depends on the type (e.g., Milgrom and Roberts (1986), Lipman and Seppi (1995)) rationalize this dependence either by the assumption that the speaker must present hard evidence or that the penalties to certain messages, which are interpreted as "lies", are severe. Such an interpretation is appropriate here, but a broader interpretation is also possible. For example, in a political debate, the potential answers which a candidate perceives in response to a question may depend on his personality, his way of thinking, or the ideology of his party. If one simply counts the number of sentences that a speaker could possibly say, it is very implausible that a speaker would consider all possibilities, and one would expect that idiosyncratic characteristics of a speaker will determine which possibilities he considers. A second function of the varying message space, as will be explained in Section 4.1, will be to model the possibility that a speaker may not be able to present all his information.

The listener selects actions from a finite set $\mathbf{A} := \{1, \dots, k\}$. It is often convenient to write $\mathbf{A} = \{a_1, \dots, a_k\}$, but formally the action a_j is identified with the number j. The speaker has a continuous strictly increasing von Neumann-Morgenstern utility function $u : \mathbf{A} \to \mathbb{R}$, which does not depend on his type. Thus, the speaker prefers higher actions independently of his type. The listener has a von Neumann-Morgenstern utility function $v : \mathbf{A} \times T \to \mathbb{R}$. Thus, the listener's utility depends both on the action, and on the speaker's type.

A *persuasion rule* is a function $f : \mathbf{M} \to \Delta\mathbf{A}$, in other words, a function which maps messages into probability measures over actions. A *speaker strategy* $\sigma : T \to \Delta(\mathbf{M})$ which maps types into probability distributions over messages. I also use the notation $\sigma(t, m)$ to denote the probability that type t plays message m. Let Σ be the set of all speaker strategies.

The timing of the game between the speaker and the listener is as follows:

(1) The listener commits to a persuasion rule f.

(2) Nature selects the speaker's type according to π

(3) The speaker selects–possibly randomly–a message $m \in M(t)$.

(4) An action is selected according to the probability distribution $f(m)$.

The idea behind this timing is that the listener searches for the best *rule* for responding to the speaker's request.

Define an outcome function to be a function $g : T \to \Delta\mathbf{A}$ mapping types into probability measures over actions. Let $E_{f(m)}$ be the expectation operator with respect to the probability measure $f(m)$. A persuasion rule f *implements* an outcome function g if there exists a sender strategy $\sigma \in \Sigma$ satisfying:

$$\begin{aligned} \forall t \in T, \forall m \in \mathbf{M}, \sigma(t,m) > 0 &\Rightarrow m \in \text{argmax}_{m \in M(t)} E_{f(m)}[u(a)] \\ \forall t \in T, g(t) &= \sum_{m \in M(t)} \sigma(t,m) f(m) \end{aligned}$$

A deterministic persuasion rule is a persuasion rule f such that $f(m)$ always puts probability one on a single action, and likewise a deterministic outcome function g is an outcome function such that $g(t)$ always puts probability one on a single action. When speaking of deterministic persuasion rules, I treat such rules as functions $f : \mathbf{M} \to \mathbf{A}$, and likewise a deterministic outcome function is treated as a function $g : T \to \mathbf{A}$. Notice that deterministic persuasion rules always implement deterministic outcome functions. Whereas implementation is generally a many-to-many relation between persuasion rules and outcome functions, when restricting attention to deterministic persuasion rules, it is a many-to-one relation. It is not difficult to see that a deterministic persuasion rule f

implements a deterministic outcome function function g if and only if

$$g(t) = \max\{f(m) : m \in M(t)\}, \tag{1.1}$$

In the deterministic case, since the implemented outcome function is unique, we write $g = g_f$ for the outcome function implemented by f. A deterministic outcome function g is *implementable* if there exists a deterministic persuasion rule f that implements it.

A persuasion rule f is *optimal* if it gets the highest possible expected utility for the listener given that the speaker chooses the strategy which is best for the listener among his best responses. An *optimal deterministic* persuasion rule is a persuasion rule which is both optimal and deterministic. A persuasion rule is *optimal among deterministic rules* if it is optimal given that the listener's choice is restricted to deterministic rules.

.

CHAPTER 2

Determinism and Credibility

This chapter extends results of Glazer and Rubinstein (2006) concerning determinism and credibility of optimal persuasion rules. Glazer and Rubinstein (2006) showed that, in the case of two actions, there is an optimal deterministic persuasion rule. Moreover, they showed that any optimal persuasion rule can be credibly implemented in the sense that there is exists a sequential equilibrium of the game in which the speaker moves first and the listener does not commit to a persuasion rule, whose outcome is the same as the outcome of the optimal rule.

In this chapter I extend these results to the case of multiple actions. This is interesting in part because the results of Glazer and Rubinstein (2006) do not hold generally in the case of multiple actions, but only under certain assumptions on the utility functions of the players. Thus an examination of the multiple action case yields additional insight as to why and when such results would be true generally. It turns out moreover that a set of conditions sufficient for the existence of optimal deterministic persuasion rules is closely related to conditions sufficient for credible implementation of optimal deterministic rules.

The determinism and credibility results play different roles with respect to the research presented in this dissertation. Both pertain to the scope of the analysis. In later chapters, I study the lattice structure underlying deterministic persuasion rules, and use this structure to develop comparative statics results, as well as results which pertain to the structure of the set of optima. Restricting attention to deterministic rules is justified in part by the fact that they are natural and have an interesting structure. However, the justification is bolstered by the analysis of this chapter which shows that there is a broad class of specifications of the model for which deterministic rules are optimal.

The credibility result plays a similar role in terms of determining the scope of the analysis. There may be instances in which the commitment to a persuasion rule is plausible; the listener may announce in advance what arguments he would find persuasive, and his commitment to such an announcement may be enforced by his reputation. Nevertheless, in many persuasive situations, commitment is not possible. The credibility result says that under certain assumptions, conclusions derived from studying optimal persuasion rules under commitment are always *consistent with* equilibrium predictions about the outcome of persuasion without commitment. Among all persuasion rules consistent with equilibrium, those which are designed to optimize the listener's objective may be of special interest.

Section 2.1 of this chapter presents conditions under which some optimal persuasion rule is deterministic. The condition is that at every state of the world, the listener's utility function is a concave transformation of the speaker's utility function. A surprising result is that regardless of whether this condition holds, there always exists an optimal persuasion rule in which every speaker type gets an expected utility equal to the utility of some pure action. Section 2.2 studies conditions under which optimal persuasion rules can be credibly implemented. More specifically, conditions are found under which persuasion rules which are *optimal among deterministic rules* can be credibly implemented. This means that conditions are found under which the best deterministic persuasion rule can be credibly implemented even when there is a random persuasion rule which dominates it. When combined with the results of Section 2.1, this implies conditions under which persuasion rules which are optimal relative to all rules–random and deterministic–can be credibly implemented. The condition under which rules which persuasion rules which are optimal among deterministic rules can be credibly implemented is that there exists a strictly increasing function of the actions r such that at every state of the world, the listener's utility function is a concave transformation of r. If r is the speaker's utility function, then the conditions of both the determinism and credibility results are simultaneously satisfied.

2.1. Optimality of Deterministic Persuasion Rules

This section will present two results. The first result shows that there always exists an optimal persuasion rule, which gives each type of speaker an expected utility equal to the utility of some pure action. The second result establishes conditions under which there is an optimal deterministic persuasion rule.

Theorem 2.1. *There is an optimal persuasion rule f such that (1) for every message $m \in \mathbf{M}$, either $f(m)$ selects a single action with probability 1, or $f(m)$ randomizes over two non-adjacent actions, and (2) in any best response to f, each type of speaker gets an expected utility equal to the utility to some pure action.*

Proof. See Section 2.3.1. □

A sketch of the proof is as follows. First, it is not difficult to see that there must be an optimal rule f^* and a speaker best response σ^* to f^* such that (i) each type t of speaker sends some message m^t with probability 1 according to σ^* (where different types may send the same message), (ii) σ^* is a best response which maximizes the expected utility of the listener among speaker best responses to f^*, and (iii) f^* assigns the lowest action with probability 1 to any message which is not equal to m_t for some type t. Therefore, the listener can restrict attention to persuasion rules which satisfy (iii) with respect to the fixed set of messages $\mathbf{M}^* = \{m^t : t \in T\}$ described in (i). Enumerate the messages in $\mathbf{M}^*$ so that $\mathbf{M}^* = \{m_1, \ldots, m_n\}$. Any persuasion rule f satisfying (iii)[1] then corresponds to a vector $\alpha = (\alpha^i_j)^{i=1,\ldots,n}_{j=1,\ldots,k} \in [0,1]^{nk}$, where α^i_j is the probability that the persuasion rule selects action a_j if message m_i is reported. Assuming that the speaker uses strategy σ^*, the listener's expected utility is linear in α. Noting that each type of speaker has the same preferences, choose α to maximize the listener's expected utility subject to the constraint that the ranking (in terms of *weak* inequalities) of the speaker's utility to sending any message in $\mathbf{M}^*$ is the same as under f^*, as well as inequalities guaranteeing that α^i is a probability distribution over actions for each i. It is easy to see that the set of points P satisfying these constraints is a polytope (i.e., the convex hull of a finite number of points in R^{nk}). P was designed specifically so that the fact that σ^* is a best response to f^* implies that σ^* is a speaker best response to any persuasion rule f corresponding to a point $\alpha \in P$. Next observe that the vector α^* corresponding to the optimal persuasion f^* belongs to P. Therefore, any α which maximizes the listener's expected utility subject to belonging to P must correspond to an optimal persuasion rule. Since, when maximizing a linear function on a polytope, some maximizer is always an extreme point of the polytope, the proof is completed by examining the constraints defining P to show that any extreme

[1]relative to the *fixed* set $\mathbf{M}^*$.

point β of P satisfies:

$$\sum_{j=1}^{k} u(a_j)\beta_j^i \in \{u(a_1), \ldots, u(a_k)\}. \tag{2.1}$$

$$|\{j : \beta_j^i > 0\}| \leq 2. \tag{2.2}$$

for all i. Note that (2.1) and (2.2) together imply that at any extreme point β of P, if $\beta_{j_1}^i > 0, \beta_{j_2}^i > 0$, then $j_2 \neq j_1 + 1$.

An alternative way to see that property (ii) must hold is to observe that given any optimal persuasion rule f^* and speaker best response σ^* satisfying (i)-(iii), then it is already the case that any message m not used in equilibrium is assigned to a single action with probability one by f^*, and for any $m = m^t$ which is such that $f^*(m^t)$ assigns positive probability to more than two actions, it is possible to take probability mass off an action a_j which gives the listener either the lowest or the second lowest expected utility conditional on m^t and transfer it other actions which give the speaker a weakly higher expected utility while keeping the speaker's expected utility to message m^t unchanged, so that σ^* is still a best response for the speaker. In fact, since f^* is already optimal, probability mass must be transferred from a_j to actions which give the listener the same expected utility as a_j conditional on m^t. Thus we can see that there is an optimal persuasion rule such that for every message m, $f^*(m)$ puts positive probability on at most two actions. To see that f^* can be constructed so that whenever $f^*(m)$ puts positive probability on exactly two actions, these can be assumed to be non-adjacent, we must appeal to the fact that the extreme points of P have property (2.1).

The following example shows that in general it is not possible to strengthen Theorem 2.1 to the statement that there is an optimal deterministic persuasion rule.

Example 2.1. Suppose that there are two types, t_1, and t_2. As always, it is assumed that each type occurs with positive probability. Assume that the message correspondence is given by $M(t_1) = M(t_2) = \{m_1, m_2\}$. Suppose that $\mathbf{A} = \{a_1, a_2, a_3\}$, $u(a_j) = j$ for all $a_j \in \mathbf{A}$. Suppose further that v is given by:

v	a_1	a_2	a_3
t_1	1	0	1
t_2	0	1	0

Figure 2.1. Listener's Utility Function

Recall that the optimal rule is defined in such a way that if there are multiple speaker best responses to the optimal rule, then the speaker selects the best response which is best for the listener. In light of this consideration, there are two optimal persuasion rules, each of which responds to one message with action a_2 and to the other by randomizing over actions a_1 and a_3 with equal probability. Notice that–consistent with Theorem 2.1–each type of speaker gets a utility of 2, which is the same as that type's utility to action a_2. □

In the example above, $v(\cdot, t_1)$ is not a concave transformation of u. The condition that at every state of the world, the speaker's utility function is a concave transformation of the listener's utility function is sufficient for a deterministic optimal persuasion rule, both in the continuous and discrete case.

Theorem 2.2. *Assume that for all $t \in T$, there exists a concave function $c_t : \mathbb{R} \to \mathbb{R}$ such that for all $a \in \mathbf{A}$, $v(a,t) = c_t(u(a))$. Then there is an optimal deterministic persuasion rule.*

Proof. As in the proof of Theorem 2.1, one can argue that there must exist an optimal persuasion rule f^* and a speaker best response σ^* to f^* satisfying (i)-(iii).[2] The listener's utility to σ^* and f^* is:

$$\sum_{t \in T} E_{f^*(m^t)}[c_t(u(a))]\pi(t). \tag{2.3}$$

Theorem 2.1 says that there is an optimal persuasion rule in which each type of speaker receives the utility to some pure action. It is clear from the proof sketch of Theorem 2.1 that this property cannot conflict with (i)-(iii). This implies that f^* could have been chosen so that for each $m^t \in \mathbf{M}^*$, there exists $a^t \in \mathbf{A}$ such that $u(a^t) = E_{f^*(m^t)}[u(a)]$.[3] Now consider the persuasion rule f^{**} which assigns the lowest action with probability 1 to any message outside $\mathbf{M}^*$, and such that for each $m^t \in \mathbf{M}^*$, $f^{**}(m^t)$ selects a^t with probability 1. Every message delivers the same utility to the speaker under both f^* and f^{**}. So σ^* is a best response to f^{**}, and the listener's utility to f^{**} and σ^* is:

$$\sum_{t \in T} E_{f^{**}(m^t)}[c_t(u(a))]\pi(t) = \sum_{t \in T} c_t(u(a^t))\pi(t) = \sum_{t \in T} c_t(E_{f^*(m^t)}[u(a)])\pi(t). \tag{2.4}$$

It follows from Jensen's inequality that (2.4) is at least as large as (2.3), so the optimality of f^* implies the optimality of f^{**}, and f^{**} is deterministic. □

[2] See the discussion after Theorem 2.1.

[3] If there exist t, t' such that $m^t = m^{t'}$, then assume $a^t = a^{t'}$.

The analysis above puts an interesting perspective on the result in Glazer and Rubinstein (2006) that in the case of two actions, there always exists an optimal deterministic rule. Consider the following two statements:

(1) There exists an optimal rule which gives each type of speaker the utility to a pure action.

(2) There exists an optimal deterministic rule.

(1) is just Theorem 2.1. (2) implies (1), but the converse is not true. In order to establish (2) with many actions, we must assume that the listener's utility function is a concave transformation of the speaker's utility function at every state. The latter condition is trivially satisfied with only two actions. Thus (1) and (2) are equivalent with two actions, which is the case studied by Glazer and Rubinstein (2006). Moreover, in the case of two actions, an optimal non-deterministic persuasion rule which gives every type of speaker the utility to a pure action can only assign non-degenerate probability distributions to messages which are not used in equilibrium, and in fact, always implements the same outcome function as some deterministic persuasion rule. This is not true in the case of many actions. Given the equivalence of (1) and (2) in the case of two actions, the proof of Glazer and Rubinstein (2006) does not reveal the role of the concavity assumption in going from (1) to (2).[4] In fact the theorem in Glazer and Rubinstein (2006) can be viewed as a special case both of Theorem 2.1 and of Theorem 2.2. Notice finally that, with two actions, the existence of an optimal deterministic rule is also equivalent to part (i) of Theorem 2.2, which says that there exists an optimal persuasion rule such that for every

[4]This is not intended as a criticism, but rather shows the value added by looking at the case of multiple actions.

message m, $f(m)$ either plays a single action with probability one or randomizes over two non-adjacent actions.

2.2. Credibility of Optimal Persuasion Rules

In the case of two actions, Glazer and Rubinstein (2006) show that any optimal persuasion rule can be credibly implemented, meaning that there is a game in which the speaker moves first, the listener does not commit, and a sequential equilibrium of that game with the same outcome as the given optimal rule.

Glazer and Rubinstein (2006) provide the following counter-example to the proposition that any rule which is optimal among deterministic rules can be credibly implemented when there are more than two actions. The following is almost identical to the example presented there, although I will extend the example to show that it is not generally possible to credibly implement an optimal persuasion rule when there are more than two actions.[5]

Example 2.2. Suppose that there are two types $T = \{t_1, t_2\}$, and the probability distribution π is such that $\pi(t_1) = .4$ and $\pi(t_2) = .6$. There are three actions $\{a_1, a_2, a_3\}$, and the message correspondence is given by $M(t_1) = \{m_1\}, M(t_2) = \{m_1, m_2\}$. The following is the listener's utility function:

v	a_1	a_2	a_3
t_1	0	-1	1
t_2	0	1	-1

Figure 2.2. Listener's Utility Function

Restricting attention to deterministic rules, any optimal rule must have the property that type t_2 attains at least as high an action as type t_1. This is because type t_2 can always

[5]This is how Glazer and Rubinstein (2006) describe what they are doing, but this is incorrect.

mimic type t_1. It follows that the unique optimal persuasion rule among deterministic rules is the rule f such that $f(m_1) = a_1$ and $f(m_2) = a_2$. However, in this case, upon seeing m_1, the listener would know that the speaker's type is t_1, and therefore would prefer to take action a_3. So the optimal rule among deterministic rules is not credibly implementable.

Glazer and Rubinstein implicitly restrict attention to deterministic rules in this example, despite the fact that deterministic rules may not be optimal when there are more than two actions.[6] In fact, in this example, as long as $u(a_1) < u(a_2) < u(a_3)$, but regardless of the exact values, an optimal rule cannot be deterministic. In particular, consider the random rule f', which responds to message m_2 by taking action a_2, and responds to m_1 by taking action a_1 with probability $\frac{u(a_3)-u(a_2)}{u(a_3)-u(a_1)}$, and taking action a_3 with probability $\frac{u(a_2)-u(a_1)}{u(a_3)-u(a_1)}$. This random rule would make the speaker indifferent between messages m_1 and m_2, and assuming that t_1 chooses m_1 (as he must), and t_2 chooses m_2, this rule improves upon the optimal deterministic rule. Note however that this random rule is also not credible because upon seeing m_2, the listener would prefer to take action a_3 with probability 1 rather than randomizing over a_1 and a_3.

In fact f' is an optimal rule, which establishes the desired conclusion, namely that with more than two actions, it may not be possible to credibly implement an optimal persuasion rule. To see that f' is an optimal persuasion rule, observe that the listener can implement an outcome function g if and only if

$$\sum_{i=1}^{3}[g(t_2)(a_i) - g(t_1)(a_i)]r_i \geq 0,$$

[6]Notice that this example cannot satisfy the sufficient condition found in the previous section for the optimality of some deterministic rule.

where $g(t)(a)$ is the probability that $g(t)$ selects a. In other words, the listener can implement an outcome function g if and only if g gives type t_2 a higher expected utility than type t_1. It follows that an outcome function such that $g(t_2)(a_1) > 0$ cannot be give the listener the highest expected utility among implementable outcome functions, because this implies that $g(t_1)(a_3) < 1$, and thus the listener could simultaneously increase the probability that type t_2 gets a_2 and the probability that type t_1 gets a_3, in the process increasing his own utility. So, at the optimum, $g(t_2)(a_1) = 0$. Likewise $g(t_1)(a_2) = 0$ because otherwise transferring the probability mass in $g(t_1)$ from a_2 to a_1 would be feasible and increase the listener's utility.

Given that the support of an optimal implementable outcome function $g(t_2)$ is contained in $\{a_2, a_3\}$, it now follows from Theorem 2.1 that there either is an optimal implementable outcome function g' such that $g'(t_2)(a_2) = 1$, or else an optimal implementable outcome function g'' such that $g''(t_2)(a_3) = 1$. If g' is optimal, then as we have seen, the listener will choose $g'(t_1)$ with support contained in $\{a_1, a_3\}$, and the listener would like to choose $g(t_1)$ to put as much probability mass as possible on a_3 subject to the constraint that $E_{g(t_1)}[u(a)] \leq u(a_2)$. This occurs when $g'(t_1)(a_1) = \frac{u(a_3)-u(a_2)}{u(a_3)-u(a_1)}$ and $g'(t_1)(a_3) = \frac{u(a_2)-u(a_1)}{u(a_3)-u(a_1)}$, so that the total expected utility for the listener given g' is:

$$.4(\frac{u(a_2) - u(a_1)}{u(a_3) - u(a_1)}) + .6 > 0.$$

On the other hand, if g'' is optimal, then we must have $g''(t_1)(a_3) = 1$, but then the listener's expected utility given g'' is $.4 - .6 = -.2$, but this means that g'' cannot be optimal, which in turn implies that g' is optimal. However, as we have seen above, g'

corresponds to a persuasion rule f' described above, which, as we have seen, cannot be credibly implemented. $\square$

I now present the formal definition of credibility.

Definition 2.1. *A deterministic persuasion rule f is credible if there exists a speaker strategy σ such that:*

(2.5)

$$\forall t \in T, \forall m \in \mathbf{M}, \sigma(t,m) > 0 \quad \Rightarrow \quad f(m) = g_f(t).$$

(2.6)

$$\forall m \in \mathbf{M}, \sum_{t \in T} \sigma(t,m) > 0 \quad \Rightarrow \quad f(m) \in \operatorname{argmax}_{a \in \mathbf{A}} \sum_{t \in T} v(a,t) \frac{\sigma(t,m)\pi(t)}{\sum_{t \in T} \sigma(t,m)\pi(t)}.$$

(2.7)

$$\forall m \in \mathbf{M}, \sum_{t \in T} \sigma(t,m) = 0 \quad \Rightarrow \quad \exists \mu \in \Delta(\{t \in T : m \in M(t)\}),$$

$$f(m) \in \operatorname{argmax}_{a \in \mathbf{A}} \sum_{t \in T} v(a,t)\mu(t).$$

Condition (2.5) defines sender strategy σ to be a best response to f. Condition (2.6) says that for any message which is played with positive probability by some sender type, f specifies a best response for the listener conditional on seeing m when the listener computes his updated belief via Bayes' rule using π and σ. (2.7) says that when the speaker sees a message m which is not sent in equilibrium, f specifies an action which is a best response to some belief μ which only puts positive probability on types who could have sent m. This amounts to a sequential equilibrium because any such μ can be generated as the limit of beliefs induced by totally mixed speaker strategies which converge to σ.

The following theorem gives conditions under which a persuasion rule which is optimal among deterministic persuasion rules can be credibly implemented.

Theorem 2.3. *Suppose that*

> **(*):** *there exists a strictly increasing function* $r : \{i, \ldots, k\} \to \mathbb{R}$*, such that for all* $t \in T$*, there exists a concave function* $c_t : \mathbb{R} \to \mathbb{R}$ *with* $v(a_i, t) = c_t(r(i))$*.*

Then any persuasion rule which is optimal among deterministic rules can be credibly implemented.

Proof. See Section 2.3.2. □

Notice that if v satisfies (*) with r defined by $r(i) := u(a_i)$, then by combining Theorem 2.2 and 2.3, there exists an optimal persuasion rule which can be credibly implemented, and moreover, every optimal persuasion rule which is deterministic can be credibly implemented. The difference between the requirements of the theorems is that Theorem 2.2 requires that the listener's utility function be a concave transformation of the speaker's utility function at every state of the world in order for there to exist an optimal deterministic rule, whereas Theorem 2.3 requires that at every state, the listener's utility function be a concave transformation of *some* strictly increasing function which is independent of the state.

The following theorem gives some insight as to the role of the concavity assumption (*):

Theorem 2.4. *Let f be a persuasion rule which is optimal among deterministic persuasion rules, and then for any $j, j+1$ such that $1 \leq j \leq k-1$, there exists a speaker strategy σ which is a best response to f, and such that for any m with $f(m) = a_j$, the listener (weakly) prefers action a_j to a_{j+1} conditional on the event that the speaker sends m and for any m' such that $f(m') = a_{j+1}$, the speaker (weakly) prefers a_{j+1} to a_j conditional on the event that the speaker sends m'.*

This theorem does not require assumption (*), and is very similar to the credibility result in Glazer and Rubinstein (2006). The proof is similar to the proof of their theorem, and also to Step 1 of the proof of Theorem 2.3 in Section 2.3.2. In view of Example 2.2, the conclusion of Theorem 2.3 does not generally hold without (*). A comparison of Theorems 2.3 and 2.4 suggests two related roles for the assumption (*). First, if f is a persuasion rule which is optimal among deterministic rules, then by Theorem 2.4 be able to find a speaker best response σ_1 to f such that (i) given σ_1, upon seeing any message m such that $f(m) = a_j$, the listener would prefer a_j to a_{j+1}, as well as a speaker best response σ_2 such that (ii) given σ_2, upon seeing any message m with $f(m) = a_j$, the listener would weakly prefer a_j to a_{j-1}. On the other hand, it may not be possible to find any speaker best response σ which satisfies both (i) and (ii) simultaneously. Assumption (*) is sufficient to ensure the existence of a speaker strategy which satisfies both (i) and (ii) simultaneously. Secondly, Theorem 2.4 does not guarantee that for $\ell < i < j$, or $\ell > i > j$, there exists a speaker best response σ such that whenever $f(m) = a_\ell$, upon seeing m, the listener prefers a_ℓ to both a_i or a_j. Assumption (*) also guarantees the existence of such a speaker best response.

One might conjecture that the condition (*) is more than is required and that it would be sufficient that the listener's utility function be "single-peaked" at every state of the world, or, more precisely, that the difference $d(j,t) = v(a_{j+1},t) - v(a_j,t)$ satisfies a single-crossing property whereby for $j_1 < j_2$, $d(j_1,t) \leq 0 \Rightarrow d(j_2,t) \leq 0$ and $d(j_1,t) < 0 \Rightarrow d(j_2,t) < 0$. However, this property would not be strong enough, as the following example demonstrates.

It is interesting to compare this with the requirement in Theorem 2.2, which requires that for each $t \in T$, there is a concave function c_t with the property that for all $t \in T$, $v(a_i,t) = c_t(r_i)$. Given that r_i is increasing in i, both of these requirements are clearly consistent with one another, so that when they both hold, they imply that there is in fact an optimal rule which is deterministic and that every optimal rule which is also deterministic can be credibly implemented, so that some optimal rule can be credibly implemented.

Example 2.3. Consider the following example. $T = \{t_1,t_2,t_3\}$. $M(t_1) = M(t_2) = \{m_1\}, M(t_3) = \{m_1,m_2\}$. Suppose that for all $t \in T$, $\pi(t) = 1/3$. Suppose that the listener's utility function is given by:

v	a_1	a_2	a_3
t_1	0	1	4
t_2	0	-2	-3
t_3	-1	0	-2

Figure 2.3. Listener's Utility Function

This listener utility function satisfies the single-crossing condition described above.

Let us now restrict attention to deterministic persuasion rules. It is easy to verify that the unique optimal deterministic persuasion rule is given by f^* with $f^*(m_1) = a_1$, $f^*(m_2) = a_2$. On the other hand, this rule is not credible because upon seeing m_1, the listener would be better off taking action a_3.

It then follows from Theorem 2.3 that there does not exists a strictly increasing function $r : \mathbb{R} \to \mathbb{R}$ and a family of concave functions $\{c_t : \mathbb{R} \to \mathbb{R}\}_{t \in T}$ such that for all $i \in \{1, 2, 3\}$ and $t \in T$, $c_t(r(i)) = v(a_i, t)$. To see this directly, without appeal to Theorem 2.3, assume for contradiction that such r and c_t exist. Then considering type t_2, we have:

$$\frac{-2}{r(2)-r(1)} = \frac{c_{t_2}(r(2)) - c_{t_2}(r(1))}{r(2)-r(1)} \geq \frac{c_{t_2}(r(3)) - c_{t_2}(r(2))}{r(3)-r(2)} = \frac{-1}{r(3)-r(2)}$$

or equivalently,

$$\frac{r(3)-r(2)}{r(2)-r(1)} \leq \frac{1}{2}.$$

On the other hand, considering t_1, we have:

$$\frac{1}{r(2)-r(1)} = \frac{c_{t_1}(r(2)) - c_{t_1}(r(1))}{r(2)-r(1)} \geq \frac{c_{t_1}(r(3)) - c_{t_1}(r(2))}{r(3)-r(2)} = \frac{3}{r(3)-r(2)},$$

which implies that

$$\frac{r(3)-r(2)}{r(2)-r(1)} \geq 3,$$

a contradiction. $\square$

2.3. Proofs

2.3.1. Proof of Theorem 2.1

Consider an optimal persuasion rule f^*, and suppose that the speaker responds with a strategy σ^* which maximizes the utility of the listener among the speaker's best responses to f^*. We may assume wlog that for all $t \in T$, there exists $m^t \in M(t)$ such that $\sigma^*(t, m^t) = 1$, because all messages which the speaker is using with positive probability must give both the speaker and the listener the same expected utility. Notice that it may be that $t \neq t'$, but $m^t = m^{t'}$. Let

$$\mathbf{M}^* := \{m \in \mathbf{M} : \exists t \in T, m = m^t\}.$$

Let F^* be the set of persuasion rules which assign each $m \notin \mathbf{M}^*$ the lowest action with probability 1. We may assume wlog that $f^* \in F^*$.

Let us write $\mathbf{M}^* = \{m_1, \ldots, m_n\}$ where the messages are enumerated in such a way that:

$$i < j \Rightarrow E_{f^*(m_i)}[u(a)] \leq E_{f^*(m_j)}[u(a)].$$

Note that any persuasion rule $f \in F^*$ can be associated with a vector $\alpha = (\alpha^i_j)^{i=1,\ldots,n}_{j=1,\ldots,k} \in [0,1]^{nk}$, where for all $m_i \in \mathbf{M}^*$, and $a_j \in \mathbf{A}$, α^i_j is the probability that $f(m_i)$ selects action a_j. Likewise, for any $\alpha \in [0,1]^{kn}$ with $\sum_j \alpha^i_j = 1$ for all i, let f_α be the corresponding persuasion rule.

It is possible that there exist two distinct messages m_{i_1} and m_{i_2} in $\mathbf{M}^*$ such that $E_{f^*(m_{i_1})}[u(a)] = E_{f^*(m_{i_2})}[u(a)]$. Group messages in $\mathbf{M}^*$–or equivalently, their indices–together into *blocks* B_ℓ such that two messages belong to the same block if and only if they give the speaker the same expected utility according to f^*. Suppose that the blocks $\{B_1, \ldots, B_p\}$ are enumerated so that types in smaller blocks receive a lower expected utility than types in higher blocks. Next, for $i = 1, \ldots, n$, define:

$$c^i_j = \sum_{t \in T_i} v(a_j, t)\pi(t),$$

where $T_i = \{t \in T : m^t = m_i\}$. Now consider the following linear program:

$$\max_{\alpha \in \mathbb{R}^{nk}} \sum_{i=1}^{n} \sum_{j=1}^{k} \alpha^i_j c^i_j \qquad \text{s.t.} \tag{2.8}$$

$$\sum_{j=1}^{k} u(j)\alpha^{i_1}_j = \sum_{j=1}^{k} u(j)\alpha^{i_2}_j \quad \text{if} \quad \exists \ell, \text{ s.t. } i_1, i_2 \in B_\ell \tag{2.9}$$

$$\sum_{j=1}^{k} u(j)\alpha^{i_1}_j \leq \sum_{j=1}^{k} u(j)\alpha^{i_2}_j \quad \text{if} \quad \exists \ell_1, \ell_2 \text{ s.t. } \ell_1 < \ell_2, i_1 \in B_{\ell_1}, i_2 \in B_{\ell_2} \tag{2.10}$$

$$\sum_{j=1}^{k} \alpha^i_j = 1 \qquad \forall i = 1, \ldots, n \tag{2.11}$$

$$\alpha^i_j \geq 0 \qquad \forall i = 1, \ldots, n, \forall j = 1, \ldots, k. \tag{2.12}$$

Let P be the polytope defined by constraints (2.9)-(2.12). (Note that P is nonempty because it contains some α^* corresponding to f^*. Also P is bounded as it is contained in the hypercube $[0, 1]^{nk}$.)

The fact that playing m^t with probability 1 was a best response to f^* along with the way that the blocks were defined implies that for any type t, and $\alpha \in P$, playing m^t with probability 1 is a best response to f_α for t. Moreover, $\sum_{i=1}^n \sum_{j=1}^k \alpha^i_j c^i_j$ is the listener's expected utility when he plays f_α and each type t of speaker responds by playing m_t. Optimality of f^* and σ^*, along with the facts that $\sigma^*(t, m^t) = 1$ for all $t \in T$ and $f^* = f_{\alpha^*}$

for some $\alpha^* \in P$ imply that any α which solves (2.8) is such that f_α is an optimal persuasion rule. Given this consideration, the theorem is implied by the following lemma.

Lemma 2.1. *If β is an extreme point of P, then for all $i = 1, \ldots, n$*

$$\sum_{j=1}^{k} u(a_j)\beta_j^i \in \{u(a_1), \ldots, u(a_k)\}. \tag{2.13}$$

$$|\{j : \beta_j^i > 0\}| \leq 2. \tag{2.14}$$

Notice in particular that (2.13) and (2.14) together imply that for any extreme point β of P, if for some i, j_1, and j_2, $\beta_{j_1}^i > 0$ and $\beta_{j_2}^i > 0$, then $j_2 \neq j_1 + 1$.

Proof of Lemma 2.1. I choose $\beta \in P$, and argue first that if β does not satisfy (2.13) for some i, then β is not an extreme point of P. First, notice that by construction, larger blocks contain types with larger indices. It then follows from constraints (2.9) and (2.10) that:

$$i_1 < i_2 \Rightarrow \sum_{j=1}^{k} u(a_j)\beta_j^{i_1} \leq \sum_{j=1}^{k} u(a_j)\beta_j^{i_2}.$$

Separate indices i into blocks $\{\widehat{B}_1, \ldots, \widehat{B}_p\}$ such that i_1 and i_2 are in the same block if $\sum_{j=1}^{k} u(a_j)\beta_j^{i_1} = \sum_{j=1}^{k} u(a_j)\beta_j^{i_2}$, and i_1 is in a smaller block (i.e., block with a smaller index) than i_2 if $\sum_{j=1}^{k} u(a_j)\beta_j^{i_1} < \sum_{j=1}^{k} u(a_j)\beta_j^{i_2}$. Notice that the blocks so constructed are (weakly) coarser than the initial blocks $\{B_1, \ldots, B_\ell\}$, and so also $p \leq \ell$. If β does not satisfy (2.13) for all i, then there must be some i_0 such that $\sum_{j=1}^{n} u(a_j)\beta_j^{i_0} = q \notin \{u(a_1), \ldots, u(a_k)\}$. Clearly $u(a_1) < q < u(a_k)$. i_0 belongs to some block $\widehat{B}_h$ and for all $i \in \widehat{B}_h$, $\sum_{j=1}^{k} u(a_j)\beta_j^i = q$. It follows that for all $i \in \widehat{B}_h$ and all $j = 1, \ldots, k$, $\beta_j^i \neq 1$, and moreover for each $i \in \widehat{B}_h$, there must exist at least two indices j_1, j_2, such that $\beta_{j_1}^i, \beta_{j_2}^i \neq 0$. In fact, for each $i \in \widehat{B}_h$, define $\underline{j}(i)$ to be the smallest index j such that $\beta_j^i \neq 0$, and $\overline{j}(i)$

to be the largest index j such that $\beta^i_j \neq 0$. By what was just argued $\underline{j}(i) < \overline{j}(i)$ for all $i \in \widehat{B}_h$. Define:

$$x = \begin{cases} \sum_{j=1}^{k} u(a_j)\beta^i_j, & \text{for some } i \in \widehat{B}_{h-1} \text{ if } h > 1; \\ u(1), & \text{if } h = 1. \end{cases}$$

$$y = \begin{cases} \sum_{j=1}^{k} u(a_j)\beta^i_j, & \text{for some } i \in \widehat{B}_{h+1} \text{ if } h < p; \\ u(n), & \text{if } h = p. \end{cases}$$

Notice that $x < q < y$. Choose some $\epsilon_{i_0} > 0$ such that:

$$[u(a_{\overline{j}(i_0)}) - u(a_{\underline{j}(i_0)})]\epsilon_{i_0} < \min\{q - x, y - q\} \tag{2.15}$$

Now I define two vectors $\underline{\beta}, \overline{\beta} \in \mathbb{R}^{nk}$. For all $i \notin \widehat{B}_p$, and all $j = 1, \ldots, k$, define $\underline{\beta}^i_j = \overline{\beta}^i_j = \beta^i_j$. For each $i \in \widehat{B}_h \setminus \{i_0\}$, define ϵ_i to the be solution to:

$$[u(a_{\overline{j}(i)}) - u(a_{\underline{j}(i)})]\epsilon_i = [u(a_{\overline{j}(i_0)}) - u(a_{\underline{j}(i_0)})]\epsilon_{i_0} \tag{2.16}$$

Notice that $\epsilon_i > 0$. For each $i \in \widehat{B}_h$ (including i_0), if $j \notin \{\underline{j}(i), \overline{j}(i)\}$, define $\underline{\beta}^i_j = \overline{\beta}^i_j = \beta^i_j$. Again for $i \in \widehat{B}_h$ (including i_0), define $\overline{\beta}^i_{\underline{j}(i)} = \beta^i_{\underline{j}(i)} - \epsilon_i, \overline{\beta}^i_{\overline{j}(i)} = \beta^i_{\overline{j}(i)} + \epsilon_i, \underline{\beta}^i_{\underline{j}(i)} = \beta^i_{\underline{j}(i)} + \epsilon_i, \underline{\beta}^i_{\overline{j}(i)} = \beta^i_{\overline{j}(i)} - \epsilon_i$. If some component of $\overline{\beta}$ or $\underline{\beta}$ is not strictly between 0 and 1, replace all the ϵ_i by $\gamma\epsilon_i$ for some sufficiently small $\gamma > 0$, so that all components of $\overline{\beta}$ and $\underline{\beta}$ are strictly between 0 and 1. Notice that this would not violate (2.15) or (2.16). It now follows that $\underline{\beta}$ and $\overline{\beta}$ satisfy all constraints of the form (2.11) and (2.12). (2.16) implies that $\underline{\beta}$ and $\overline{\beta}$ satisfy all constraints of the form (2.9). Moreover if the block $\widehat{B}_h$ contains several blocks in $\{B_1, \ldots, B_\ell\}$, then (2.16) also implies that the corresponding constraints of the form (2.10) are satisfied. Finally, that the rest of the constraints of the form (2.10) are satisfied follows from (2.15). Thus $\underline{\beta}$ and $\overline{\beta}$ both belong to P. On the other hand, by construction, $\frac{1}{2}\underline{\beta} + \frac{1}{2}\overline{\beta} = \beta$. It follows that β is not an extreme point of P. This establishes that any extreme point of P satisfies (2.13).

Next, assume that $\beta \in P$ does not satisfy (2.14), or in other words, for some $i_0 = 1, \ldots, k$, there exist $j_1 < j_2 < j_3$ such that for $j \in \{j_1, j_2, j_3\}$, $\beta_j^{i_0} > 0$. For $j' \in \{j_1, j_2, j_3\}$, define $\widehat{\beta}_{j'} := \beta_{j'}^{i_0} / \sum_{j \in \{j_1, j_2, j_3\}} \beta_j^{i_0}$. So for $\widehat{\beta}_j \in [0,1]$ and $\sum_{j \in \{j_1, j_2, j_3\}} \widehat{\beta}_j = 1$. Define

$$z := \widehat{\beta}_{j_1} u(a_{j_1}) + \widehat{\beta}_{j_2} u(a_{j_2}) + \widehat{\beta}_{j_3} u(a_{j_3}). \tag{2.17}$$

First, I consider the possibility that $z = u(a_{j_2})$. Then define $\delta, \gamma \in \mathbb{R}^{nk}$ such that whenever either $i \neq i_0$, or $j \notin \{j_1, j_2, j_3\}$, $\gamma_j^i = \delta_j^i = \beta_j^i$. On the other hand, $\gamma_{j_2}^{i_0} = \sum_{j \in \{j_1, j_2, j_3\}} \beta_j^{i_0}$, and for $j \in \{j_1, j_3\}$, $\gamma_j^{i_0} = 0$, $\delta_{j_2}^{i_0} = 0$, and for $j \in \{j_1, j_3\}$, $\delta_j^{i_0} = \beta_j^{i_0} / (1 - \widehat{\beta}_{j_2})$. Then notice that for all i:

$$\sum_{j=1}^{k} u(a_j) \beta_j^i = \sum_{j=1}^{k} u(a_j) \gamma_j^i = \sum_{j=1}^{k} u(a_j) \delta_j^i \tag{2.18}$$

It then follows from the fact that $\beta \in P$, that both γ and δ belong to P. Notice next that $\beta = \widehat{\beta}_{j_2} \gamma + (1 - \widehat{\beta}_{j_2}) \delta$, which implies that β is not an extreme point of P.

Next, consider the possibility that $z \neq u(a_{j_2})$, and assume moreover that $u(a_{j_2}) < z$. (The case in which $z < u(a_{j_2})$ is similar.) Define $\zeta, \eta \in [0,1]$ by the following equations:

$$\zeta u(a_{j_1}) + (1 - \zeta) u(a_{j_3}) = z \tag{2.19}$$

$$\eta u(a_{j_2}) + (1 - \eta) u(a_{j_3}) = z \tag{2.20}$$

Now define $\gamma, \delta \in \mathbb{R}^{kn}$ such that whenever either $i \neq i_0$ or $j \notin \{j_1, j_2, j_3\}$, $\beta^i_j = \gamma^i_j = \delta^i_j$. Moreover, $\gamma^{i_0}_{j_1} = (\sum_{j\in\{j_1,j_2,j_3\}} \beta^{i_0}_j)\zeta, \gamma^{i_0}_{j_2} = 0, \gamma^{i_0}_{j_3} = (\sum_{j\in\{j_1,j_2,j_3\}} \beta^{i_0}_j)(1-\zeta), \delta^{i_0}_{j_1} = 0, \delta^{i_0}_{j_2} = (\sum_{j\in\{j_1,j_2,j_3\}} \beta^{i_0}_j)\eta, \delta^{i_0}_{j_3} = (\sum_{j\in\{j_1,j_2,j_3\}} \beta^{i_0}_j)(1-\eta)$. Notice that this implies (2.18) holds in this case as well, which implies that since β belongs to P, so do γ and δ. Notice next that it must be the case that $\eta > \widehat{\beta}_{j_1}$. So define $\theta = \widehat{\beta}_{j_1}/\zeta \in [0,1]$. It then follows from (2.19) and (2.20) that

$$\theta\zeta u(a_{j_1}) + (1-\theta)\eta u(a_{j_2}) + [\theta(1-\zeta) + (1-\theta)(1-\eta)]u(a_{j_3}) = z,$$

where $\theta\zeta = \widehat{\beta}_{j_1}$. On the other hand, notice that:

$$1-\theta = 1 - \frac{\widehat{\beta}_{j_1}}{\zeta} = \frac{\zeta - \widehat{\beta}_{j_1}}{\zeta} = \frac{\frac{u(a_{j_3})-z}{u(a_{j_3})-u(a_{j_1})} - \widehat{\beta}_{j_1}}{\frac{u(a_{j_3})-z}{u(a_{j_3})-u(a_{j_1})}} = \frac{u(a_{j_3}) - z - \beta_{j_1}(u(a_{j_3}) - u(a_{j_1}))}{u(a_{j_3}) - z}$$

$$= \frac{u(a_{j_3}) - \widehat{\beta}_{j_1}u(a_{j_1}) - \widehat{\beta}_{j_2}u(a_{j_2}) - \widehat{\beta}_{j_3}u(a_{j_3}) - \widehat{\beta}_{j_1}(u(a_{j_3}) - u(a_{j_1}))}{u(a_{j_3}) - z}$$

$$= \frac{(\widehat{\beta}_{j_1} + \widehat{\beta}_{j_2})u(a_{j_3}) - \widehat{\beta}_{j_1}u(a_{j_1}) - \widehat{\beta}_{j_2}u(a_{j_2}) - \widehat{\beta}_{j_1}(u(a_{j_3}) - u(a_{j_1}))}{u(a_{j_3}) - z}$$

$$= \frac{\widehat{\beta}_{j_2}(u(a_{j_3}) - u(a_{j_2}))}{u(a_{j_3}) - z} = \frac{\widehat{\beta}_{j_2}}{\frac{u(a_{j_3})-z}{u(a_{j_3})-u(a_{j_2})}} = \frac{\widehat{\beta}_{j_2}}{\eta},$$

where (2.17) was invoked during the course of the definition. It follows that $(1-\theta)\eta = \widehat{\beta}_{j_2}$. Recalling that $\theta\zeta = \widehat{\beta}_{j_1}$, it follows from (2.17) that $\theta(1-\zeta) + (1-\theta)(1-\eta) = \widehat{\beta}_{j_3}$. It follows from these considerations that $\beta = \theta\gamma + (1-\theta)\delta$, implying that β is not an extreme point of P. $\square$

2.3.2. Proof of Theorem 2.3

2.3.2.1. Preliminaries. Here, I will present several definitions which will be useful in the course of the proof.

For each type t, let $\underline{\jmath}(t)$ (resp., $\bar{\jmath}(t)$) be the smallest (resp., greatest) index of an action in $\text{argmax}_{a\in\mathbf{A}}v(a,t)$.

Next, for any any deterministic persuasion rule f, and $1 \leq k \leq \ell$, define:

$$\begin{aligned}
M(f,\ell) &:= \{m \in \mathbf{M}, f(m) = a_\ell\} \\
T(f,\ell) &:= \{t \in T : g_f(t) = a_\ell\} \\
\Sigma(f,\ell) &:= \{\sigma \in \Sigma : \forall m \in M(f,\ell) : \sigma(t,m) > 0 \Rightarrow g_f(t) = a_\ell\}
\end{aligned}$$

$M(f,\ell)$ is that set of messages which lead to action a_ℓ under persuasion rule f. $T(f,\ell)$ is the set of types who receive action a_ℓ when best responding to persuasion rule f. $\Sigma(f,\ell)$ is the set of speaker strategies in which any type t who would get action a_ℓ in a best response to f is best responding. Clearly, $\Sigma(f,\ell)$ is nonempty.

For $1 < \ell < k$, $\sigma \in \Sigma, m \in \mathbf{M}$, and deterministic persuasion rule f, define:

$$\begin{aligned}
V_\ell(m,\sigma) &:= \sum_{t\in T}[v(a_\ell,t) - v(a_{\ell-1},t)]\sigma(t,m)\pi(t) \\
W_\ell(m,\sigma) &:= \sum_{t\in T}[v(a_{\ell+1},t) - v(a_\ell,t)]\sigma(t,m)\pi(t)
\end{aligned}$$

and define:

$$\begin{aligned}
\overline{N}_{W_\ell}(f,\ell,\sigma) &:= \{m \in M(f,\ell) : W_\ell(m,\sigma) > 0\} \\
N^*_{W_\ell}(f,\ell,\sigma) &:= \{m \in M(f,\ell) : W_\ell(m,\sigma) = 0\} \\
\underline{N}_{W_\ell}(f,\ell,\sigma) &:= \{m \in M(f,\ell) : W_\ell(m,\sigma) < 0\}
\end{aligned}$$

Let $\overline{N}_{V_\ell}(f,\ell,\sigma), N^*_{V_\ell}(f,\ell,\sigma)$, and $\underline{N}_{V_\ell}(f,\ell,\sigma)$ be defined similarly, except that the function V_ℓ takes the place of W_ℓ in the definitions above. Notice that if $\sum_{t\in T}\sigma(t,m) = 0$, then m belongs to both $N^*_{V_\ell}(f,\ell,\sigma)$ and $N^*_{W_\ell}(f,\ell,\sigma)$.

For any $\sigma \in \Sigma, t^* \in T, m_1, m_2 \in M(t)$, and $\epsilon \in (0, \sigma(t,m_1))$, define:

$$\sigma^{(t^*,m_1,m_2,\epsilon)} := \begin{cases} \sigma(t,m) - \epsilon, & \text{if } (t,m) = (t^*,m_1)\ ; \\ \sigma(t,m) + \epsilon, & \text{if } (t,m) = (t^*,m_2); \\ \sigma(t,m), & \text{otherwise.} \end{cases}$$

Thus $\sigma^{(t,m_1,m_2,\epsilon)}$ is a strategy which–conditional on type t–shifts ϵ probability mass from m_1 to m_2, and is defined provided that such a redistribution is possible.

We may assume without loss of generality that for all $m \in \mathbf{M}$, $f(m) \notin \{a_1, a_k\}$, because if not, it is possible to add actions a_0 and a_{k+1} such that the listener's utility to a_0 and a_{k+1} is so low for every type of speaker that these actions would never be used in an optimal rule. This can be done in such a way that $v(a_i,t)$ will still satisfy the assumptions of the theorem, and the optimal rules with and without the actions a_0 and a_{k+1} will coincide.

2.3.2.2. Step 1. I will now show that there exists $\sigma \in \Sigma(f,\ell)$ such that for all $m \in M(f,\ell), W_\ell(m,\sigma) \leq 0$. To this end, consider the program:

$$\begin{aligned}
&\min_{\sigma \in [0,1]^{T(f,\ell)\times M(f,\ell)}} \sum_{m\in M(f,\ell)} [\max\{0, \sum_{t\in T(f,\ell)} \{[v(a_{\ell+1},t) - v(a_{\ell,t})]\pi(t)\}\sigma(t,m)\}] \\
\text{s.t.}\quad & m \notin M(t) \Rightarrow \sigma(t,m) = 0 \\
& \forall t \in T, \sum_{m\in M(t)\cap M(f,\ell)} \sigma(t,m) = 1.
\end{aligned}$$

This is the minimization of a continuous function over a compact set. Therefore, it takes on a minimum value. The problem can equivalently be written as:

$$Y^* = \min_{\sigma\in\Sigma(f,\ell)} I(\sigma)$$

where

$$I(\sigma) := \sum_{m\in M(f,\ell):W_\ell(\sigma,m)>0} W_\ell(\sigma,m)$$

Clearly $Y^* \geq 0$. I would like to argue that $Y^* = 0$. This would establish that there is a speaker best response σ to f such that upon seeing a message m played with positive probability by σ, such that $f(m) = a_\ell$, the listener would never have an incentive to select action $a_{\ell+1}$ instead.

By continuity of the objective and compactness of the constraint set, in order to show that $Y^* = 0$, it is sufficient to argue that for any $\sigma \in \Sigma(f,\ell)$, if $I(\sigma) > 0$, then there exists $\sigma' \in \Sigma(f,\ell)$ such that $I(\sigma') < I(\sigma)$.

Suppose that

i: there exists $t^* \in T$ with $\bar{\jmath}(t^*) \leq \ell$, such that for some $m_1 \in \underline{N}_{W_\ell}(f, \ell, \sigma)$, $\sigma(t^*, m_1) > 0$, and there exists $m_2 \in M(t^*) \cap N^*_{W_\ell}(f, \ell, \sigma)$.

Then for some small $\epsilon > 0$, define a new strategy $\sigma_0 := \sigma^{(t^*, m_1, m_2, \epsilon)}$. Notice that if ϵ is chosen sufficiently small, then $W_\ell(\sigma_0, m_1) < 0$. Since $\bar{\jmath}(t^*) < \ell$, and by (*),

$$v(a_{\ell+1}, t^*) - v(a_\ell, t^*) < 0.$$

Recall also that we assume that $\pi(t^*) > 0$. So:

$$W_\ell(\sigma_0, m_2) = W_\ell(\sigma, m_2) + \epsilon[v(a_{\ell+1}, t^*) - v(a_\ell, t^*)]\pi(t^*) < W_\ell(\sigma, m_2) = 0.$$

Notice that $I(\sigma_0) = I(\sigma)$. This procedure is repeated until **i** is no longer true. Let σ_1 be the strategy that results. Notice that $I(\sigma_1) = I(\sigma)$.

Next, assume that

ii: there exists $t^* \in T$ with $\bar{\jmath}(t^*) \leq \ell$, such that for some $m_1 \in \underline{N}_{W_\ell}(f, \ell, \sigma_1)$, $\sigma_1(t^*, m_1) > 0$, and there exists $m_2 \in M(t^*) \cap \overline{N}_{W_\ell}(f', \ell, \sigma_1)$.

Then, define a new strategy $\sigma_2 := \sigma_1^{(t^*, m_1, m_2, \epsilon)}$. Notice that if ϵ is chosen sufficiently small, σ_2 is well-defined, and $W_\ell(\sigma_2, m_1) < 0$ and $W_\ell(\sigma_2, m_2) > 0$.

Since $\jmath(t^*) \leq \ell$, $v(a_{\ell+1}, t^*) - v(a_\ell, t^*) < 0$. So since, $\pi(t^*) > 0$,

$$\begin{aligned} I(\sigma_2) &= I(\sigma_1) + [W_\ell(\sigma_2, m_2) - W_\ell(\sigma_1, m_2)] \\ &= I(\sigma_3) + \epsilon[v(a_{\ell+1}, t) - v(a_\ell, t)] < I(\sigma_3) = I(\sigma). \end{aligned}$$

So we have found a strategy $\sigma_2 \in \Sigma(f, \ell)$ such that $I(\sigma_2) < I(\sigma)$, as desired.

So we derive the conclusion we want if **ii** is true. On the other hand, assume for contradiction that **ii** is not true. Then consider f' defined by:

$$f'(m) := \begin{cases} a_{\ell+1}, & \text{if } m \in O; \\ f(m), & \text{otherwise;} \end{cases}$$

where $O := \overline{N}_{W_\ell}(f, \ell, \sigma_1) \cup N^*_{W_\ell}(f, \ell, \sigma_1)$. Then define:

$$S := \{t \in T : g_f(t) = a_\ell, \exists m \in M(t) \cap O\},$$

Then for each $t \in S$, define:

$$\alpha(t) := \sum_{m \in \underline{N}_{W_\ell}(f, \ell, \sigma_1)} \sigma(t, m)$$

Then notice that the payoff to the listener given rule f' conditional on the event $\{t \in S\}$ is:

$$\frac{1}{\pi(S)}[\sum_{m \in O} \sum_{t \in T} v(a_{\ell+1}, t)\sigma_1(t, m)\pi(t) + \sum_{t \in S} v(a_{\ell+1}, t)\pi(t)\alpha(t)]$$

On the other hand the payoff to f conditional on $\{t \in S\}$ is:

$$\frac{1}{\pi(S)}[\sum_{m \in O} \sum_{t \in T} v(a_\ell, t)\sigma_1(t, m)\pi(t) + \sum_{t \in S} v(a_\ell, t)\pi(t)\alpha(t)]$$

Notice that by the definition of O:

$$\sum_{m \in O} \sum_{t \in T} v(a_{\ell+1}, t)\sigma_1(t, m)\pi(t) - \sum_{m \in O} \sum_{t \in T} v(a_\ell, t)\sigma_1(t, m)\pi(t) > 0,$$

where the strict inequality follows from the fact that $I(\sigma_1) > 0$, so the probability that a message in $\overline{N}_{W_\ell}(f, \ell, \sigma_1)$ is used is nonzero. On the other hand, recall that σ_1 does not satisfy **i** or **ii**. It follows that if $t \in S$, $\alpha(t) > 0$, it is not possible that $\bar{\jmath}(t) \leq \ell$. So it must be the case that if $\alpha(t) > 0$, then $\bar{\jmath}(t) > \ell$, but this means that:

$$\sum_{t \in S} v(a_{\ell+1}, t)\pi(t)\alpha(t) - \sum_{t \in S} v(a_\ell, t)\pi(t)\alpha(t) \geq 0.$$

So–recalling that every best speaker best response to a deterministic persuasion rule gives the listener the same expected utility–conditional on $\{t \in S\}$ f' attains a higher utility for the listener than f. On the other hand, conditional on $\{t \notin S\}$, f' and f attain the same utility. So f' attains a higher utility than f a contradiction.

2.3.2.3. Step 2. In this step, I will show that there exists $\sigma \in \Sigma(f, \ell)$ such that for all $m \in M(f, \ell), W_\ell(m, \sigma) \leq 0$ and $V_\ell(m, \sigma) \geq 0$.

It follows from the above argument that there exists $\sigma \in \Sigma(f, \ell)$ such that $I(\sigma) = 0$, which means that for all $m \in M(f, \ell), W_\ell(m, \sigma) \leq 0$. Define:

$$\Sigma^*(f, \ell) := \{\sigma \in \Sigma(f, \ell) : \forall m \in M(f, \ell), W_\ell(m, \sigma) \leq 0\}$$

Then $\Sigma^*(f, \ell) \neq \emptyset$ by the above argument. Now consider the following program:

$$\begin{aligned}
&\max_{\sigma \in [0,1]^{T(f,\ell) \times M(f,\ell)}} \sum_{m \in M(f,\ell)} [\min\{0, \sum_{t \in T(f,\ell)} \{[v(a_\ell, t) - v(a_{\ell-1}, t)]\pi(t)\}\sigma(t, m)\}] \\
\text{s.t.} \quad & m \notin M(t) \Rightarrow \sigma(t, m) = 0 \\
& \forall t \in T(f, \ell), \sum_{m \in M(t) \cap M(f,\ell)} \sigma(t, m) = 1 \\
& \forall m \in M(f, \ell), \sum_{t \in T(f,\ell)} [v(a_{\ell+1}, t) - v(a_\ell, t)]\pi(t)\sigma(t, m) \leq 0.
\end{aligned}$$

Notice that this is the maximization of a continuous function on a nonempty compact set. (the above argument established non-emptiness). Therefore, it attains a maximum, and that maximum can be at most 0.

Notice that this can be rewritten as

$$X^* = \max_{\sigma \in \Sigma^*(f,\ell)} K(\sigma)$$

where

$$K(\sigma) = \sum_{m \in M(f,\ell): V_\ell(m,\sigma)<0} V_\ell(m, \sigma)$$

I want to show that $X^* = 0$. This would establish that there is a speaker best response σ to f such that upon seeing a message m played with positive probability by σ, such that $f(m) = a_\ell$, the listener would never have an incentive to select action $a_{\ell+1}$ or $a_{\ell-1}$ instead.

Since, as argued above $K(\sigma)$ attains a maximum on $\Sigma^*(f,\ell)$, in order to establish that $X^* = 0$, it is sufficient to argue that if $K(\sigma) < 0$, then it is possible to find $\sigma' \in \Sigma^*(f,\ell)$ with $K(\sigma') > K(\sigma)$.

So suppose that $K(\sigma) < 0$. Before proceeding, I must prove a lemma:

Lemma 2.2. *For all $\sigma' \in \Sigma(f,\ell)$ and all $m \in M(f,\ell)$:*

$$V_\ell(m, \sigma') < 0 \Rightarrow W_\ell(m, \sigma') < 0$$

Proof. Using (*), we have:

$$
\begin{aligned}
0 > V_\ell(m,\sigma') &= \sum_{t\in T}[v(a_\ell,t)-v(a_{\ell-1},t)]\sigma'(t,m)\pi(t)\\
&= \sum_{t\in T}[c_t(r(\ell))-c_t(r(\ell-1))]\sigma'(t,m)\pi(t)\\
&= (r(\ell)-r(\ell-1))\sum_{t\in T}\Big[\frac{c_t(r(\ell))-c_t(r(\ell-1))}{r(\ell)-r(\ell-1)}\Big]\sigma'(t,m)\pi(t)\\
&> (r(\ell)-r(\ell-1))\sum_{t\in T}\Big[\frac{c_t(r(\ell+1))-c_t(r(\ell))}{r(\ell+1)-r(\ell)}\Big]\sigma'(t,m)\pi(t) \qquad (2.21)\\
&= \frac{r(\ell)-r(\ell-1)}{r(\ell+1)-r(\ell)}\sum_{t\in T}[c_t(r(\ell+1))-c_t(r(\ell))]\sigma'(t,m)\pi(t)\\
&= \frac{r(\ell)-r(\ell-1)}{r(\ell+1)-r(\ell)}W_\ell(m,\sigma'),
\end{aligned}
$$

where the key inequality (2.21) follows from the concavity of c_t, and the fact that r is strictly increasing. Again, since r is a strictly increasing function, the above implies that $W_\ell(m,\sigma') < 0$. $\Box$

Now assume that:

iii: There exists $t^* \in T$ with $\bar{\jmath}(t) < \ell$ and $m_1 \in \underline{N}_{V_\ell}(f,\ell,\sigma)$ such that $\sigma(t^*,m_1) > 0$ and and there exists $m_2 \in M(t) \cap N^*_{V_\ell}(f,\ell,\sigma)$.

The consider the strategy $\sigma_0 := \sigma^{(t,m_1,m_2,\epsilon)}$. Notice that $V_\ell(m_1,\sigma) < 0$, and by Lemma 2.2 $W_\ell(m_1,\sigma) < 0$. If ϵ is chosen sufficiently small, then $V_\ell(m_1,\sigma_0) < 0$ and $W_\ell(m_1,\sigma_0) < 0$. Notice moreover that $V_\ell(m_2,\sigma) = 0, W_\ell(m_2,\sigma) \le 0$, and by (*), $V_\ell(m_2,\sigma_0) < 0$, and so by Lemma 2.2, $W_\ell(m_2,\sigma) < 0$. Finally, notice that $K(\sigma_0) = K(\sigma)$ and $\sigma_0 \in \Sigma^*(f,\ell)$. Now iterate the process until it can no longer be iterated, arriving at a strategy $\sigma_1 \in \Sigma^*(f,\ell)$ for which **iii** does not hold. Note that $K(\sigma) = K(\sigma_1)$.

Next suppose that:

iv: There exists $t^* \in T$ with $\bar{\jmath}(t^*) < \ell$ and $m_1 \in \underline{N}_{V_\ell}(f, \ell, \sigma_1)$ and there exists $m_2 \in M(t) \cap \overline{N}_{V_\ell}(f, \ell, \sigma_1)$.

For some small $\epsilon > 0$, define $\sigma_2 = \sigma_1^{(t,m_1,m_2,\epsilon)}$. Notice that if ϵ is chosen sufficiently small, then $V_\ell(m_1, \sigma_2) < 0$, and so by Lemma 2.2, $W_\ell(m_1, \sigma_2) < 0$. Also if ϵ is chosen sufficiently small, then $V_\ell(m_2, \sigma_2) > 0, W_\ell(m_2, \sigma_2) < W_\ell(m_2, \sigma_1) \leq 0$. So $\sigma_2 \in \Sigma^*(f, \ell)$. On the other hand $V_\ell(m_1, \sigma_2) > V_\ell(m_1, \sigma_1)$, so $K(\sigma_2) > K(\sigma_1)$, as desired.

So, we derive the conclusion we want if **iv** is true. Now, assume for contradiction that **iv** is false. Note that since $\sigma_1 \in \Sigma^*(f, \ell)$, σ_1 is a best response to f for all types $t \in T(f, \ell)$. We may also assume that σ_1 is a best response for all types $t \notin T(f, \ell)$. Now, consider the strategy f'' defined by:

$$f''(m) = \begin{cases} a_{\ell-1}, & \text{if } m \in \underline{N}_{V_\ell}(f, \ell, \sigma_1) \text{ ;} \\ f(m), & \text{otherwise.} \end{cases}$$

Now consider a speaker strategy σ'', such that for all $t \notin T(f, \ell)$ and all $m \in \mathbf{M}$, $\sigma''(t, m) = \sigma_1(t, m)$. Next define

$$Z := \{t \in T(f, \ell) : M(t) \cap M(f, \ell) \subseteq \underline{N}_{V_\ell}(f, \ell, \sigma)\}.$$

If $t \in Z$, then for all $m \in \mathbf{M}$, $\sigma''(t, m) = \sigma_1(t, m)$. Finally, if $t \in T(f, \ell) \setminus Z$, then

$$\sum_{m \in \overline{N}_{V_\ell}(f, \ell, \sigma_1)} \sigma''(t, m) = 1.$$

It is easy to see that σ'' is a best response to f''. Notice that the listener's payoff to f'' conditional on the event $\{t \in Z\}$ is:

(2.22)

$$\sum_{\{m \in \underline{N}_{V_\ell}(f,\ell,\sigma)\}} \sum_{t \in T} v(a_{\ell-1}, t)\sigma_1(t,m)\pi(t) - \sum_{\{t \in T(f,\ell) \setminus Z\}} \sum_{\{m \in \underline{N}_{V_\ell}(f,\ell,\sigma)\}} v(a_{\ell-1}, t)\sigma_1(t,m)\pi(t)$$

In contrast, the listener's payoff to f conditional on $\{t \in Z\}$ is:

$$\sum_{\{m \in \underline{N}_{V_\ell}(f,\ell,\sigma)\}} \sum_{t \in T} v(a_\ell, t)\sigma_1(t,m)\pi(t) - \sum_{\{t \in T(f,\ell) \setminus Z\}} \sum_{\{m \in \underline{N}_{V_\ell}(f,\ell,\sigma)\}} v(a_\ell, t)\sigma_1(t,m)\pi(t). \tag{2.23}$$

It follows from the definition of $\underline{N}_{V_\ell}(f, \ell, \sigma)$ that

$$\sum_{\{m \in \underline{N}_{V_\ell}(f,\ell,\sigma)\}} \sum_{t \in T} v(a_{\ell-1}, t)\sigma_1(t,m)\pi(t) > \sum_{\{m \in \underline{N}_{V_\ell}(f,\ell,\sigma)\}} \sum_{t \in T} v(a_\ell, t)\sigma_1(t,m)\pi(t), \tag{2.24}$$

where the inequality is strict because $K(\sigma_1) < 0$. Next notice that it follows from the definition of Z, the fact that **iii** is false, and the assumption that **iv** is false, that if $t \in T(f, \ell) \setminus Z$, then $\bar{\jmath}(t) \geq \ell$. But this implies that:

(2.25)

$$\sum_{\{t \in T(f,\ell) \setminus Z\}} \sum_{\{m \in \underline{N}_{V_\ell}(f,\ell,\sigma)\}} v(a_\ell, t)\sigma_1(t,m)\pi(t) \geq \sum_{\{t \in T(f,\ell) \setminus Z\}} \sum_{\{m \in \underline{N}_{V_\ell}(f,\ell,\sigma)\}} v(a_{\ell-1}, t)\sigma_1(t,m)\pi(t).$$

Recalling again that the listener's payoff given any best response to a deterministic persuasion rule is the same, equations (2.22)-(2.25) imply that the listener receives a higher payoff in response to persuasion rule f'' than to f, which contradicts the assumption that f is optimal among deterministic persuasion rules. So $K(\sigma) = K(\sigma_1) < 0$ implies that **iv** must be true, in which case, we derived the desired conclusion.

2.3.2.4. Step 3. In this step, I establish that there exists a speaker strategy σ which satisfies (2.5) and (2.6) relative to f.

Recall that we are assuming wlog that the is a persuasion rule f which is optimal among deterministic rules and such that for all $m \in \mathbf{M}$, $f(m) \notin \{a_1, a_k\}$. Steps 1 and 2 imply that there exists a speaker best response σ to f such that for all $m \in \mathbf{M}$ such that $f(m) = a_\ell$ with $1 < \ell < k$,

$$\begin{aligned}\sum_{t \in T}[v(a_\ell, t) - v(a_{\ell+1}, t)]\sigma(t, m)\pi(t) &\geq 0 \\ \sum_{t \in T}[v(a_\ell, t) - v(a_{\ell-1}, t)]\sigma(t, m)\pi(t) &\geq 0.\end{aligned}$$

In other words, upon seeing a message m which is sent with positive probability according to σ, it is never in the listener's interest to take either the next highest or the next lowest action, as opposed to the action dictated by the persuasion rule f.

Next consider a_j with $a_j > a_{\ell+1}$. Suppose that $j = \ell + h$. Notice that (*) implies that:

$$\frac{v(a_{\ell+1}, t) - v(a_\ell, t)}{r(\ell+1) - r(\ell)} \geq \frac{v(a_j, t) - v(a_\ell, t)}{r(j) - r(\ell)},$$

or equivalently,

$$\frac{r(j) - r(\ell)}{r(\ell+1) - r(\ell)}[v(a_{\ell+1}, t) - v(a_\ell, t)] \geq v(a_j, t) - v(a_{\ell+1}, t)$$

So invoking the fact that r is increasing:

$$0 \geq \frac{r(j) - r(\ell)}{r(\ell+1) - r(\ell)} \sum_{t \in T}[v(a_{\ell+1}, t) - v(a_\ell, t)]\sigma(t, m)\pi(t) \geq \sum_{t \in T}[v(a_j, t) - v(a_\ell, t)]\sigma(t, m)\pi(t).$$

So the listener is never better off taking any action which is higher than the one dictated by f when he sees a message m which is played with positive probability by σ.

On the other hand for $a_i < a_{\ell-1}$:

$$\frac{r(\ell) - r(i)}{r(\ell) - r(\ell-1)}[v(a_\ell, t) - v(a_{\ell-1}, t)] \leq v(a_\ell, t) - v(a_i, t)$$

So:

$$0 \leq \frac{r(\ell) - r(i)}{r(\ell) - r(\ell-1)} \sum_{t\in T}[v(a_\ell, t) - v(a_{\ell-1}, t)]\sigma(t, m)\pi(t) \leq \sum_{t\in T}[v(a_\ell, t) - v(a_i, t)]\sigma(t, m)\pi(t)$$

This implies that the listener is never better off taking any action which is lower than the one dictated by f when he sees a message which is played with positive probability by σ.

2.3.2.5. Step 4. Say that a persuasion rule f is *weakly credible* if there exists a speaker strategy σ which satisfies (2.5) and (2.6) relative to f. So far I have shown that every persuasion rule f which is optimal among deterministic rules is weakly credible. It remains to strengthen this to credibility. Notice that in so doing, the statement of the theorem is such that we are allowed to alter the persuasion rule as long as this does not alter the induced outcome function. This is achieved by the following lemma.

Lemma 2.3. *Assume that the listener's utility function satisfies (*), and let f be a weakly credible persuasion rule which is optimal among deterministic rules. Then, if the listener's utility function satisfies (*), then there exists a credible deterministic persuasion rule f' which implements the same outcome function as f.*

Proof. Since f is weakly credible, there exists a sender strategy σ such that for all $m \in \mathbf{M}$:

$$\sum_{t\in T}\sigma(t,m) > 0 \Rightarrow f(m) \in \text{argmax}_{a\in\mathbf{A}}\sum_{t\in T}v(a,t)\sigma(t,m)\pi(t)$$

For every $m \in \mathbf{M}$, define:

$$h(m) := \min\{a_j : \exists t \in T, m \in M(t) \text{ and } a_j \in \text{argmax}_{a\in\mathbf{A}}v(a,t).\}$$

Then define:

$$f'(m) := \begin{cases} f(m), & \text{if } \sum_{t\in T}\sigma(t,m) > 0 \text{ ;} \\ h(m), & \text{otherwise.} \end{cases}$$

For each $m \in \mathbf{M}$, let $\mu_m(t) = \sigma(t,m)$ if $\sum_{t\in T}\sigma(t,m) > 0$ and suppose that μ_m puts probability one on some type t with $h(m) \in \text{argmax}_{a\in\mathbf{A}}v(a,t)$ and $m \in M(t)$ otherwise. By the definition of h, μ_m is well defined for each $m \in \mathbf{M}$. Note that for all $m \in \mathbf{M}$,

$$f'(m) \in \text{argmax}_{a\in\mathbf{A}}\sum_{t\in T}v(a,t)\mu_m(t),$$

by the credibility of f and the definition of f'. To complete the proof, it is sufficient to show that σ is a best response to f'. Since σ is a best response to f, if σ is not a best response to f', there must be some type t^* and some message $m^* \in M(t^*)$ with $\sum_{t\in T}\sigma(t,m^*) = 0$, and such that $h(m^*) = f'(m^*) > g_f(t^*)$. Let $S \neq \emptyset$ be the set of all types t such that $m^* \in M(t)$ and $h(m^*) > g_f(t)$. It follows σ' defined by:

$$\sigma'(t,m) = \begin{cases} \sigma(t,m), & \text{if } t \notin S; \\ 1, & \text{if } t \in S \text{ and } m = m^*; \\ 0, & \text{otherwise.} \end{cases}$$

is a best response to the persuasion rule f'' defined by

$$f''(m) = \begin{cases} h(m^*), & \text{if } m = m^*; \\ f(m), & \text{otherwise.} \end{cases}$$

Notice that for all $t \in S$, $a_{\underline{\jmath}(t)} \geq h(m^*)$ by the definition of h. By (*), it then follows that f'' and σ' attain a higher utility for the listener than f and σ, a contradiction.[7] So σ is a best response to f', completing the proof. $\square$

[7]Recall that all best responses to f–which is deterministic–induce the same outcome function and hence the same utility for the listener.

CHAPTER 3

The Lattice Structure of the Persuasion Problem

The purpose of this chapter is to reveal an underlying lattice theoretic structure to the persuasion problem. In Section 2.1, I found conditions under which there exists an optimal deterministic rule, and therefore also deterministic outcome functions. In this chapter, I restrict attention to deterministic persuasion rules. The justification for this restriction can either be taken to be that the assumption that the conditions of Theorem 2.2 hold, namely that the listener's utility function is a concave transformation of the speaker's utility function at every state of the world, or else that the class of deterministic persuasion rules induces a natural structure on the problem. The lattice structure which is studied in this section will be seen to be useful in Chapter 5 for analyzing qualitative properties of the persuasion problem within a very broad class of message structures.

In the model described in Section 1.1, the listener's choice set is viewed as the set of persuasion rules. However, the listener's assessment of a persuasion rule is dependent on the message correspondence; what the listener really cares about is which outcome function will be implemented. The interaction of the message correspondence and the set of persuasion rules is not very transparent. It would be advantageous to represent the listener's choice set directly as the set of implementable outcome functions. In fact, once this is done, the persuasion rules and the message correspondence become redundant. It is possible to recover all important information about the message correspondence from the family of implementable outcome functions. In particular, one can recover what I will refer to as the *essential* messages from the set of implementable outcome functions; messages which are not essential cannot be recovered, but also have no effect on what is feasible for the listener given the essential messages. Essential messages are identified in terms of a certain lattice-theoretic notion, and therefore this approach–which eschews the message correspondence–enables the use of lattice theory in order to characterize the structure of messages.

In order to pursue the approach just described, I will study properties which are common to *all* persuasion situations. In particular, I will give a complete answer the question: when is a family of outcome functions *exactly* the family of implementable outcome functions for some message correspondence? This has several consequences. In particular, I provide a method for inferring what outcome functions are implementable for k actions from knowledge of what outcome functions are implementable with 2 actions. Moreover, I show that the family of implementable outcome functions has certain closure properties. In other words, I will show how choices are linked for the listener in the sense that having certain choices entails having certain other choices. Finally, the analysis allows me to pursue the fine-grained analysis of messages.

In contrast to Chapter 2, where T and $\mathbf{M}$ were assumed to be finite, it is assumed in this chapter that T and $\mathbf{M}$ are of arbitrary (finite or infinite) cardinality.[1] Notice that since in this chapter, we are concerned only with which deterministic outcome functions are implementable the probability measure π is irrelevant, and therefore it is unnecessary

[1]Of course, restricting attention to deterministic persuasion rules when T is infinite is not justified by Theorem 2.2. However, since the theory is very similar when T is infinite as when T is finite, it is worth developing in the more general setting. Moreover, the infinite case allows one to discuss some distinctions which are absent in the finite case.

to define a sigma algebra on T. The *definition* of implementation can be taken to be (1.1) regardless of the cardinality of T and $\mathbf{M}$.

The chapter is split into four parts. Section 3.1 will analyze the case of two actions, which will be used as a building block for the main results in Section 3.2 which apply for an arbitrary number of actions. Section 3.3 then studies the consequences of the previous sections for the structure of messages. Section 3.4 presents an example that demonstrates some of the main ideas introduced in the previous sections. Section 3.5 contains proofs not contained in the main body of the chapter.

3.1. 2 Actions

With two actions, imagine that the speaker makes a request of the listener. Then a_2 corresponds to *acceptance* of the request, and a_1 corresponds to *rejection.* Each persuasion rule can be represented by the set $Q = \{m \in \mathbf{M} : f(m) = a_2\}$ of messages it accepts. So there is a one-to-one correspondence between persuasion rules and subsets of $\mathbf{M}$. The following terminology will prove to be convenient: think of a subset Q of $\mathbf{M}$ as a *question* with the wording: "Can you make some statement in Q? If so, I will accept your request." A message $m \in Q$ is an *answer* to Q. To the question Q, there corresponds the *acceptance set* A_Q defined as:

$$A_Q = \{t \in T : M(t) \cap Q \neq \emptyset\}.$$

So A_Q is the set of types who have some answer to the question Q. As Q corresponds to a persuasion rule, A_Q corresponds to the outcome function implemented by Q.

If $A_{Q_1} \subseteq A_{Q_2}$ then Q_1 is a *more difficult* question than Q_2, in that fewer types can answer Q_1 than Q_2. A more difficult question is worse for the speaker ex ante, and also at least as bad ex post. Define

$$\mathcal{I} = \{A_Q : Q \subseteq \mathbf{M}\} \tag{3.1}$$

to be the set of all acceptance sets as questions vary. $\mathcal{I}$ corresponds to the family of implementable outcome functions when $k = 2$.

Notice that $T \in \mathcal{I}$ because $T = A_{\mathbf{M}}$, the acceptance set of the question which accepts all messages. Likewise $\emptyset \in \mathcal{I}$ because $\emptyset = A_{\emptyset}$, the acceptance set corresponding to the question which rejects all messages. The following is an important observation.

Observation 3.1. *$\mathcal{I}$ is closed under union.*

To see this notice that for any $\mathcal{J} \subseteq \mathcal{I}$, it is possible to write $\mathcal{J} = \{A_Q : Q \in \mathcal{Q}\}$ where $\mathcal{Q} \subseteq 2^{\mathbf{M}}$ is some family of questions. Then consider the question Q^* which accepts a message if and only if it answers some question in $\mathcal{Q}$, or in other words, $Q^* = \bigcup \mathcal{Q}$. Then $A_{Q^*} = \bigcup \{A_Q : Q \in \mathcal{Q}\}$.

Observation 3.1 provides a simple but abstract property. What is its significance in intuitive terms? As explained above, each question Q corresponds to a persuasion rule $f : \mathbf{M} \to \{a_1, a_2\}$. In general, such persuasion rules may be simple or complex, in that they may depend on any property of messages. Observation 3.1 holds because no restriction is placed on these rules. This means that there are no limits on the listener's ability to ask "complex" questions. If some rules were possible, but other rules were not, then Observation 3.1 would not hold.

An immediate consequence of Observation 3.1 is:

Corollary 3.1. *$(\mathcal{I}, \subseteq)$ is a complete lattice.*

This follows because every family of sets which contains the empty set and is closed under union is a complete lattice. Equating equivalent questions (i.e., questions Q_1 and Q_2 such that $A_{Q_1} = A_{Q_2}$), this means that for every family of questions $\mathcal{Q}$, there is a unique question Q^* which is the easiest question that is at least as hard as every question in $\mathcal{Q}$, and a unique question Q_* which is the hardest question which is no harder than any question in $\mathcal{Q}$. If there were limits on the listener's ability to ask questions, then even after equating equivalent questions, this would not be true.

Observation 3.1 and Corollary 3.1 hold also when T and $\mathbf{M}$ are infinite (and of any infinite cardinality). It is important to note that the supremum in $(\mathcal{I}, \subseteq)$ always corresponds to the union; on the other hand, the infimum does not generally correspond to the intersection, but is always contained within the intersection.

Example 3.1. Imagine that the speaker observes a pair of signals, $x_i, i = 1, 2$, taking on values in $\{0, 1\}$. Suppose that the speaker must show exactly one of the two signals to the listener, and cannot lie. The types and message correspondence can be represented as follows:

$$\begin{aligned} M(t_1) &= \{(x_1, 0), (x_2, 0)\} \\ M(t_2) &= \{(x_1, 1), (x_2, 0)\} \\ M(t_3) &= \{(x_1, 0), (x_2, 1)\} \\ M(t_4) &= \{(x_1, 1), (x_2, 1)\}. \end{aligned} \tag{3.2}$$

The message $(x_1, 0)$ proves that the value of the first signal is 0, and the other messages are defined analogously. It possible to represent the set $\mathcal{I}$ using a device from lattice theory known as a *Hasse diagram*:

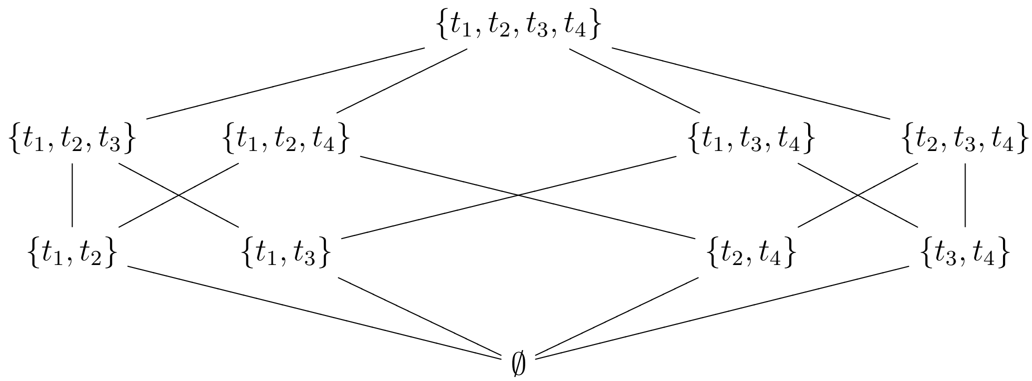

Figure 3.1. Hasse Diagram for an Interior System

To see the relation of this diagram to the message correspondence presented in (3.2), notice that $\{t_1, t_2\}$ corresponds to the acceptance set for the question which only accepts the message $(x_2, 0)$, since only t_1 and t_2 have this message. Likewise $\{t_1, t_2, t_3\}$ corresponds to the acceptance set for the question which accepts $(x_1, 0)$ and $(x_2, 0)$. Every set of types in the diagram is the acceptance set for some question. In general, the idea behind a Hasse diagram is that a line moving up from one set, say $\emptyset$, to another, say $\{t_1, t_2\}$ means that $\{t_1, t_2\}$ is a next largest element after $\emptyset$, in that $\{t_1, t_2\}$ is larger than $\emptyset$, and there is no *other* set in $\mathcal{I}$ which contains $\emptyset$, and is contained in $\{t_1, t_2\}$. As stated by Corollary 3.1, the set $\mathcal{I}$ is a lattice. The join (least upper bound) of any pair of elements in $\mathcal{I}$ is the union of those elements. For example, the join of $\{t_1, t_2\}$ and $\{t_2, t_4\}$, written $\{t_1, t_2\} \vee \{t_2, t_4\}$ is equal to the union $\{t_1, t_2\} \cup \{t_2, t_4\} = \{t_1, t_2, t_4\}$. By contrast the meet (greatest lower bound) of any two elements is not necessarily the intersection. For example the meet of

$\{t_1, t_2\}$ and $\{t_2, t_4\}$, written $\{t_1, t_2\} \wedge \{t_2, t_4\}$, is not their intersection $\{t_2\}$, but rather is $\emptyset$. Likewise, $\{t_1, t_2, t_3\} \wedge \{t_2, t_3, t_4\}$ does not equal the intersection $\{t_2, t_3\}$ but instead is $\emptyset$. This reflects the fact that acceptance of certain types may be tied to a choice between acceptance of other types. For example, the fact that $\{t_1, t_2\} \wedge \{t_2, t_4\} = \emptyset$ reflects the fact that in order to accept type t_2, it is necessary to accept either t_1 or t_4. Once we draw the Hasse diagram, or more generally, find a description of the set $\mathcal{I}$, it is no longer necessary to consider the speaker's messages at all, when asking questions such as: "what is feasible?" or "what is optimal?" Later, it will be shown that it is not even necessary to draw the entire Hasse diagram, but rather we can focus on certain special elements. □

3.2. k Actions

I will now show how the case with two actions serves as a foundation for the case with an arbitrary number of actions. I will present two results–Theorems 3.1 and 3.2–that together will answer the question posed in the beginning of this chapter, namely: what properties characterize the family of implementable outcome functions corresponding to some message correspondence? I will also demonstrate the consequences described at the beginning of this chapter, i.e., show how one can infer what outcome functions are implementable with k actions from what outcome functions are implementable with 2 actions (see Corollary 3.2), as well as demonstrate some closure properties of the set of implementable outcome functions (see Corollary 3.3).

The first step is to extend the definitions which were presented in the previous section to the case of k actions. Consider any persuasion rule $f : \mathbf{M} \to \mathbf{A}$. Such a persuasion rule can be represented by a sequence of sets of messages $(M_1, M_2, \ldots, M_k)$ where M_j is the set of messages m such that $f(m) = a_j$. In order to show the parallelism with the case $k = 2$ for each $j = 1, \dots, k$, define:

$$Q_j = \bigcup_{\ell=j}^{k} M_\ell,$$

and replace $(M_1, M_2, \dots, M_k)$ by $(Q_1, Q_2, \dots, Q_k)$. Thus the sequence $(Q_1, Q_2, \dots, Q_k)$ represents $(M_1, M_2, \dots, M_k)$ cumulatively. In fact, it is possible to eliminate the first component and write $(Q_2, \dots, Q_k)$ because $Q_1 = \mathbf{M}$. I refer to the tuple $(Q_2, \dots, Q_k)$ as a k-*question*. Notice that it must be the case that $Q_2 \supseteq Q_3 \supseteq \cdots \supseteq Q_k$. Given this representation, it is possible to think of any persuasion rule as a sequence of progressively more difficult questions, yielding progressively higher rewards.

To every k-question $\overline{Q} = (Q_2, \dots, Q_k)$, there corresponds the sequence $(A_{Q_2}, \dots, A_{Q_k})$ where A_{Q_j} is the acceptance set corresponding to the question Q_j. Notice here that even with k actions, I still refer to sets A_Q as acceptance sets, and $\mathcal{I} = \{A_Q : Q \subseteq \mathbf{M}\}$ as the family of acceptance sets.

Setting $A_{Q_1} = T$ and $A_{Q_{k+1}} = \emptyset$, k-question $\overline{Q}$ implements outcome function $g : T \to \{a_1, \dots, a_k\}$ iff:

$$g(t) = a_j \Leftrightarrow t \in A_{Q_j} \setminus A_{Q_{j+1}} \tag{3.3}$$

In other words, the k-question $\overline{Q}$ implements the outcome function g if g assigns a_j to t whenever Q_j is the hardest question in $\overline{Q}$ that t can answer. This follows from the definition of implementation (see 1.1) when $\overline{Q}$ is associated with the persuasion rule f from which it was derived.

Summarizing what has been accomplished, we have represented implementable outcome functions in a way which is parallel to the case $k = 2$. That is to say, in the case $k = 2$, an implementable outcome function was associated with the set of types who got the high action. The construction above says that in the general case, we may represent an implementable outcome functions as a *sequence* of sets of types (i.e., acceptance sets), which are decreasing according to inclusion. The next step is to make this parallelism more precise.

Call a family $\mathcal{F}$ of subsets of T an *interior system* if (i) $\emptyset \in \mathcal{F}$, (ii) $T \in \mathcal{F}$, and (iii) $\mathcal{F}$ is closed under union.[2] In Section 3.1, we saw that with two actions, the family of implementable outcome functions can be represented as an interior system. For any interior system define:

$$\mathbf{C}(\mathcal{F}, k) := \{(F_2, \ldots, F_k) : F_2, \ldots, F_k \in \mathcal{I}, F_2 \supseteq F_3 \supseteq \cdots \supseteq F_k\}.$$

Thus $\mathbf{C}(\mathcal{F}, k)$ is the family of decreasing sequences of sets from $\mathcal{F}$ of length $k-1$, possibly with repetitions. A decreasing sequence of sets in $\mathcal{F}$ of length $k-1$ is called a k-*chain* on $\mathcal{F}$, so that $\mathbf{C}(\mathcal{F}, k)$ is the set of k-chains on $\mathcal{F}$. Notice that $\mathbf{C}(\mathcal{F}, 2) = \mathcal{F}$.

The following theorem employs the construction presented above to provide important structural information about the family of implementable outcome functions. In particular, it shows how the question of what is implementable in the case of arbitrary k can be reduced to the question of what is implementable when $k = 2$. Moreover, it will end

[2]The term "interior system" comes from the fact that every interior system is naturally associated with an interior operator, similar to interior operators found in topology. The dual of an interior system is a closure system. Each closure system is associated with a closure operator. Although both closure systems and interior systems have been extensively studied in lattice theory, closure systems are more common. However, every statement about closure systems corresponds to a statement about interior systems by the principle of duality from lattice theory.

up providing half of the answer to the question posed at the beginning of the section, namely: how can one know whether a family of outcome functions is the implementable family for some message correspondence?

Theorem 3.1. *Fix a message structure* $(T, \mathbf{M}, M(\cdot))$*, and let* $\mathcal{I} = \{A_Q : Q \subseteq \mathbf{M}\}$ *be the induced interior system of acceptance sets. Then if there are* k *actions, the implementable outcome functions correspond one-to-one with the elements of* $\mathbf{C}(\mathcal{I}, k)$*.*

Proof. It was shown above how, starting with an implementable outcome function and the persuasion rule which implements it, one can derive a k-chain on $\mathcal{I}$. It is clear from (3.3) that two distinct k-chains cannot correspond to the same outcome function. Finally, I want to show that every k-chain $(I_2, \ldots, I_k) \in \mathbf{C}(\mathcal{I}, k)$ corresponds to some implementable outcome function. Since for $j = 2, \ldots, k$, $A_j \in \mathcal{I}$, there exists Q_j such that $I_j = A_{Q_j}$. Notice that the sets Q_j may not be decreasing (according to inclusion).[3] Nevertheless, one can define a persuasion rule g by the formula in (3.3), and it is easy to verify that the persuasion rule f such that $f(m) = a_j \Leftrightarrow m \in Q_j \setminus \bigcup_{i=j+1}^{k} Q_i$ implements g. $\square$

There are several consequences of Theorem 3.1.[4] The following consequence is notable:

Corollary 3.2. *Fix a message structure. If one knows what outcome functions are implementable when there are two actions but does not directly know the message structure, then one can infer what outcome functions are implementable when there are* k *actions.*

[3] One could argue that the Q_j *can* be chosen so that they are decreasing, although this step is unnecessary in order to complete the argument.

[4] Equation (3.3)–which was used in the proof–can also be used to show that an outcome function g is implementable if and only if each type has a message which is not available to any type which is assigned a lower action, a theorem proved for a similar preference structure by Singh and Wittman (2001).

This is an immediate consequence of the theorem, but to see how to actually make the inference, note that as was seen in Section 3.1, and as stated by Theorem 3.1, the set of implementable outcome functions with two actions corresponds one-to-one with the family $\mathcal{I} = \mathbf{C}(\mathcal{I}, 2)$ of acceptance sets. If one knows the set of implementable outcome functions with two actions, it is easy to construct $\mathcal{I}$: for each implementable outcome function g, $\mathcal{I}$ contains the sets $I = \{t \in T : g(t) = a_2\}$, and does not contain any sets which cannot be formed in this way. Of course, if one does not know the message structure, then one will not know how to represent I in the form A_Q for some $Q \subseteq \mathbf{M}$. To discover whether an outcome function g' is implementable when there are k actions, form the family of k-chains $\mathbf{C}(\mathcal{I}, k)$. We saw in Section 3.1 that $\mathcal{I}$ must contain T, which means that T can actually be inferred by looking at the graphs of the implementable outcome functions. Setting $I_0 = T$, $I_{k+1} = \emptyset$, and using Theorem 3.1, it follows that in order to check whether outcome function g' is implementable with k actions, it suffices to check whether there exists $(I_2, \ldots, I_k) \in \mathbf{C}(\mathcal{I}, k)$ satisfying $g'(t) = a_j \Leftrightarrow t \in I_j \setminus I_{j+1}$. Similarly, it is clear that if one starts with the outcome functions which are implementable with $k \geq 2$ actions, one can infer which outcome functions are implementable with k' actions for any k'. What this reasoning highlights is that once the set $\mathcal{I}$ is formed, one no longer needs to consider messages at all in order to study the question of implementability.

Another consequence of Theorem 3.1 is that the set of implementable outcome functions satisfies certain closure properties. This is easiest to demonstrate by means of example.

Example 3.2. Suppose that when there are three types and three actions, we know that the outcome functions below are implementable.

g_1	g_2
$t_1 \mapsto a_1$	$t_1 \mapsto a_3$
$t_2 \mapsto a_2$	$t_2 \mapsto a_2$
$t_3 \mapsto a_3$	$t_3 \mapsto a_1$

Figure 3.2. Implementable Outcome Functions

Then we can infer that the outcome function:

g_3
$t_1 \mapsto a_3$
$t_2 \mapsto a_1$
$t_3 \mapsto a_2$

Figure 3.3. Implementable Outcome Function

is also implementable. To see this, note that g_1 corresponds to the 3-chain $(\{t_2, t_3\}, \{t_3\})$, and g_2 corresponds to the 3-chain $(\{t_1, t_2\}, \{t_1\})$. From this, it follows that $\{t_1\}$ and $\{t_3\}$ both belong to $\mathcal{I}$. Since $\mathcal{I}$ is closed under union, $\{t_1, t_3\}$ belongs to $\mathcal{I}$. Therefore, $(\{t_1, t_3\}, \{t_1\})$ is a 3-chain in $\mathbf{C}(\mathcal{I}, k)$. This 3-chain corresponds to the outcome function g_3, and therefore g_3 is implementable. $\square$

The reasoning from the example can be generalized:

- Start with a set $\mathcal{G}$ of outcome functions. Consider the family

$$\mathcal{E} = \{\{t \in T : g(t) \geq a_j\} : g \in \mathcal{G}, j = 2, \ldots, k\} \cup \{T, \emptyset\},$$

  where actions a_j are ordered by $\leq$ according to their indices. Let $\mathcal{F}$ be the closure of $\mathcal{E}$ under arbitrary union. Define:

$$\tau(\mathcal{G}) = \{g \in \{a_1, \ldots, a_k\}^T : \exists (I_2, \ldots, I_k) \in \mathbf{C}(\mathcal{F}, k), \forall t, \forall j, g(t) = a_j \Leftrightarrow t \in I_j \setminus I_{j+1}\},$$

where as usual, $I_1 = T, I_{k+1} = \emptyset$.

It follows from Theorem 3.1 that if $\mathcal{H}$ is the family of implementable outcome functions induced by some message correspondence and if $\mathcal{G} \subseteq \mathcal{H}$, then $\tau(\mathcal{G}) \subseteq \mathcal{H}$. However, what is not yet clear is whether $\tau(\mathcal{G})$ is the family of implementable outcome functions corresponding to some message correspondence. In other words, Theorem 3.1 tells us that families of implementable outcome functions are closed under the operation τ, but it leaves open the possibility that such families must satisfy other requirements. The following theorem shows that there are no other requirements, and so completes the answer to the question with which we started: how can one know whether a family of outcome functions is the family of implementable outcome functions corresponding to some message correspondence?

Theorem 3.2. *Let $\mathcal{F}$ be any interior system. Then there exists a message structure whose induced family of acceptance sets is $\mathcal{F}$.*

Proof. Choose $\mathbf{M}$ so that $|\mathbf{M}| = |\mathcal{F}|$, and choose a one-to-one map $F \mapsto m_F$ from $\mathcal{F}$ to $\mathbf{M}$. Then define a message correspondence by:

$$M(t) = \{m_F : t \in F, F \in \mathcal{F}\}.$$

In other words, t has one message for each element F of $\mathcal{F}$ to which he belongs. Since $T \in \mathcal{F}$, $m_T \in M(t)$, for all $t \in T$, so $M(t) \neq \emptyset$. As usual, let $\mathcal{I} = \{A_Q : Q \subseteq \mathbf{M}\}$. I want to prove that $\mathcal{I} = \mathcal{F}$. For any $F \in \mathcal{F}$, $A_{\{m_F\}} = F$, so $\mathcal{F} \subseteq \mathcal{I}$. Next consider $Q \subseteq \mathbf{M}$.

$$A_Q = \bigcup \{A_{\{m_F\}} : m_F \in Q\} = \bigcup \{F : m_F \in Q\} \in \mathcal{F},$$

where membership in $\mathcal{F}$ follows because $\mathcal{F}$ is closed under union. So $\mathcal{I} \subseteq \mathcal{F}$. So $\mathcal{F} = \mathcal{I}$.

$\square$

With this theorem, we have answered the question posed above. To summarize:

Corollary 3.3. *For any family of outcome functions $\mathcal{G}$, the smallest family of outcome functions containing $\mathcal{G}$ and corresponding to the set of implementable outcome functions for some message correspondence is $\tau(\mathcal{G})$. In particular $\mathcal{G}$ is the family of implementable outcome functions corresponding to some message correspondence if and only if $\mathcal{G} = \tau(\mathcal{G})$.*

Thus families of implementable outcome functions for some message correspondence are precisely the fixed points of τ. As a consequence of Tarski's fixed point theorem, the family of such families is itself a complete lattice when ordered by inclusion. Moreover, and more importantly, in the finite case, τ provides a method for checking whether a family of outcome functions is the implementable family for some message correspondence.

Example 3.3. This continues Example 3.2. Let g_1 and g_2 be as in Example 3.2, and assume again that g_1 and g_2 are implementable. Next consider the outcome function:

g_4
$t_1 \mapsto a_1$
$t_2 \mapsto a_3$
$t_3 \mapsto a_2$

Figure 3.4. Outcome Function

Can we infer whether g_4 is implementable? Theorem 3.2 can be used to answer this question. Consider the following pair of interior systems:

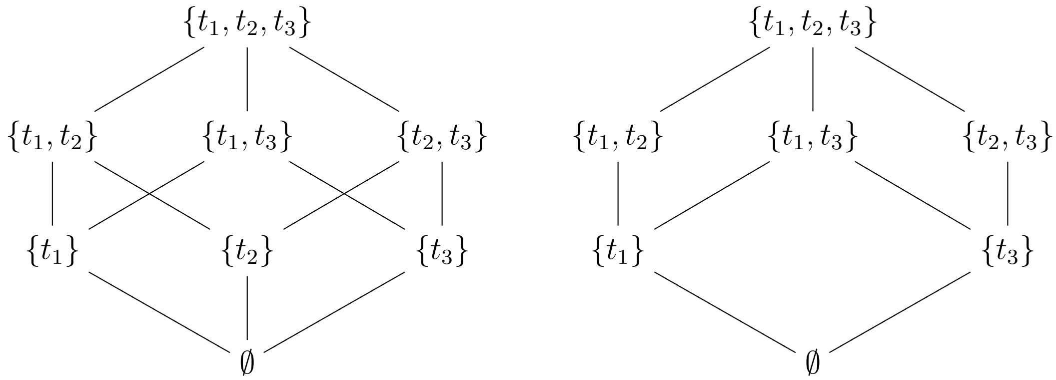

Figure 3.5. Two Interior Systems

Using the interior system on the left, it is possible to build 3-chains corresponding to g_1, g_2, and g_4 (and in fact, corresponding to any outcome function). On the other hand, in the interior system on the right, it is possible to build 3-chains corresponding to g_1 and g_2 but it is not possible to build the 3-chain $(\{t_2, t_3\}, \{t_2\})$ which corresponds to g_4. Using Theorem 3.2, and verifying that the Hasse diagram on the right represents an interior system (i.e., is closed under union, and contains T and $\emptyset$), it follows that there is some message structure which generates it. But then by Theorem 3.1, g_4 is not implementable given this message structure. So we can conclude that implementability of g_1 and g_2 does not imply implementability of g_4 (nor, of course, does it imply that g_4 is not implementable). $\square$

Before concluding this section, I extend the notion of difficulty of question from the case of two actions to k as well. In particular define the ordering $\leq$ on $\mathbf{C}(\mathcal{I}, k)$ by:

$$(A_2, \dots, A_k) \leq (B_2, \dots, B_k) \Leftrightarrow A_j \subseteq B_j, \forall j = 2, \dots, k. \tag{3.4}$$

Thus, we can say that k-question $(Q_2, \ldots, Q_k)$ is more difficult than k-question $(R_1, \ldots, R_k)$ if $(A_{Q_2}, \ldots, A_{Q_k}) \leq (A_{R_2}, \ldots, A_{R_k})$. In other words, a more difficult k-question is one which assigns to every type a lower action, and hence makes every type at least weakly worse off.

Observation 3.2. $(\mathbf{C}(\mathcal{I}, k), \leq)$ *is a complete lattice.*

To see this, note first that the product of complete lattices is a complete lattice under the product ordering, and then that $\mathbf{C}(\mathcal{I}, k)$ is a subset of the k-fold product $\mathcal{I} \times \cdots \times \mathcal{I}$ which, moreover, is closed under supremum and infimum of subsets within the product lattice.

3.3. Join-Irreducible Elements and Essential Messages

In this section, I will complete the task of showing that the message correspondence and persuasion rules are redundant. To do this, I will show how one can recover all important information about the message correspondence from the family of implementable outcome functions. A basic notion from lattice theory which will be employed for this purpose is that of a join-irreducible element. Within an interior system–or any lattice, for that matter–join-irreducible elements play a role which is analogous to the role played by prime numbers relative to the natural numbers. In particular, join-irreducible elements are the basic elements out of which all other elements of the interior system can be built. Join-irreducible elements will be shown to correspond to *essential* messages, i.e., messages which cannot be eliminated from the message correspondence without altering the family of implementable outcome functions.

A *non-empty* element F of a family of sets $\mathcal{F}$ is said to be *join-irreducible* in $\mathcal{F}$ iff $\forall G, H \in \mathcal{F}$:

$$F = G \cup H \Rightarrow F = G \text{ or } F = H.$$

Thus an element F of an interior system is join-irreducible if it cannot be expressed as the union of two elements G and H, both of which are distinct from F. Let $J(\mathcal{F})$ be the set of join-irreducible elements of $\mathcal{F}$. The following fact shows that the set of join-irreducible elements of an interior system is exactly the set of elements out of which all other elements of the interior system can be built.

Fact 3.1. *Let $\mathcal{F}$ be an interior system on a finite set T. Then every element F of $\mathcal{F}$ is a union of some subset of $J(\mathcal{F})$, and no element F of $J(\mathcal{F})$ can be expressed as a union of elements in $\mathcal{F} \setminus \{F\}$.*

This means that the set $J(\mathcal{F})$ summarizes all information about the interior system. In particular, in the finite case, the entire interior system can be recovered through the equation:

$$\mathcal{F} = \{\bigcup \mathcal{G} : \mathcal{G} \subseteq J(\mathcal{F})\}$$

Moreover, for any set $\mathcal{K} \subseteq \mathcal{F}$ such that $J(\mathcal{F}) \nsubseteq \mathcal{K}$, we have $\mathcal{F} \neq \{\bigcup \mathcal{G} : \mathcal{G} \subseteq \mathcal{K}\}$. In fact there is no way to recover all the information about an interior system from any proper subset of $J(\mathcal{F})$.

The following lemma says that every join-irreducible element is the acceptance set for a one-message question.

Lemma 3.1. *Suppose T is finite. Let $\mathcal{I}$ be the family of acceptance sets induced by message structure $(\mathbf{M}, T, M(\cdot))$, and let $F \in J(\mathcal{I})$. Then there exists $m \in \mathbf{M}$ such that $A_{\{m\}} = F$.*

Proof. Since $F \in \mathcal{I}$, there exists $Q \subseteq \mathbf{M}$ such that $F = A_Q$. But then $F = \bigcup\{A_{\{m\}} : m \in Q\}$. Since $F \in J(\mathcal{I})$, there is $m \in Q$ such that $F = A_{\{m\}}$. □

Notice that if $F \in \mathcal{F}$ but $F \notin J(\mathcal{F})$, there may not exist $m \in \mathbf{M}$ such that $A_{\{m\}} = F$. In other words, the conclusion of the lemma can only be derived when F is join-irreducible. Call a message $m \in \mathbf{M}$ *essential* if $A_{\{m\}} \in J(\mathcal{I})$. In other words, a message is essential if the set of types who can send it is a join-irreducible element of $\mathcal{I}$. Call a collection $\mathbf{M}_0$ of essential messages *complete* if $\{A_{\{m\}} : m \in \mathbf{M}_0\} = J(\mathcal{I})$. Lemma 3.1 implies that a complete set of essential messages always exists. Finally, call a complete collection $\mathbf{M}_0$ of essential messages *minimal* if no proper subset of $\mathbf{M}_0$ is a complete set of essential messages. The following lemma shows that if $\mathbf{M}_0$ is a complete and minimal set of essential messages, then all other messages are redundant.

Theorem 3.3. *Suppose T is finite. Consider a message structure $(T, \mathbf{M}, M(\cdot))$, and assume k actions. Let $\mathbf{M}_0 \subseteq \mathbf{M}$ be a minimal and complete set of essential messages, and define $M_0(t) = M(t) \cap \mathbf{M}_0$. Then $(T, \mathbf{M}, M(\cdot))$ and $(T, \mathbf{M}_0, M_0(\cdot))$ generate the same set of implementable outcome functions.*

Proof. Let $\mathcal{I}$ be the family of acceptance sets induced by $(T, \mathbf{M}, M(\cdot))$. Since $\mathbf{M}_0$ is a minimal and complete set of essential messages, one can construct a bijection $F \mapsto m_F$ from $J(\mathcal{I})$ to $\mathbf{M}_0$. Let $\mathcal{I}_0$ be the family of acceptance sets generated by $(T, \mathbf{M}_0, M_0(\cdot))$. Clearly, $\mathcal{I}_0 \subseteq \mathcal{I}$. Next choose $I \in \mathcal{I}$. Then, by Fact 3.1, there exists $\mathcal{J} \subseteq J(\mathcal{I})$ such that

$\bigcup \mathcal{J} = I$. But then defining $Q = \{m_F : F \in \mathcal{J}\}$, $A_Q = I$.[5] So $I \in \mathcal{I}_0$. So $\mathcal{I} = \mathcal{I}_0$, and the proof is completed by Theorem 3.1. □

The proof of Theorem 3.3 provides a method for recovering all non-redundant messages from a family $\mathcal{G}$ of implementable outcome functions. In particular, first identify the corresponding interior system $\mathcal{I}$. Then for each element F of $J(\mathcal{I})$, select a message which is available exactly to the types in F. The resulting message correspondence will generate $\mathcal{G}$. Each message in the range of the message correspondence so constructed will be equivalent to a message in the message correspondence which originally generated $\mathcal{G}$. The original message correspondence may have had additional messages, but these messages were redundant, and for that reason it is impossible to know whether these messages existed on the basis of the information in $\mathcal{G}$. So, we have now shown that the message correspondence is dispensable in the sense that all important information about it is contained in $\mathcal{G}$.

Essential messages have been defined globally relative to the entire interior system. This is useful for establishing that the acceptance sets corresponding to the essential messages provide all the information necessary for constructing the set of implementable outcome functions. It will also prove useful for analyzing the maximum possible number of non-redundant messages. On the other hand, one can characterize the essential messages in a local manner. This will be done by means of the notion of a *maximally informative* message. A message m is maximally informative for a type t speaker if t does not have a message which rules out strictly more types than m. Formally, the set $M^{\#}(t)$ of maximally

[5] In particular, considering $I = T$, this establishes that for all $t \in T$, $M_0(t) \neq \emptyset$, so that $M_0(\cdot)$ is a legitimate message correspondence.

informative messages for type t is defined:

$$M^{\#}(t) := \{m \in M(t) : \forall m' \in M(t), A_{\{m'\}} \not\subset A_{\{m\}}\},$$

where "$\subset$" means proper subset. Note that nothing rules out the possibility that there is some redundancy in the speaker's set of maximally informative messages, in the sense that two maximally informative messages may correspond to the same acceptance set.

Even when there is no redundancy in $M^{\#}(t)$, the speaker may have several maximally informative messages. The presence of more than one maximally informative message is caused by constraints which limit the speaker's ability to present all his information. For example, suppose that the speaker has ten minutes to present as much information as he can. Ten minutes is not enough time for the speaker to present all the information that he has. The maximally informative messages would correspond to what the speaker could communicate if he used the ten minutes to convey as much information as possible. There would generally be several distinct ways to do this, depending on which information the speaker chooses to convey in the time allotted. On the other hand, if the speaker stops talking after only five minutes, he has not provided a maximally informative message.

The following theorem relates a speaker's maximally informative messages to the set of essential messages.

Theorem 3.4. *Suppose T is finite. A message is essential if and only if it is maximally informative for some type. Consequently*

$$J(\mathcal{I}) = \{A_{\{m\}} : m \in \bigcup_{t \in T} M^{\#}(t)\}. \tag{3.5}$$

The proof relies on a standard fact about interior systems and is presented in Section 3.4. The theorem says that if a message m is one of type t's maximally informative messages, then m is essential. The converse is not true: it is not necessarily the case that every essential message belonging to $M(t)$ is maximally informative for t. Some essential messages are more informative than others, but each essential message is maximally informative for *some* type.

Theorems 3.3 and 3.4 together imply that in implementing any outcome function, it is without loss of generality that each type uses messages which are maximally informative to some type. The theorems do not directly say that it is without loss of generality that each type t uses a message that is maximally informative for t, but this is also true. In particular, any implementable outcome function can be implemented by means of a k question $\overline{Q} = (Q_2, \ldots, Q_k)$. Now consider the k-question $\overline{R} = (R_2, \ldots, R_k)$ such that

$$R_j = \{m : A_{\{m\}} \subseteq A_{Q_j}, m \text{ is essential}\}.$$

It is easy to see that $\overline{R}$ implements the same outcome function as $\overline{Q}$, and that every type t who answers the most difficult question in $\overline{R}$ which he can answer, can answer it using a message in $M^{\#}(t)$. Returning to the example of a time constrained speaker, this means that it is without loss of generality that the listener will only grant the speaker an action which is better than the worst action if the speaker uses all of his time to convey as much information as possible, regardless of the listener's objective. There is some analogy between this observation and the revelation principle. Intuitively, the revelation principle says that it is without loss of generality that a speaker will report his type–i.e., provide all his information.[6] Clearly, this cannot happen under time constraints, but it is still without loss of generality that the speaker gives some maximally informative report. When time constraints are lifted, and there is a unique maximally informative report, one recovers the revelation principle.

The use of join-irreducible elements makes it relatively easy to answer questions which might be quite difficult to address otherwise. For example, suppose that one knows the number of types is n, but knows nothing else about the message structure. Suppose now that one learns that there are ℓ messages. Will this piece of knowledge provide any information about what is implementable? Clearly, if $\ell = 1$, then it is possible to infer that the only implementable outcome functions are the ones which assign the same action to each type. How big does ℓ have to be so that no information is provided? In other words, what is the smallest number of messages which would not limit the generality of the model when there are n types?

Define:

$$b(n) = \max\{|J(\mathcal{F})| : \mathcal{F} \text{ is an interior system on } \{1, \ldots, n\}\}.$$

Thus $b(n)$ is the maximum number of join-irreducible elements in an interior system on a ground set of size n.

Fact 3.2. $\binom{n}{\lfloor n/2 \rfloor} \leq b(n) \leq \binom{n}{\lfloor n/2 \rfloor}(1 + \frac{1}{n^{1/2}})$.

[6] As discussed by Bull and Watson (2007), in an environment with evidence in which the revelation principle holds, it may generally not be sufficient for a type to present all their evidence, but there may also be a role for a "truthful" cheap talk message. However, the speaker's preferences in this model coupled with the restriction to deterministic persuasion rules eliminates any productive role for cheap talk.

This fact follows from theorems of Sperner (1928) and Kleitman (1976), and is explained further in Section 3.5.2. It is therefore clear that $b(n)/2^n \to 0$. In other words, as n gets large, $b(n)$ is much smaller than the size of the powerset of $\{1, \ldots, n\}$.

Formally, say ℓ messages are *unrestrictive* for n types if every family of outcome functions which are implementable given some message structure with n types are implementable given some message structure with n types and ℓ messages.

Theorem 3.5. *The smallest number of messages which are unrestrictive for n types is $b(n)$.*

Proof. This theorem is an immediate consequence of Theorem 3.3. $\square$

The bound $b(n)$ given in the theorem is independent of the number k of actions. From another perspective, $b(n)$ is the maximum possible number of essential messages when there are n types.

This subsection has focused on the finite case. When T is infinite, Fact 3.1 is no longer true. In fact join-irreducible elements play a much less important role when T is infinite, and it is no longer without loss of generality that each type communicates one of his maximally informative messages. The analog of Theorem 3.5 when T is infinite (of any infinite cardinality) is:

Theorem 3.6. *Suppose T is infinite. The smallest cardinality of the set of messages which is unrestrictive is $|2^T|$.*

Proof. See Section 3.5.3. $\square$

3.4. An Example

This section will demonstrate some of the main concepts introduced above by means of an example. Below is an example which shows an interior system with the maximum possible number of join-irreducible elements when there are three types ($b(3) = 4$):

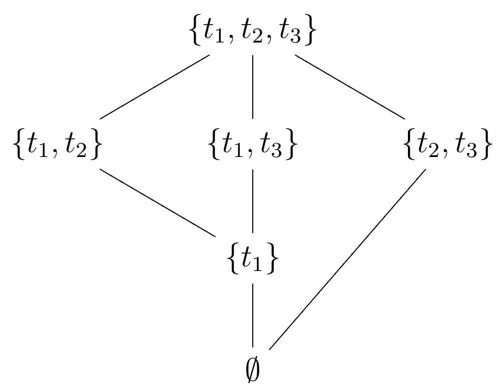

Figure 3.6. Maximum Number of Join Irreducible Elements

Every element in the interior system displayed above except for $\emptyset$ and $\{t_1, t_2, t_3\}$ is join-irreducible, so that there are four join-irreducible elements. In general, one can read off the join-irreducible elements–and hence also the essential messages–from a Hasse diagram, by looking for those elements which have exactly one immediate predecessor, in other words, have exactly one edge moving *down* to another element. In order for this to be the interior system induced by some message correspondence, there must be one message per join-irreducible element (i.e., which can be used by types contained in that join-irreducible element), and the rest of the messages are redundant. Looking back at Example 3.1 for join-irreducible elements, it is clear that there are four of these elements, since there are four elements with exactly one immediate predecessor. Since Example 3.1 contains four types, and $4 < 6 = \binom{4}{2}$, Example 3.1 is not one which requires the maximum number of messages when there are four types.

I will now use the diagram to analyze what is implementable when there are k actions. First, consider the question of whether the decision to assign t_1 an action *at least* as high as a_i puts any constraint on what must be assigned to other types. The answer is no. To see this start with any outcome function g_1 such that $g_1(t_1) = a_\ell$ with $\ell < i \leq j$. Form the corresponding k-chain $(I_2, \ldots, I_\ell, \ldots, I_j, \ldots, I_k)$. Then t_1 only occurs in sets I_h with index at most ℓ. It is easy to see from the diagram that

$$(I_2, \ldots, I_\ell, I_{\ell+1} \cup \{t_1\}, \ldots, I_{j-1} \cup \{t_1\}, I_j \cup \{t_1\}, I_{j+1}, \ldots, I_k)$$

is also a k-chain, and the corresponding outcome function g_2 only differs from g_1 in that $g_2(t_1) = a_j$. So it is always possible to grant type t_1 a higher action without changing the action of any other type.

Next, consider the question of whether the decision to assign t_2 an action *at most* as high as a_i puts any constraint on what must be assigned to other types. The answer is yes. If $(I_2, \ldots, I_k)$ is the k-chain corresponding to an outcome function g_1 which gives t_1 an action a_j with $j \leq i$, then I_j must be some set in the diagram containing t_1, and any set I_ℓ with $j < \ell$ in the k-chain must be a subset of I_ℓ not containing t_1. I_ℓ can be nonempty only if $I_\ell = \{t_2, t_3\}$ and $I_j = \{t_1, t_2, t_3\}$. This means is that if t_1 is assigned at most a_i, this requires that types t_2 and t_3 are given at most a_i *unless* they are assigned the same action.

The above discussion shows that the join-irreducible elements simultaneously correspond to the messages and to the constraints faced by the listener.

3.5. Proofs

3.5.1. Proof of Theorem 3.4

Suppose $m \in M(t)$, but $A_{\{m\}} \notin J(\mathcal{I})$. Then there exist $I, I' \in \mathcal{I}$ such that $A_{\{m\}} = I \cup I', I \neq A_{\{m\}}, I' \neq A_{\{m\}}$. Then w.l.o.g., $t \in I$. So for some $Q \subseteq \mathbf{M}$, $I = A_Q$. So for some $m' \in Q$, $t \in A_{\{m'\}}$, and $A_{\{m'\}}$ is a proper subset of $A_{\{m\}}$, implying that $m \notin M^{\#}(t)$. So a maximally informative message for some type is essential.

Next suppose that $m \notin \bigcup_{t \in T} M^{\#}(t)$. If $A_{\{m\}} = \emptyset$, then m is not essential, since $\emptyset \notin J(\mathcal{I})$. So suppose $A_{\{m\}} \neq \emptyset$. Then for each $t \in A_{\{m\}}$, there exists $m_t \in M(t)$ such that $A_{\{m_t\}} \subset A_{\{m\}}$ and $A_{\{m\}} = \bigcup \{A_{\{m_t\}} : t \in A_{\{m\}}\}$. Since T is finite, it is possible to find some $B \subseteq A_{\{m\}}$ such that $A_{\{m\}} = \bigcup \{A_{\{m_t\}} : t \in B\}$, but for all $s \in B$, $A_{\{m\}}$ is a proper subset of $\bigcup \{A_{\{m_t\}} : t \in B \setminus \{s\}\}$. So choose some $s \in B$, and $A_{\{m\}} = A_{\{m_s\}} \cup \bigcup \{A_{\{m_t\}} : t \in B \setminus \{s\}\}$, and $A_{\{m\}} \neq A_{\{m_s\}}, A_{\{m\}} \neq \bigcup \{A_{\{m_t\}} : t \in B \setminus \{s\}\}$. So m is not essential. So being essential implies being maximally informative to some type.

(3.5) then follows from Lemma 3.1. $\square$

3.5.2. Explanation of Fact 3.2

The lower bound comes from Sperner (1928), who proved that the maximum number of sets in an antichain included in a powerset lattice on a ground set of size n is $\binom{n}{\lfloor n/2 \rfloor}$. The closure of this antichain under union is an interior system and the set of join-irreducible elements in this interior system is the antichain.

The upper bound is based on a paper of Kleitman (1976), which found an upper bound on the maximum size of a subset of a powerset lattice on a ground set of size n

containing no two sets and their union. The precise form of the bound comes from Greene and Kleitman (1978).

3.5.3. Proof of Theorem 3.6

In the infinite case, the relevant notion is no longer the cardinality of the collection of join-irreducible elements, but rather the minimum cardinality of a join-dense set, where a join-dense set $\mathcal{J}$ is a set in an interior system $\mathcal{F}$ is a set $\mathcal{J} \subseteq \mathcal{F}$ such that every element of $\mathcal{F}$ is a union of elements of $\mathcal{J}$. In the finite case, the unique minimum cardinality join dense set is the set of join-irreducible elements, but in the infinite case these two concepts diverge. When T is infinite, The minimum cardinality join-dense set in the family of acceptance sets $\mathcal{I}$ is clearly bounded above by $|2^T|$. In fact, it is exactly $|2^T|$. This follows from the fact that for any infinite set of cardinality T, there exists an antichain in 2^T (i.e., a collection of sets no two of which are ordered by inclusion) of cardinality 2^T and whose union is T. Let $\mathcal{X}$ be this antichain, then $\mathcal{I} = \{\bigcup \mathcal{Y} : \mathcal{Y} \subseteq \mathcal{X}\}$ is an interior system whose minimal cardinality join-dense set is $\mathcal{X}$, and in fact, $\mathcal{X}$ *is* the set of join-irreducible elements of $\mathcal{I}$ in this case. $\square$

CHAPTER 4

Normal Message Structures

The analysis of the previous chapter allowed for arbitrary message correspondences. In Section 4.1, I will explain how message correspondences may be used to represent limitations on a speaker's ability to summarize information, as occurs when a speaker faces time constraints, or the listener is subject to attention constraints. In Section 4.2, I will present a condition–known as *normality*–which was originally introduced by Bull and Watson (2007), and which is related to the *nested range condition* of Green and Laffont (1986), and which imposes the condition that a speaker does not face any constraints which limit his ability to summarize his information. Then I will study the lattice structure of families of implementable outcome functions–much in the spirit of the previous chapter–under the assumption of normality. It will be found that under the assumption of normality the family of implementable outcome functions can always be represented as the set of monotone outcome functions with respect to some ordering relation, whereas in the absence of normality, the family of implementable outcome functions can never be represented in this way. Section 4.3 will introduce the notion of *weak normality*, which intuitively means that the speaker can summarize small (i.e. finite) but not large (i.e. infinite) amounts of information, and study properties of families of implementable outcome functions which satisfy this assumption.

4.1. How Message Correspondences Represent Inability to Completely Summarize Information

There are many circumstances in which a speaker has the option of making each of several statements, but does not have the option of making all these statements. For example, in making a speech, time constraints or the limits imposed by the audience's attention span, may force a speaker to choose which statements to make and which statements to omit. When the speaker is limited in this way, I say that he is forced to be *selective* in the presentation of information. On the other hand when the speaker has the option of presenting all his information, either using one statement or many, I say the speaker can *summarize* his information. The notion of summary often has the connotation of a brief statement which makes one's main points. However, that is not how it is used here. I use "summary" to mean *complete* summary, or in other words, a statement or series of statements which provide *all* of one's information. In the sense in which the term is used here, the speaker can summarize his information when there are no time, attention, or other constraints which prevent this. It is also possible to talk about summarizing a specific part of one's information, for example summarizing one's information about a topic. This would correspond to presenting all of one's information about that topic. I sometimes use the term *combine* instead of summarize, as it is sometimes significantly more natural, especially when discussing a part of the speaker's information. In this section, I will explain how the message correspondence can be used to formally model these notions.

Consider the following example of a message correspondence:

$$
\begin{aligned}
M(t_1) &= \{m_0, m_1, m_2\} \\
M(t_2) &= \{m_0, m_1\} \\
M(t_3) &= \{m_0, m_2\} \\
M(t_4) &= \{m_0\}
\end{aligned}
\tag{4.1}
$$

The messages m_1 and m_2 are independent in this example in the sense that knowledge that type t can send m_1 does not help one to predict whether he can send m_2, and vice versa. If type t_2 were eliminated, then an asymmetry would be created in that ability to send of m_1 would imply ability to send of m_2, but the converse would not be true. If t_3 were eliminated as well, then a stronger dependence of the messages would be created in that ability to send m_1 depends on ability to send m_2 and vice versa.

The ability to summarize information can be understood in terms of a kind of dependence as well. Suppose again that all four types in (4.1) exist. Message m_1 means that the type is either t_1 or t_2, and m_2 means that the type is either t_1 or t_3, but there is no message which summarizes these two messages, and says that the type is t_1. If there are are two rounds of communication, over which the message correspondence remains constant, then t_1 could combine these messages by first sending m_1 and then m_2. Equivalently, if t_1 could combine his messages with the word "and" (written &), he could send message $m_1\&m_2$. Both of these methods provide the speaker with a communication that is available exactly when he has both messages m_1 and m_2. Suppose now that we rewrite

t_1's message space and add two types, t_5 and t_6, as follows:

$$\begin{aligned} M(t_1) &= \{m_0, m_1, m_2, m_3\} \\ M(t_5) &= \{m_0, m_1, m_2, m_3, m_4\} \\ M(t_6) &= \{m_0, m_4\} \end{aligned}$$

Then m_3 can be interpreted as $m_1 \& m_2$, since it is available to exactly the types who have m_1 and m_2. This example also shows that the ability to combine messages may be partial, as t_5 can combine messages m_1 and m_2 by using m_3, but cannot combine m_1 and m_4. Below, when the effect of being able to completely summarize information information is studied, I will assume only that the speaker can summarize all of his information, not that he can combine any part.

I conclude this section with a final caveat: the slogan here is that "not all messages are created equal." Imagine that the speaker read part of the newspaper this morning. Different types of speaker read different parts of the newspaper where these parts may overlap in all kinds of ways (different stories, different parts of the same story, different stories on the same topic, etc.). The speaker is only willing to report facts that he actually read, and this is how the messages depend on the type. Then messages may stand in complex relationships, and there may be no intuitively independent messages like m_1 and m_2 above.

4.2. Normality

The analysis of Chapter 3 allowed for the possibility that the speaker faces time, attention, or other constraints which force him to be selective in the presentation of information. In this section, I will analyze precisely what changes when such constraints are lifted, and therefore the speaker has the capacity to completely summarize his information. The construction and results of Chapter 3 will prove to be very useful for this purpose. In this section as well as the next chapter, I will show that the ability to combine information is critical for multiple properties simultaneously.

The first task is to define formally what it means for the speaker to be able to summarize his information. Recall from Section 4.1 that a message here is interpreted to mean a total communication. For example, in a job interview, if an applicant has ten qualifications, and has the option of presenting any five or fewer qualifications, then one could model every set of five or fewer qualifications as a message. If the speaker were constrained in terms of the order in which the qualifications could be presented, then one would model every sequence respecting these constraints as a message. Given that messages are treated as total communications, in the job interview example, a definition of the ability to summarize information which would say that the speaker could present as many qualifications as he has would not be appropriate because it is a definition in terms of the "parts" of messages rather than in terms of the messages themselves.

The definition which I use to impose the condition that a speaker can summarize all of his information comes from Bull and Watson (2007), and is similar to the *nested range condition* of Green and Laffont (1986).

Definition 4.1. A message structure $(T, \mathbf{M}, M(\cdot))$ is *normal* if $\forall t, \exists m(t) \in M(t)$, such that

$$\forall s \in T, m(t) \in M(s) \Rightarrow M(t) \subseteq M(s).$$

$m(t)$ is referred to as t's *maximal message.*

To relate this to the discussion of Section 4.1, recall that in that section, it was explained that if there are two messages, m_1 and m_2, and one wanted to have a message which means "m_1 and m_2", the way that this would be done is to give a message m_3 exactly to the types who have both m_1 and m_2. m_3 is then a message which communicates both m_1 and m_2. What normality says is simply that for every collection $\{m_1, m_2, \ldots, m_\ell\}$ of messages such that for some type t, $\{m_1, m_2, \ldots, m_\ell\} = M(t)$, there exists a message $m(t)$ which is available to exactly those types who can say $\{m_1, m_2, \ldots, m_\ell\}$. This means that $m(t)$ actually belongs to $\{m_1, m_2, \ldots, m_\ell\}$. To understand this, consider again the example above in which m_3 was used to mean "m_1 and m_2". Then a type can send m_3 if and only if that type can send m_1 and m_2. Equivalently, a type can send m_3 if and only if that type can send every message in the set $\{m_1, m_2, m_3\}$. Notice that if $\{m_1, m_2, \ldots, m_\ell\}$ is not equal to $M(t)$ for any type t, then normality does not say anything about the ability to summarize $\{m_1, m_2, \ldots, m_\ell\}$. For example $\{m_1, m_2, \ldots, m_\ell\}$ may be a proper subset of $M(t)$ for some type t. Normality would then not imply that the speaker has a message which which is equivalent to "m_1 and m_2 and $\ldots$ and m_ℓ", but it does imply that the speaker has a message which is equivalent to the conjunction of all the messages in the larger set $M(t)$. Section 4.5.1 contains an example which demonstrates this point, namely that when the message correspondence satisfies normality, there may still be subsets of a speaker's information which the speaker cannot exactly summarize.

In a multi-agent mechanism design environment with provability, Bull and Watson (2007) show that normality implies a version of the revelation principle (their *strong revelation principle*). In particular, if normality holds, then every implementable outcome function can be implemented in such a way that each agent reports his maximal message $m(t)$. In their framework an agent may also report a cheap talk message, but in the current framework, since all types of speaker have the same preference, and once we restrict attention to deterministic persuasion rules, there can be no productive role for cheap talk. In a related paper, Deneckere and Severinov (2001) assume that agents can send as many messages as they want to the mechanism designer and show that a version of the revelation principle holds in such an environment. As the preceding discussion makes clear, in the current model this assumption is stronger than normality. In their framework, Deneckere and Severinov (2001) show that the mechanism designer only has to consider incentive constraints for a type t which keep him from pretending to be type s where $M(s) \subseteq M(t)$. It is not hard to see that this conclusion continues to hold if one makes the weaker assumption that the message structure is normal. Forges and Koessler (2005) also study a related model, and emphasize the point, which translated into my terminology, is that in order for the revelation principle to hold, it is only necessary that the speaker be able to summarize all his information, not that he can summarize any part of his information. In what follows, and unlike the papers described above, my concern is not with the revelation principle. Although I will not treat the revelation principle formally, I will point out where the revelation principle could have been used instead of the reasoning I employ.

Corollary 3.2 established that in order to know what outcome functions are implementable with k actions, it is sufficient to know which outcome functions are implementable with 2 actions. In other words, it is sufficient to know the family $\mathcal{I}$ of acceptance sets. The next theorem characterizes the families of acceptance sets induced by normal message structures. The proof uses two ingredients. The first ingredient is the definition of normality, namely that the speaker has a message which summarizes all his information. The second ingredient is the fact that $\mathcal{I}$ is closed under union. Recall that this is true because the listener can choose to assign actions to messages according to any rule he wishes, no matter how complex. Thus the proof of the theorem combines the speaker's ability to summarize information, and the listener's ability to ask any question he wishes.

Theorem 4.1. *Let $\mathcal{I}$ be the family of acceptance sets induced by $(T, \mathbf{M}, M(\cdot))$. $\mathcal{I}$ is closed under intersection if and only if $(T, \mathbf{M}, M(\cdot))$ is normal.*

Proof. See Section 4.5.2. □

Above, we saw that in general, the only property–other than containing T and $\emptyset$–which is shared by all families of acceptance sets is that they are closed under union. Now we see that the property which is shared by all families of acceptance sets corresponding to normal message structures is that they are closed under union *and* intersection. A family of sets closed under union and intersection is sometimes referred to as a "ring of sets." Since a ring of sets can have at most n join-irreducible elements, an immediate corollary of the theorem–invoking Theorem 3.3–is:

Corollary 4.1. *Under normality, n messages are without loss of generality for n types.*

Compare this corollary with Theorem 3.5, which said that in the general case $b(n)$– that is, approximately $\binom{n}{n/2}$–messages is the smallest number of messages which is without loss of generality when there are n types. There are in fact, several ways to derive the corollary. For example, one could observe that under normality the set $M^{\#}(t)$ of maximally informative messages is essentially unique in the sense that any pair of messages that it contains are equivalent. This observation and Theorem 3.4 also imply the corollary. To summarize the consideration which is important here:

Observation 4.1. *Under normality, $\{m(t) : t \in T\}$ is a minimal and complete set of essential messages.*

A third method of deriving the corollary would be to appeal to the versions of the revelation principle described above. However, the methods of deriving this corollary which do not appeal to the revelation principle have the advantage of explaining Corollary 4.1 using the same principles which were used to derive Theorem 3.5, where the revelation principle would not have been available. Therefore these other methods explain the relationship between the numbers $b(n)$ when normality may not hold and n when normality is assumed to hold.

I now provide an alternative characterization of the implementable outcome functions. A *quasi-order* is a relation which is reflexive and transitive. Henceforth, I assume a ring of sets $\mathcal{R}$ on T contains both T and $\emptyset$, as well as being closed under union and intersection. Given any interior system $\mathcal{I}$, one can define a quasi-order $\preccurlyeq^{\mathcal{I}}$ by:

$$t_1 \preccurlyeq^{\mathcal{I}} t_2 \Leftrightarrow \forall I \in \mathcal{I}, t_1 \in I \Rightarrow t_2 \in I.$$

On the other hand for any quasi-order $\preccurlyeq$, one can define a ring of sets $\mathcal{R}_{\preccurlyeq}$ by:

$$\mathcal{R}_{\preccurlyeq} = \{\{t \in T : \exists s \in S : s \preccurlyeq t\} : S \subseteq T\}$$

It is not hard to see that for any ring of sets $\mathcal{R}$, $\mathcal{R}_{\preccurlyeq^{\mathcal{R}}} = \mathcal{R}$ and for all quasi-orders $\preccurlyeq$, $\preccurlyeq^{\mathcal{R}_{\preccurlyeq}} = \preccurlyeq$. It then follows from the fact that every ring of sets is an interior system that there is a one-to-one correspondence between rings of sets and quasi-orders. On the contrary, if $\mathcal{I}$ is an interior system but not a ring of sets, then of course $\mathcal{R}_{\preccurlyeq^{\mathcal{I}}} \neq \mathcal{I}$. For any quasi-order $\preccurlyeq$, say that an outcome function g is $\preccurlyeq$-*monotone* if:

$$\forall t_1, t_2 \in T, t_1 \preccurlyeq t_2 \Rightarrow g(t_1) \leq g(t_2),$$

Moreover, for any quasi-ordering, there is a natural message correspondence $M_{\preccurlyeq}$ with range T, satisfying:

$$M_{\preccurlyeq}(t) = \{s \in T : s \preccurlyeq t\} \tag{4.2}$$

That normality is related to transitivity can be seen from its ancestor, the *nested range condition* of Green and Laffont (1986).[1] Among other things, the following theorem provides an alternative characterization of families of implementable outcome functions induced by normal message correspondences.

[1]In the first part of their paper, Green and Laffont (1986) assume that the type space is equal to the message space. The nested range condition is then defined to be the condition that $t_3 \in M(t_2)$ and $t_2 \in M(t_1)$ implies $t_3 \in M(t_1)$. This is essentially a transitivity condition but keep in mind that when writing $t_i \in M(t_j)$, t_i is a message and t_j is a type. The authors also assume that the speaker can tell the truth: for all $t, t \in M(t)$, which is essentially reflexivity. They also provide an example of a message structure of the form (4.2), and state that it implies the nested range condition, although do not specify whether $\preccurlyeq$ is a quasi, partial, or linear order, nor do they make the stronger observation that the set of message structures (with message space T) satisfying the nested range condition and the condition that for all $t, t \in M(t)$ correspond one-to-one with the set of quasi-orders.

Theorem 4.2. *Fix a message structure $(T, \mathbf{M}, M(\cdot))$. Let $\mathcal{I}$ be the induced family of acceptance sets.*

(i) $\forall t_1, t_2, t_1 \preccurlyeq^{\mathcal{I}} t_2 \Leftrightarrow M(t_1) \subseteq M(t_2)$.

(ii) Every implementable outcome function is $\preccurlyeq^{\mathcal{I}}$-monotone.

(iii) Every $\preccurlyeq^{\mathcal{I}}$-monotone outcome function is implementable if and only if $(T, \mathbf{M}, M(\cdot))$ is normal.

(iv) $(T, \mathbf{M}, M(\cdot))$ and $(T, T, M_{\preccurlyeq^{\mathcal{I}}}(\cdot))$ induce the same set of implementable outcome functions if and only if $(T, \mathbf{M}, M(\cdot))$ is normal.

(v) $\mathcal{G}$ is the set of $\preccurlyeq$-monotone outcome functions for some quasi-order $\preccurlyeq$ on T if and only if $\mathcal{G}$ is the set of implementable outcome functions for some normal message structure.

Proof. See Section 4.5.3. □

The notion that under normality, the implementable outcome functions are exactly those in which types with more messages (according to inclusion) get higher actions could have been derived from the more general observation discussed above that in mechanism design environments with provability, under normality, one must only consider incentive constraints involving types who can say more mimicing types who can say less, combined with the particular preference structure in this model. However, Theorem 4.2 says much more. In particular, it says that (for a fixed k), families of implementable outcome functions induced by *normal* message structures correspond one-to-one with the set of quasi-orders on the set of types. Thus under normality, a quasi-order $\preccurlyeq$ on T summarizes *all* exogenous information determining what is implementable. One may assume the message space is T, and that the message correspondence is $M_{\preccurlyeq}(\cdot)$. If the message structure is not normal, then $\preccurlyeq$ does not summarize all information about implementability, but it does provide some information about it. In particular, every implementable outcome function is monotone, but monotonicity is not sufficient for implementability. Thus quasi-orders serve a role which is analogous with respect to normal message structures to the role served by interior systems with respect to arbitrary message structures. So interior systems generalize quasi-orders and are necessary for representing limitations on the ability to summarize information.

If we start from a message structure $(T, \mathbf{M}, M(\cdot))$ violating normality with acceptance sets $\mathcal{I}$, and form the message structure $(T, T, M_{\preccurlyeq^{\mathcal{I}}}(\cdot))$ and consider the set $\mathcal{G}_{\preccurlyeq}$ of $\preccurlyeq^{\mathcal{I}}$-monotone outcome functions, how do these objects relate to the situation with which we started? $(T, T, M_{\preccurlyeq^{\mathcal{I}}}(\cdot))$ is equivalent to the message structure which we would get if we allowed types to send a message corresponding to all their messages, so that we restored the ability to summarize, and $\mathcal{G}_{\preccurlyeq}$ is the set of outcome functions that would be implementable if it was possible for types to summarize. Thus, we can exactly identify which outcome functions are not implementable because of the inability to summarize in the original structure: namely the $\preccurlyeq^{\mathcal{I}}$-monotone functions which do not correspond to any k-chain with elements drawn form $\mathcal{I}$.

4.3. Weak Normality

In this section, I discuss a weaker notion than normality in the case that the speaker may be able to summarize small amounts of information, but he cannot summarize large pieces of information. Here "small" will mean finite and "large" will mean infinite. The definition is as follows:

Definition 4.2. *A signal structure* $(T, M, M(\cdot))$ *is* ***weakly normal*** *if for every for all* $t \in T$*, and all finite* $M_0 \subseteq M(t)$*, there exists a message* $m(M_0, t) \in M(t)$ *such that for all* $s \in T$*, if* $m(M_0, t) \in M(s)$*, then* $M_0 \subseteq M(s)$*.*

The following lemma shows that normality and weak normality coincide when T is finite:

Lemma 4.1. *(i) Normality implies weak normality.*

(ii) If T *is finite, normality and weak normality are equivalent.*

Proof. (i) If normality holds, then for all finite $M_0 \subseteq M(t)$, we can set $m(M_0, t) = m(t)$.

(ii) Suppose that T is finite. Define an equivalence relation on messages such that two messages are equivalent if they are accessible to exactly the same set of types. Notice that if a type has one message in an equivalence class, that type has all messages in that equivalence class. Let $\overline{M}$ be a selection of exactly one message from each equivalence class. If T is finite, then $\overline{M}$ must be finite. Thus $\overline{M} \cap M(t)$ is finite and nonempty. To complete the proof, set $m(t) = m(\overline{M} \cap M(t), t)$. □

Thus, the difference between normality and weak normality only comes out in the case where T is infinite. The following theorem is the analog of Theorem 4.1.

Theorem 4.3. *Let* $\mathcal{I}$ *be the family of acceptance sets induced by* $(T, \mathbf{M}, M(\cdot))$*.* $\mathcal{I}$ *is closed under finite intersection if and only if* $(T, \mathbf{M}, M(\cdot))$ *is weakly normal.*

Proof. See Section 4.5.4. □

Since Theorem 4.1 holds also when T is infinite, the difference between the implications of normality and weak normality for $\mathcal{I}$ when T is infinite is that normality implies that $\mathcal{I}$ is a ring of sets (closed under arbitrary union and *arbitrary* intersection), whereas weak normality only implies that $\mathcal{I}$ is a topology (closed under arbitrary union and *finite* intersection). Employing Theorem 3.2 we can see that every topology corresponds to some message structure so that there is quite a rich set of structures to choose from. One can use this representation to show that in general there are monotone outcome functions which are not implementable under weak normality. Since topologies are not closed under arbitrary intersection, types of speaker will generally not have a unique maximally informative message. The most dramatic illustration of this occurs when assuming a countable number of types, there may be a message structure which can only be generated of the set of messages has the cardinality of the continuum:

Example 4.1. Consider the Cartesian product $\{0,1\}^{2^{\aleph_0}}$ endowed with the product topology. By Theorem 2.3.15 in Engelking (1989), $\{0,1\}^{2^{\aleph_0}}$ has a countable dense set X. Let the set of types $T = X$. Let $\mathcal{I}$ be the induced topology on X. It is shown in Example 2.3.37 of Engelking (1989) that $\mathcal{I}$ does not have a countable base, and a similar argument can be used to show directly that the smallest cardinality of a base of $\mathcal{I}$ is $2^{\aleph_0}$.[2] This means that the minimal cardinality of a join-dense set in $\mathcal{I}$ is $2^{\aleph_0}$, which means that one cannot construct the corresponding message correspondence without using $2^{\aleph_0}$ messages.

The following theorem characterizes normality and weak normality in purely lattice-theoretic terms.

[2]Since $2^{\aleph_0}$ is obviously an upper bound on the cardinality of a base of any topology on a countable set, if one assumes the continuum hypothesis, it follows from the fact that $\mathcal{I}$ does not have a countable base that the minimum cardinality of a base of $\mathcal{I}$ is $2^{\aleph_0}$. What I mean by using a similar argument to show this "directly" is providing an argument which does not invoke the continuum hypothesis.

Theorem 4.4. *Let* $(T, \mathbf{M}, M(\cdot))$ *be a message structure and let* $\mathcal{G}$ *be the induced family of implementable outcome functions. Then:*

(i) $(\mathcal{G}, \leq)$ *is a sublattice of* $(\mathbf{A}^T, \leq)$ *if and only if* $(T, \mathbf{M}, M(\cdot))$ *is weakly normal.*

(ii) $(\mathcal{G}, \leq)$ *is a subcomplete sublattice of* $(\mathbf{A}^T, \leq)$ *if and only if* $(T, \mathbf{M}, M(\cdot))$ *is normal.*

Proof. See Section 4.5.5. □

Several remarks about this theorem are in order. First notice that it follows form Observation 3.2 that the set of implementable outcome functions $(\mathcal{G}, \leq)$ is always a complete lattice. However, as Theorem 4.4 shows, $(\mathcal{G}, \leq)$ is not a subcomplete sublattice of $(\mathbf{A}^T, \leq)$ when weak normality fails, and is not even a sublattice of $(\mathbf{A}^T, \leq)$ when weak normality fails. This is because when weak normality fails the meet of some pair of elements in $\mathcal{G}$ will fail to coincide with the meet of that pair in $\mathbf{A}^T$, and when normality fails the meet of some collection of elements in $\mathcal{G}$ will fail to coincide with the meet of that collection in $\mathbf{A}^T$. Notice, however that an arbitrary sublattice (resp., subcomplete sublattice) of $\mathbf{A}^T$ may not correspond to a weakly normal (resp., normal) family of implementable outcome functions relative to some message correspondence because it may not correspond to *any* family of implementable outcome functions. A final remark concerns an interesting relation between part (ii) of Theorem 4.4 and Theorem 4.2. An alternative proof the proposition that normality implies that the family of implementable outcome functions $\mathcal{G}$ is a subcomplete sublattice of $\mathbf{A}^T$ would be as follows. Part (v) of Theorem 4.2 implies that $\mathcal{G}$ is the subset of $\mathbf{A}^T$ which is monotone with respect to some quasi-order $\preccurlyeq$. Moreover, taking into account that $\mathbf{A}$ is finite, the minimum and maximum of an arbitrary collection of $\preccurlyeq$-monotone functions are also $\preccurlyeq$-monotone.

4.4. Max, Min, Join and Meet

In this section, I will briefly discuss the relations between the operations of componentwise maximum and minimum and the join and meet within the lattice of implementable outcome functions, as well as the effect of applying the operations of componentwise maximum and minimum to persuasion rules to the join and the meet of the implemented outcome functions. The property of normality studied above is relevant to these relations.

Let $\mathcal{G}$ be a family of implementable outcome functions for some message structure. Then as discussed in the previous section $(\mathcal{G}, \leq)$ is a complete lattice where $\leq$ is the componentwise order. Let $\mathcal{H} \subseteq \mathcal{G}$. Then define $\max \mathcal{H}$ to be the componentwise maximum of the elements of $\mathcal{H}$, and likewise define $\min \mathcal{H}$ to be the componentwise minimum of the elements of $\mathcal{H}$. Notice that there is a strong analogy between max with many actions and union with two actions; likewise there is strong analogy between min and intersection. The following theorem confirms this:

Theorem 4.5. *Let $\mathcal{H}$ be a collection of implementable outcome functions. Then for all $t \in T$:*

$$\bigvee \mathcal{H} = \max \mathcal{H} \tag{4.3}$$

$$\bigwedge \mathcal{H} \leq \min \mathcal{H} \tag{4.4}$$

If the message structure is normal, then:

$$\bigwedge \mathcal{H} = \min \mathcal{H} \tag{4.5}$$

Proof. See subsection 4.5.6. □

Equation (4.3) gives a fact for the operation max which is analogous to the fact that an interior system is closed under union. (4.4) gives a property of min which is analogous to the fact that an interior system may not be closed under intersection, so that the meet of a collection of set within an interior system may be a strict subset of the intersection of the sets in that collection. However, under normality, the meet does coincide with the intersection, and the analogous fact for min is presented in (4.5).

Notice that for a family of outcome functions to be closed under max is a weaker property than for the family to be a fixed point of τ, and thus not every family of outcome functions closed under max is the set of implementable outcome functions for some message correspondence. Likewise, closure under max and min is a weaker property than being the implementable family for some normal message correspondence.

Next, we examine the properties of the operations max and min when operating on persuasion rules rather than outcome functions.

Theorem 4.6. *Let $\mathcal{F}$ be a collection of persuasion rules, and define:*

$$f^*(m) := \max\{f(m) : f \in \mathcal{F}\} \tag{4.6}$$

$$f_*(m) := \min\{f(m) : f \in \mathcal{F}\} \tag{4.7}$$

Then:

$$g_{f^*} = \bigvee\{g_f : f \in \mathcal{F}\} \tag{4.8}$$

$$g_{f_*} \leq \bigwedge\{g_f : f \in \mathcal{F}\} \tag{4.9}$$

Proof. See Section 4.5.7. $\square$

The following example shows that (4.9) cannot generally be strengthened to an equality even under normality.

Example 4.2. Suppose that $T = \{t\}, M(t) = \{m_1, m_2\}$ and that there are two actions. Then consider the persuasion rules f_1, f_2 defined by $f_1(m_1) = a_2, f_1(m_2) = a_1, f_2(m_1) = a_1, f_2(m_2) = a_2$. Then $g = g_{f_1} = g_{f_2}$ where $g(t) = a_2$, so $g_{f_1} \wedge g_{f_2} = g$. But defining f_* as in (4.7), $g_{f_*}(t) = a_1$, so $g_{f_*} < g$. $\square$

4.5. Proofs and Examples

4.5.1. Example of Normal Message Structure in Which the Speaker Cannot Summarize Parts of His Information

Suppose that $T = \{t_0, t_1, t_2, t_3, t_4\}, \mathbf{M} = \{m(t_0), m(t_1), m(t_2), m(t_3), m(t_4)\}$, and that:

$$\begin{aligned} M(t_0) &= \{m(t_0), m(t_1), m(t_2), m(t_3), m(t_4)\} \\ M(t_1) &= \{m(t_1)\} \\ M(t_2) &= \{m(t_2)\} \\ M(t_3) &= \{m(t_1), m(t_2), m(t_3)\} \\ M(t_4) &= \{m(t_1), m(t_2), m(t_4)\} \end{aligned}$$

It is easy to verify that the above message structure satisfied normality. The following is a Hasse diagram which only displays the acceptance sets which correspond to questions that only accept a single message, or in other words, acceptance sets of the form $A_{\{m\}}$:

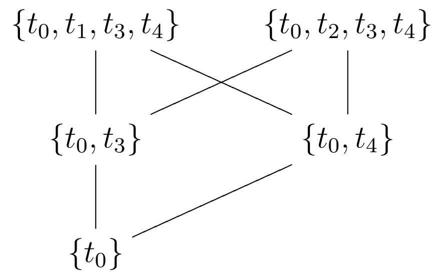

Figure 4.1. Single Message Acceptance Sets

Notice that since type t_0 can send any message, this is also a diagram of the acceptance sets corresponding to one-message questions which t_0 can answer. Thus the sets in this diagram can be thought of as the "meanings" of the messages which type t_0 can use. Notice that these meanings are not closed under intersection. Thus normality does not mean that a type can combine any part of this information, only that a type can summarize all of his information. In particular, t_0 can summarize all of his information by sending message m_0, which corresponds to the set $\{t_0\}$.

4.5.2. Proof of Theorem 4.1

First assume that the message structure is normal. Consider $\mathcal{B} \subseteq \mathcal{I}$. Now, consider a question Q which accepts exactly the messages $\{m(t) : t \in \bigcap \mathcal{B}\}$. Then clearly $\bigcap \mathcal{B} \subseteq A_Q$. On the other hand consider type $s \notin \bigcap \mathcal{B}$. This means that for all $t \in \bigcap \mathcal{B}$, there exists $A \in \mathcal{I}$ such that $t \in A, s \notin A$. So by the definition of $\mathcal{I}$, for all $t \in \bigcap \mathcal{B}$, there exists $m \in M(t)$ such that $m \notin M(s)$. Normality then implies that for all $t \in \bigcap \mathcal{B}$, $m(t) \notin M(s)$. So $s \notin A_Q$. So $\bigcap \mathcal{B} = A_Q$. So $\bigcap \mathcal{B} \in \mathcal{I}$, and $\mathcal{I}$ is closed under intersection.

Next assume $\mathcal{I}$ is closed under intersection. I argue that the message structure is normal. Consider any $t \in T$, and consider the family $\mathcal{D} := \{A \in \mathcal{I} : t \in A\}$. $\mathcal{D}$ is not empty, since $T \in \mathcal{D}$. Closure under intersection implies that $\bigcap \mathcal{D} \in \mathcal{I}$. It follows that there exists some question Q such that $\bigcap \mathcal{D} = A_Q$. Since $t \in A_Q$, there must exist some message $m_* \in M(t)$ such that $m_* \in Q$. I will argue that m_* can be treated as t's maximal message $m(t)$. In particular consider any $s \in T$ such that $m_* \in M(s)$. Then $s \in A_Q = \bigcap \mathcal{D}$. Now consider any $m \in M(t)$. Then $A_{\{m\}} \in \mathcal{D}$. So since $s \in \bigcap \mathcal{D}$, $s \in A_{\{m\}}$. So $m \in M(s)$. So $M(t) \subseteq M(s)$. This validates the identification of m_* and $m(t)$, establishing normality. $\square$

4.5.3. Proof of Theorem 4.2 and Related Facts

First I discuss some facts relating to $\mathcal{R}_{\preccurlyeq}$ and $\preccurlyeq^{\mathcal{R}}$. It is well-known in lattice theory that every partial order corresponds to a ring of sets. Every ring of sets also corresponds to a partial order on its join-irreducible elements. The join-irreducible elements do not exactly correspond to the ground set of the ring of sets, but the join irreducible-elements can be put in one-to-one correspondence with blocks which contain elements of the ground set that always co-occur within the elements of the ring of sets. This explains why we associate rings of sets with quasi-orders rather than with partial orders (which are also anti-symmetric). If we imposed the condition–which is inconvenient for other reasons–that two types cannot have exactly the same set of messages, then quasi-orders would be replaced by partial orders.

Now I prove the parts of Theorem 4.2.

(i): Suppose that $t_1 \preccurlyeq^{\mathcal{I}} t_2$. Then for all $m \in \mathbf{M}$, $t_1 \in A_{\{m\}} \Rightarrow t_2 \in A_{\{m\}}$. So $M(t_1) \subseteq M(t_2)$. On the other hand, if $M(t_1) \subseteq M(t_2)$, then for all $Q \subseteq \mathbf{M}$, $t_1 \in A_Q \Rightarrow t_2 \in A_Q$. So $t_1 \preccurlyeq^{\mathcal{I}} t_2$.

(ii): If $t_1 \preccurlyeq^{\mathcal{I}} t_2$, then in any k-chain $(I_2, \ldots, I_k)$, the highest index set I_h which contains t_1 must also contain t_2.

(iii): First suppose that $(T, \mathbf{M}, M(\cdot))$ is normal. Then consider a monotone outcome function g. Then g is implemented by the k-question such that:

$$Q_j = \{m(t) : g(t) \geq a_j\}$$

To see this consider any type such that $g(t) = a_j$. Then $t \in A_{Q_j}$, but by monotonicity for any $s \in T$, if $g(s) = a_{j+1}$, then $M(s) \setminus M(t) \neq \emptyset$, so $m(s) \notin M(t)$. So t cannot answer Q_{j+1}.

On the other hand, suppose that normality is violated. Then there must be some type t_0 and some collection S of types such that $M(t_0) \subseteq \bigcup_{s \in S} M(s)$ but $M(s) \not\subseteq M(t_0)$ for all $s \in S$. So consider the following monotone outcome function:

$$g(t) = \begin{cases} a_2, & \text{if } t_0 \preccurlyeq^{\mathcal{I}} t; \\ a_1, & \text{otherwise.} \end{cases}$$

This outcome function cannot be implemented because any question Q which accepts t_0 must accept some $s \in S$, but $t_0 \not\preccurlyeq^{\mathcal{I}} s$.

(iv) Notice that $s \preccurlyeq^{\mathcal{I}} t \Leftrightarrow M_{\preccurlyeq^{\mathcal{I}}}(s) \subseteq M_{\preccurlyeq^{\mathcal{I}}}(t)$. This part of the theorem then follows from (i), (ii) and (iii).

(v) Let $\mathcal{G}$ be the set of monotone outcome functions for $\preccurlyeq$, and consider the message structure $(T, T, M_{\preccurlyeq}(\cdot))$, with family of acceptance sets $\mathcal{I}$. Then invoking (i), it is easy to see that $\preccurlyeq = \preccurlyeq^{\mathcal{I}}$, and since $(T, T, M_{\preccurlyeq}(\cdot))$ is normal, it follows from (ii) and (iii) that $\mathcal{G}$ is exactly the implementable family for $(T, T, M_{\preccurlyeq}(\cdot))$. On the other hand, suppose that $\mathcal{G}$ is the set of implementable outcome functions for some message correspondence $(T, \mathbf{M}, M(\cdot))$, and let $\mathcal{I}$ be the corresponding family of acceptance sets. Then by (ii) and (iii), $\mathcal{G}$ is the family of outcome functions which is implementable relative to $\preccurlyeq^{\mathcal{I}}$. $\square$

4.5.4. Proof of Theorem 4.3

Assume that $(T, \mathbf{M}, M(\cdot))$ is weakly normal. I want to prove that $\mathcal{I}$ is closed under finite intersection, for which it is sufficient to prove that it is closed under pairwise intersection. Choose $A_1, A_2 \in \mathcal{I}$. Then there exist $Q_1, Q_2 \subseteq \mathbf{M}$ such that $A_1 = A_{Q_1}, A_2 = A_{Q_2}$. Consider the question $R \subseteq \mathbf{M}$ defined by:

$$R = \{m(\{m_1, m_2\}, t) : m_1 \in Q_1 \cap M(t), m_2 \in Q_2 \cap M(t), t \in A_1 \cap A_2\}.$$

Such a question exists by weak normality. Suppose that $t \in A_1 \cap A_2$. Then there exists $m_1 \in Q_1 \cap M(t)$ and $m_2 \in Q_2 \cap M(t)$. Then since by weak normality $m(\{m_1, m_2\}, t) \in M(t)$, t can answer R. So $A_1 \cap A_2 \subseteq A_R$. Now consider a type $s \notin A_1 \cap A_2$. Then either s has no answer to Q_1 or s has no answer to Q_2. In any event, weak normality implies that for all $m(\{m_1, m_2\}, t) \in R$, $m(\{m_1, m_2\}, t) \notin M(s)$. So $t \notin A_R$. So $A_R \subseteq A_1 \cap A_2$. So $A_R = A_1 \cap A_2$. So $A_1 \cap A_2 \in \mathcal{I}$, implying that $\mathcal{I}$ is closed under finite intersection.

Next assume $\mathcal{I}$ is closed under finite intersection. I want to prove that the message structure is weakly normal. Now choose $t \in T$ and a finite nonempty $M_0 \subseteq M(t)$. Then weak normality implies that there exists $Q \subseteq \mathbf{M}$ such that

$$A_Q = \bigcap\{A_{\{m\}} : m \in M_0\}.$$

Since $M_0 \subseteq M(t)$, $t \in A_Q$. This means that there must be some $m^* \in Q$ such that $m^* \in$

$M(t)$. I will argue that we can treat m^* as $m(M_0, t)$. This means that I want to show that for all $s \in T$, if $M_0 \not\subseteq M(s)$, then $m^* \notin M(s)$. However if $M_0 \not\subseteq M(s)$, then there exists $m \in M_0$ such that $s \notin A_{\{m\}}$. So $s \notin A_Q$. So $m^* \notin M(s)$. $\square$

4.5.5. Proof of Theorem 4.4

The proof requires a lemma:

Lemma 4.2. *If $\mathcal{G}$ is the family of implementable outcome functions for some message correspondence, and let $\mathcal{H} \subseteq \mathcal{G}$. Then define $g^* \in \mathbf{A}^T$ by:*

$$g^*(t) := \max\{g(t) : g \in \mathcal{H}\}.$$

Then $g^* \in \mathcal{H}$.

Proof. Let $\mathcal{I}$ be the corresponding interior system. Then for every $g \in \mathcal{H}$, let $(I_2^g, \ldots, I_n^g)$ be the corresponding n-chain in $\mathbf{C}(\mathcal{I}, n)$. Then notice that $(\bigcup_{g\in\mathcal{H}} I_2^g, \ldots, \bigcup_{g\in\mathcal{H}} I_n^g)$ also belongs to $\mathbf{C}(\mathcal{I}, n)$ and corresponds to g^*. $\square$

Proof of (i). Let $g_1, g_2 \in \mathcal{G}$. Then notice that the join of g_1 and g_2 in $\mathbf{A}^T$ is $\max(g_1, g_2)$, and the meet of g_1 and g_2 in $\mathbf{A}^T$ is $\min(g_1, g_2)$. By Lemma 4.2, $\max(g_1, g_2) \in \mathcal{G}$. Thus, to prove that weak normality implies that $\mathcal{G}$ is a sublattice of $\mathbf{A}^T$, it is sufficient to show that $\min(g_1, g_2)$ belongs to $\mathcal{G}$. So take g_1, g_2, and assume weak normality. Let $(I_2^1, \ldots, I_n^1), (I_2^2, \ldots, I_2^n)$ be the corresponding n-chains in $\mathbf{C}(\mathcal{I}, n)$. Then by Theorem 4.3, $(I_2^1 \cap I_2^2, \ldots, I_n^1 \cap I_n^2) \in \mathbf{C}(\mathcal{I}, n)$. It is easy to see that $(I_2^1 \cap I_2^2, \ldots, I_n^1 \cap I_n^2)$ corresponds to $\min(g_1, g_2)$. So $\min(g_1, g_2) \in \mathcal{G}$.

Now suppose that the message structure is not weakly normal. Then by Theorem 4.3, there exist $I^1, I^2 \in \mathcal{I}$ such $I^1 \cap I^2 \notin \mathcal{I}$. It is then easy to see that the minimum of the outcome functions corresponding to the n-chains $(I^1, \ldots, I^1)$ and $(I^2, \ldots, I^2)$ do not belong to $\mathcal{G}$, implying that $\mathcal{G}$ is not a sublattice of $\mathbf{A}^T$.

The proof of (ii) is similar, except that weak normality is replaced by normality, pairwise minimum is replaced by arbitrary minimum, pairwise intersection is replaced by arbitrary intersection, and Theorem 4.3 is replaced by Theorem 4.1. □

4.5.6. Proof of Theorem 4.5

Notice since $\max \mathcal{H}$ is the join of $\mathcal{H}$ in $\mathbf{A}^T$. Since $\mathcal{G} \subseteq \mathbf{A}^T$, it follows that in order to show (4.3), it is sufficient to establish that $\max \mathcal{H} \in \mathcal{G}$. For each $g \in \mathcal{H}$, let $(I_2^g, \ldots, I_k^g)$ be the corresponding k-chain in $\mathbf{C}(\mathcal{I}, k)$. Then notice that $\max \mathcal{H}$ corresponds to $(\bigcup_{g \in \mathcal{H}} I_2^g, \ldots, \bigcup_{g \in \mathcal{H}} I_k^g)$, since in the latter each type gets at least action a_j if and only if that type gets at least a_j in some $g \in \mathcal{H}$. But $(\bigcup_{g \in \mathcal{H}} I_2^g, \ldots, \bigcup_{g \in \mathcal{H}} I_k^g) \in \mathbf{C}(\mathcal{I}, k)$, so $\max \mathcal{H} \in \mathcal{G}$.

(4.4) follows from the fact that $\min \mathcal{H}$ is the join of $\mathcal{H}$ in $\mathbf{A}^T$, and $\mathcal{G} \subseteq \mathbf{A}^T$.

Notice that since by Theorem 4.1, $\mathcal{I}$ is closed under intersection if and only if the message structure is normal. Hence an argument like that establishing (4.3) with intersection taking the place of union can be used to establish (4.5) when the message structure is normal. □

4.5.7. Proof of Theorem 4.6

Let M_j^f be the set of messages which attain at least action a_j given persuasion rule f.

Then the k-chain corresponding to f^* is:

$$(A_{\bigcup_{f\in\mathcal{F}} M_2^f}, \ldots, A_{\bigcup_{f\in\mathcal{F}} M_k^f}) = (\bigcup_{f\in\mathcal{F}} A_{M_2^f}, \ldots, \bigcup_{f\in\mathcal{F}} A_{M_k^f}),$$

where the right hand side corresponds to $\max\{g_f : f \in \mathcal{F}\}$, which by Theorem 4.5 implies (4.8).

(4.9) follows because for each $f \in \mathcal{F}$, $g_{f_*} \leq g_f$. $\square$

CHAPTER 5

Comparative Statics and Symmetry

In this chapter, I will define the listener's utility directly as a function of the set of implementable outcome functions. I will find properties of the objective, properties of the set of optima, and comparative statics for which normality of the signal structure is critical. This will show qualitatively how the listener's decision is affected by the speaker's ability to summarize information.

The chapter is split into five sections. Section 5.1 develops the main results on properties of the listener's utility function defined directly on the lattice of implementable outcome functions, properties of the set of optima, and comparative statics. The analysis takes place under the assumption that the set of types is finite, and normality is found to be a critical property. Section 5.2 briefly considers the infinite case, and describes how in this case weak normality plays the role which normality plays in the finite case. The significance of the infinite case is also discussed. Section 5.3 shows the significance of the comparative statics derived in Section 5.1 in terms of types for the persuasiveness of individual messages. Section 5.4 shows how normality is critical for the symmetry properties of optimal persuasion rules. This is relevant to the question of when messages are interpreted according to their literal meaning, as opposed to in terms their of conversational meaning, a question addressed by the field of pragmatics. The results complement earlier results of that section complement earlier results by Glazer and Rubinstein (2003, 2004,

2006) by showing that failure of normality is critical for the sort of pragmatic phenomena studied in those papers.

5.1. Properties of the Listener's Utility Function

Throughout this chapter, I assume that both T and $\mathbf{M}$ are finite, and moreover that there is some fixed number n of actions. Let $\mathcal{G}$ be the family of implementable outcome functions induced by some message structure. $(\mathcal{G}, \leq)$ can be considered as a partially ordered set where $g_1 \leq g_2$ if and only if for all t, $g_1(t) \leq g_2(t)$. In other words, $g_1 \leq g_2$ if every type gets a higher action under g_2 than under g_1. We know from Theorem 3.1 that if $\mathcal{I}$ is the family of acceptance sets, the elements of $\mathcal{G}$ correspond one-to-one with the elements of $\mathbf{C}(\mathcal{I}, k)$. Moreover the ordering relation $\leq$ on $\mathcal{G}$ coincides with the ordering relation in terms of difficulty–with smaller elements being more difficult–defined in (3.4) on $\mathbf{C}(\mathcal{I}, k)$, when elements of $\mathcal{G}$ are associated with the corresponding elements in $\mathbf{C}(\mathcal{I}, k)$. It then follows from Observation 3.2 that $(\mathcal{G}, \leq)$ is a lattice. We will go on to define the listener's utility directly on the lattice $\mathcal{G}$. The following standard definitions provide properties of functions defined on a lattice.

Definition 5.1. A function h on a lattice $(X, \leq)$ is

- *modular* if $h(x) + h(y) = h(x \vee y) + h(x \wedge y)$.
- *supermodular* if $h(x) + h(y) \leq h(x \vee y) + h(x \wedge y)$.
- *quasi-supermodular* if

$$h(x \wedge y) \leq h(x) \quad \Rightarrow \quad h(y) \leq h(x \vee y)$$

$$h(x \wedge y) < h(x) \quad \Rightarrow \quad h(y) < h(x \vee y)$$

Notice that modularity implies supermodularity, which in turn, implies quasi-supermodularity.

Next I define the listener's expected utility directly as a function of the implementable outcome functions. In particular for any $g \in \mathcal{G}$, let:

$$V(g; v, \pi) := \sum_{t \in T} v(g(t), t)\pi(t). \tag{5.1}$$

The theorems below will generalize over listener utility functions v in a set $\mathbf{V}$. The theorems will be true if $\mathbf{V}$ is any of the following three sets:

$$\mathbb{R}^{\mathbf{A} \times T} \tag{5.2}$$

$$\{v \in \mathbb{R}^{\mathbf{A} \times T} : \exists u \in \mathbb{S}, \forall t \in T, \exists \text{ concave } c_t \in \mathbb{R}^{\mathbb{R}}, v(\cdot, t) = c_t \circ u\} \tag{5.3}$$

$$\{v \in \mathbb{R}^{\mathbf{A} \times T} : \forall t \in T, \exists \text{ concave } c_t \in \mathbb{R}^{\mathbb{R}}, v(\cdot, t) = c_t \circ u\}, \tag{5.4}$$

where $\mathbb{S}$ is the set for strictly increasing functions from $\mathbf{A}$ to $\mathbb{R}$, and in (5.4), u is any element of $\mathbb{S}$. The theorems in this chapter will generally state the equivalences among several conditions, and the theorems will be strongest when $\mathbf{V}$ is interpreted as either (5.2) or (5.4), depending on the direction of the proof. The reason that (5.3) is of interest is that it shows that theorems of this chapter hold when restricting attention to utility functions satisfying the condition given in Theorem 2.3, guaranteeing that the persuasion rule which is optimal among deterministic persuasion rules can be credibly implemented. The reason that (5.4) is of interest is that it shows that the theorems of this chapter hold when restricting attention to utility functions satisfying the condition given in Theorem 2.2 guaranteeing that there is an optimal deterministic persuasion rule when the speaker's utility function is some fixed $u \in \mathbb{S}$. Likewise, the theorems of this chapter generalize over a set Π of probability measures which can either be $\Delta(T)$ or the elements of $\Delta(T)$ which give each type a positive probability.

For any $v \in \mathbf{V}$ and $\pi \in \Pi$, let:

$$B(v, \pi) = \text{argmax}_{g \in \mathcal{G}} V(g; v, \pi).$$

First I will prove that the speaker's ability to summarize his information can be represented as a property of the listener's objective function.

Theorem 5.1. *Suppose that T is finite. Let $\mathcal{G}$ be the family of implementable outcome functions induced by $(T, \mathbf{M}, M(\cdot))$. The following three statements are equivalent:*

(1) $(T, \mathbf{M}, M(\cdot))$ *is normal.*

(2) *For all $v \in \mathbf{V}$ and $\pi \in \Pi$, $V(g; v, \pi)$ is a modular function of g on $(\mathcal{G}, \leq)$.*

(3) *For all $v \in \mathbf{V}$ and $\pi \in \Pi$, $V(g; v, \pi)$ is a quasi-supermodular function of g on $(\mathcal{G}, \leq)$.*

Proof. See Section 5.5.2. □

This theorem says that under normality, the listener's objective is always a modular function, but when weak normality fails, it is always possible to find a specification of the listener's objective such that it is not even quasi-supermodular. Notice that $V(\cdot; v, \pi)$ is obviously modular on $\mathbf{A}^T$, and therefore on any sublattice of $\mathbf{A}^T$. Theorem 4.4 says that normality is sufficient for $\mathcal{G}$ to be a sublattice of $\mathbf{A}^T$. Given the finiteness of T, normality is also necessary. Given that $\mathcal{G}$ is a sublattice of $\mathbf{A}^T$, the modularity is preserved. When the lattice $\mathcal{G}$ is not a sublattice of $\mathbf{A}^T$, one can specify v and π such that the weaker property of quasi-supermodularity is not preserved. This is the main idea underlying the theorem.

To emphasize that counter-examples to modularity, and moreover, the stronger property of supermodularity are not difficult to generate when normality fails, but on the contrary, are the norm, I present the following theorem, which shows that interesting examples of supermodular listener utility functions may be difficult to come by when the message structure is not normal. Notice, in particular, that by Theorem 4.1, the assumptions on the message structure made by the following theorem are inconsistent with normality.

Theorem 5.2. *Suppose that for all $t \in T$, there exist $m_1, m_2 \in \mathbf{M}$ such that $A_{\{m_1\}} \cap A_{\{m_2\}} = \{t\}$, but for all $t \in T$ and $m \in \mathbf{M}$, $A_{\{m\}} \neq \{t\}$. Then $V(\cdot; v, \pi)$ is supermodular on $\mathcal{G}$ only if the persuasion rule which assigns every type the lowest action is optimal.*

Proof. See Section 5.5.3. □

An example of a message structure which would satisfy the assumptions of the previous theorem would be one in which the speaker observes an element x of some product set $\prod_{i=1}^{\ell} X_i$, and can report at least half of the components of x (both the value x_i and the index i to which it corresponds), but cannot report all of the components. The type of the speaker is then x. Notice that under these assumptions, type x is the only type who can report *both* that the first $\ell/2$[1] components are $(x_1, \ldots, x_{\ell/2})$, and that the last $\ell/2$ components are $(x_{\ell/2}+1, \ldots, x_\ell)$, but it is not possible for type x to report his type. This shows that the hypotheses of the theorem are satisfied. The theorem then says that the

[1]Assume for simplicity that ℓ is even.

listener's utility function can be a supermodular function of his choice set in this case only if the persuasion rule which assigns the lowest action to every message is optimal.

The next theorem characterizes normality in terms of the structure of the set of optima.

Theorem 5.3. *Suppose T is finite. Let $\mathcal{G}$ be the family of implementable outcome functions induced by $(T, \mathbf{M}, M(\cdot))$. The following three conditions are equivalent:*

(1) $(T, \mathbf{M}, M(\cdot))$ *is normal.*

(2) *For every $v \in \mathbf{V}$ and $\pi \in \Pi$, $(B(v, \pi), \leq)$ is a sublattice of $(\mathcal{G}, \leq)$.*

(3) *For every $v \in \mathbf{V}$ and $\pi \in \Pi$, $(B(v, \pi), \leq)$ has a greatest element.*

(4) *For every $v \in \mathbf{V}$ and $\pi \in \Pi$, $(B(v, \pi), \leq)$ has a least element.*

Proof. See Section 5.5.4. □

The proof that (1) implies (2) uses a standard result on maximizing supermodular functions on a lattice to show that under normality, the set of optimal decisions (i.e. outcome functions) for the listener forms a lattice. The greatest element of this lattice corresponds to the optimal decision which is best for the speaker and ant the least element corresponds to the optimal decision which is worst for the speaker. In contrast, when normality fails, the theorem says that one can always find a counter-example, i.e., an objective function for the listener such that the optimal decisions do not form a lattice, and moreover an objective such that the set of maximizers does not have a greatest element, and an objective such that the set of maximizers does not have a least element. In fact, the proof shows that one can construct a single counter-example in which the set of maximizers has neither a greatest nor a least element. This is easiest to explain in the case of two actions. In this case, we know from Theorem 4.1 that the family of sets of types who receive the higher action is not closed under intersection. The proof of the theorem analyzes what must be happening locally in $\mathcal{I}$ where closure under intersection fails. In particular, this analysis shows that it is always possible to link acceptance of some collection of "good" types (i.e., types the listener would like to accept) to a choice among several collections of "bad" types; in order to accept the good types, the listener must accept at least some of these bad types, and in fact the listener's objective can be chosen so that he is indifferent between at least two of the collections of bad types. The listener's desire not to accept both of these collections implies that there is no greatest optimal decision, and his inability to accept the good types without accepting the bad types implies that there is no least optimal decision.

The properties described above have comparative statics consequences. Intuitively, one might expect that if the interests of the speaker and listener become more aligned for every type of speaker, then the listener will ask the speaker an easier question, in the sense that he will choose a persuasion rule which gives each type of speaker a weakly higher action in every state. Formally, consider two specifications of the listener's utility function v_1, v_2, satisfying the following increasing differences[2] relationship:

$$v_1(a_{j+1}, t) - v_1(a_j, t) \leq v_2(a_{j+1}, t) - v_2(a_j, t), \forall t \in T, \forall j = 1, \ldots, k-1 \tag{5.5}$$

If v_1 and v_2 stand in the relation (5.5), we say that the interests of the speaker and listener are *more aligned* given v_2 than given v_1. Of course, in essence, all that matters is whether this relationship holds after v_2 is multiplied by some positive constant. (5.5)

[2]This may be modeled formally as increasing differences between the action a_j and a parameter θ by modeling the listener's utility function as a function $w(a_j, t, \theta)$ parameterized by θ, where there is some ordering relation on the parameter set, and assuming that $w(a_{j+1}, t, \theta) - w(a_j, t, \theta)$ is increasing in θ.

says that the slope of $v_2(\cdot,t)$ is greater than the slope of $v_1(\cdot,t)$ everywhere.[3] This means that in moving from v_1 to v_2, the interests of the speaker and the listener have become more aligned, because on the one hand the speaker always prefers a higher action, and when the listener benefits from choosing a higher action is more beneficial under v_1 than under v_2, and when it is costly for the listener to choose a higher action, then it is less costly under v_2 than under v_1. So one would naturally expect that the listener would ask the speaker an easier question under v_2 than under v_1. However, things are a bit more complicated because the question of whether the speaker can summarize his information turns out to be important for this reasoning.

In order to express this relationship, it is important to note that the there may not be a unique optimal decision for the listener, but rather there may be multiple optimal decisions, where formally, the choice set is viewed as the set of implementable outcome functions rather than as the set of persuasion rules. I treat comparisons between sets of optima in the standard way, in terms of a relation between subsets of a lattice, known as the *strong set order*, and written $\sqsubseteq$. To define this order, consider a lattice $(X,\leq)$, and let $Y_1,Y_2\subseteq X$. Then the strong set order $\sqsubseteq$ is defined by:

$$Y_1 \sqsubseteq Y_2 \Leftrightarrow \forall y_1\in Y_1, \forall y_2 \in Y_2, y_1\wedge y_2\in Y_1, y_1\vee y_2\in Y_2.$$

When Y_1 and Y_2 are singletons, then $\sqsubseteq$ coincides with $\leq$, so that if there are unique optima, in Theorem 5.4 below, $\sqsubseteq$ reduces to $\leq$. To get a sense for $\sqsubseteq$ when the sets being compared are not singletons, consider, for example, the case where the lattice is

[3]Formally, for *any* mapping $q:\mathbf{A}\times T\to\mathbb{R}$ which is increasing in its first argument: $\frac{v_2(a_{j+1},t)-v_2(a_j,t)}{q(a_{j+1},t)-q(a_j,t)}\geq\frac{v_1(a_{j+1},t)-v_1(a_j,t)}{q(a_{j+1},t)-q(a_j,t)}$, so that the "distance" between successive actions is not an issue.

the interval $[0,1]$ with the usual order. Then two subintervals $[\underline{a},\overline{a}]$ and $[\underline{b},\bar{b}]$ satisfy $[\underline{a},\overline{a}] \sqsubseteq [\underline{b},\bar{b}]$ if $\underline{a} \leq \underline{b}$ and $\overline{a} \leq \bar{b}$.

Theorem 5.4. *Suppose T is finite, and let $\mathcal{G}$ be the family of implementable outcome functions induced by $(T, \mathbf{M}, M(\cdot))$.*

(i) If v_1 and v_2 are related as in (5.5), then if $g_2 \in B(v_2,\pi)$, $g_1 \in B(v_1,\pi)$, and $g_2 \leq g_1$, then $g_1 \in B(v_2,\pi)$ and $g_2 \in B(v_1,\pi)$.

(ii) If v_1 and v_2 are related as in (5.5) and $(T, \mathbf{M}, M(\cdot))$ is normal, then $B(v_1,\pi) \sqsubseteq B(v_2,\pi)$.

(iii) If $(T, \mathbf{M}, M(\cdot))$ is not normal, then it is possible to find v_1 and v_2 satisfying (5.5) such that $\forall g_1 \in B(v_1,\pi), \forall g_2 \in B(v_2,\pi)$, $g_1 \not\leq g_2$ and $g_2 \not\leq g_1$.

Proof. See Section 5.5.5.

With regard to the intuitive comparative static described above–namely, that as the interests of the speaker and listener become more aligned for every type of speaker, then the optimal outcome function will be such that the listener always grants the speaker a higher action–the theorem shows that this result depends on the speaker's ability to summarize information. In describing the parts of the theorem, I will talk as if $B(v_1,\pi)$ and $B(v_2,\pi)$ are singletons, so that there is a single optimal decision given both v_1 and v_2, although in general, this need not be the case. Part (i) says that regardless of the properties of the message structure, if the interests of the speaker and listener become more aligned as described above, then the listener will not uniformly give the speaker a lower action. Part (ii) says that if the message structure is normal, then when the interests of the speaker and listener become more aligned, the listener will grant every type of speaker a higher action. Once it is established that the listener's expected utility is a supermodular function of the implementable outcome functions, then part (ii) follows from standard comparative statics results; thus the key insight behind (ii) is actually Theorem 5.1. Finally, part (iii) says that when the message structure is not normal, it is always possible to find a pair of utility functions v_1 and v_2, such that the movement from v_1 to v_2 represents a change in which the interests of the speaker and listener become aligned, but some types of speaker get a higher action under v_2 and others get a higher action under v_1. The basic intuition is that when the listener decides whether to grant a speaker of type t a higher action, then when the speaker cannot summarize his information, the listener must decide on the basis of what information to grant the speaker a higher action, where this decision impacts which other types will receive a higher action as well. It may be the case that the listener's and speaker's interests become more aligned for every type of speaker, but that this change also influences the trade-off that the listener faces in terms of what information to use to grant a certain type a higher action; if the listener decides to request different information of a type t in order to grant him a higher action, this may cause other types, who had access to the old but not the new information, to get a lower action.

The theorem then shows that the ability to summarize information is critical for an intuitive comparative statics result, in the sense that when normality does not hold, it is always possible to find a counter-example. It is natural to ask whether there may be some assumption about the relationship between the listener's utility function and the message structure which would rule out such counter-examples. I now present one such assumption, although it is rather strong. This exercise does however reveal something about the nature of the counter-examples which I found above.

Imagine a heterogenous population of types, some of which have the capacity to summarize their information, and others who do not. There could be several reasons for such a heterogeneity. For example, the time constraints faced by a speaker may stem from his personal circumstances rather than from some external constraint. Alternatively, one might imagine that some types have information that comes in a concise form and others have information that comes in a diffuse form. In principle, the inability to summarize information could be taken as either a positive or negative signal by the listener. I will consider a case in which it is taken as a negative signal. Of course, the listener may or may not be able to infer whether a type has the capacity to summarize from the information that he presents. Formally, I will say that type t *can summarize* if

$$\exists m(t) \in M(t), \forall s \in T, m(t) \in M(s) \Rightarrow M(t) \subseteq M(s) \tag{5.6}$$

Otherwise, type t *cannot summarize*. Notice that condition (5.6) is the same as the condition in Definition 4.1–the definition of normality. The difference is that condition (5.6) applies locally, in the sense that some types may satisfy it and others may not, whereas normality is the condition that all types can summarize. Now consider the following assumption on the listener's preferences:

$$\text{if } t \text{ cannot summarize, then } v(a_{j+1}, t) - v(a_j, t) \leq 0, \forall j = 1, \ldots, k-1. \tag{5.7}$$

This says that when the speaker cannot summarize then the listener and speaker's interests are completely opposed. The model allows both for the possibility that when the speaker cannot summarize, then this can be deduced from the message which he sends, so that some messages are inherently incomplete, and for the possibility that the listener cannot deduce from the message alone, whether or not the speaker can summarize. The latter case is more interesting when assumption (5.7) is imposed, since when the listener recognizes a message as incomplete he can always choose the lowest action.[4] For an example of a simple situation in which the listener cannot infer from the messages alone whether or not the speaker can summarize, assume that $T = \{t_1, t_2, t_3\}$, and $M(t_1) = \{m_1\}, M(t_2) = \{m_2\}, M(t_3) = \{m_1, m_2\}$. Then types t_1 and t_2 can summarize, and type t_3 cannot, but the listener cannot infer from the message that he receives whether or not the speaker can summarize.

Theorem 5.5. *If v satisfies (5.7), then for all π, $V(g; v, \pi)$ is a supermodular function of g on $(\mathcal{G}, \leq)$.*

Proof. See appendix.

It is an immediate consequence of this theorem that in this case, one can establish a comparative statics result as in part (ii) of Theorem 5.4.

Corollary 5.1. *Suppose that v_1 and v_2 both satisfy (5.7). Suppose moreover that v_1 and v_2 are related as in (5.5). Then for all π, $B(v_1, \pi) \sqsubseteq B(v_2, \pi)$.*

This shows that normality is not necessary for the result in (ii) of Theorem 5.4 provided that one makes strong enough assumptions about the listener's utility function, and in particular, about how the listener's utility relates to the message structure.

[4]Implicitly this reasoning relies on the argument presented in Section 3.3 that it is without loss of generality that each type of speaker presents a maximally informative message.

5.2. Infinitely Many Types

The previous section assumed that the set T of types was finite. The main theorem was Theorem 5.1. Theorem 5.1 is, in large part, a consequence of Theorem 4.4, which says that under normality, the lattice $\mathcal{G}$ of implementable outcome functions is a sublattice of $\mathbf{A}^T$. Modularity of the listener's utility function when defined on $\mathbf{A}^T$ is then preserved as we move to $\mathcal{G}$. However, Theorem 4.4 shows that a weaker property then normality, namely, weak normality, is both necessary and sufficient for $\mathcal{G}$ to be a sublattice of $\mathbf{A}^T$ when the set of types is infinite. Recall that normality and weak normality coincide when the set of types is finite. It is therefore worthwhile to ask which of the properties studied in the previous section survive when we consider infinite sets of types.

The reason for considering the infinite case is twofold. The first–and most important–reason is that it shows that results in this chapter do not essentially depend on the fact that under normality, it is without lost of generality that the speaker presents his essentially unique[5] maximally informative message; weak normality does not imply an essentially unique maximally informative message when there are infinitely many types. Rather, the critical fact underlying the results of this chapter is lattice-theoretic and is presented in Theorem 4.4, namely that the set of implementable outcome functions is a *sub*lattice of $\mathbf{A}^T$ *if and only if* the message structure satisfies weak normality.[6] The second reason for considering the infinite case is that it models the interesting situation in which a speaker can summarize small, but not necessarily large, collections of statements.

---

[5]This maximally informative message is "essentially unique" in the sense that under normality, any other maximally informative message m satisfies $A_{\{m\}} = A_{\{m(t)\}}$.

[6]However, as implied by Observation 3.2, the set of implementable outcome functions is always a complete lattice.

Suppose for simplicity that the set T of types is countably infinite. I assume that for all $a \in \mathbf{A}$, $\sum_{t \in T} |v(a,t)|\pi(t) < \infty$, which implies that for all $g \in \mathcal{G}$, $|V(g; v, \pi)| < \infty$.

I will now briefly discuss, which of the results which were presented in the previous section survive in the infinite case when the property of normality is replaced with that of weak normality. To begin with, as mentioned above, Theorem 5.1, which was the main theorem of the previous section survives. Theorem 5.3 partially survives. Certainly weak normality implies that the set of optima is a (possibly empty) sublattice. This corresponds to the implication (1) implies (2). I do not know whether (2) would still imply (1), although I conjecture that it does. The proof of this implication in the finite case depends on the finiteness assumption. One would not expect either (1) or (2) to imply either (3) or (4), when weak normality is substituted in normality, because unlike in the finite case, in the infinite case, a lattice may not have a greatest or least element. Conditions under which the set of maximizers is a subcomplete sublattice–which would imply (3) and (4)–are known, but one would have to apply them within this model. Parts (i) and (ii) of Theorem 5.4 would go through under weak normality in the infinite case. The proof of (iii) in the finite case relies on finiteness, and so it is not clear whether it would be preserved in the infinite case, although I conjecture that it would. Finally, Theorem 5.6–and hence Corollary 5.1–would survive in the infinite case, under weak normality.

5.3. Persuasiveness of Messages

In this section, I translate Theorem 5.4 into a theorem which directly addresses the question of how persuasive individual messages are, and what happens to the persuasiveness of messages as interests become more aligned.

Throughout this section I will use $\mathcal{F}^*(v, \pi)$ to denote the set of optimal persuasion rules given v and π. This notation assumes that the message correspondence held fixed. I sometimes write $\mathcal{F}^*(v)$ when π is assumed to be held fixed.

Definition 5.2. *Let $\mathbf{M}^* \subseteq \mathbf{M}$. The messages in $\mathbf{M}^*$ are **jointly more persuasive** given v_2 than given v_1 if:*

(5.8)

$$\forall f_1 \in \mathcal{F}^*(v_1), \exists f_2 \in \mathcal{F}^*(v_2), \forall m \in \mathbf{M}^*, f_1(m) \leq f_2(m)$$

(5.9)

$$\forall f_2 \in \mathcal{F}^*(v_2), \exists f_1 \in \mathcal{F}^*(v_1), \forall m \in \mathbf{M}^*, f_2(m) \in \{g_{f_2}(t) : m \in M(t)\} \Rightarrow f_1(m) \leq f_2(m).$$

*The messages in $\mathbf{M}^*$ are **jointly less persuasive** given v_2 than given v_1 if the messages in $\mathbf{M}^*$ are (strictly) jointly more persuasive given v_1 than given v_2.*

*A message $m \in \mathbf{M}$ is **strictly more persuasive** given v_2 than given v_1 if:*

$$\max\{f_1(m) : f_1 \in \mathcal{F}(v_1)\} \leq \max\{f_2(m) : f_2 \in \mathcal{F}(v_2)\} \tag{5.10}$$

(5.11)

$$\min\{f_1(m) : f_1 \in \mathcal{F}(v_1), f_1(m) = g_{f_1}(t)\} \leq \min\{f_1(m) : f_2 \in \mathcal{F}(v_1), f_2(m) = g_{f_2}(t)\},$$

*where either (5.10) or (5.11) holds with a strict inequality. m is **strictly less persuasive** given v_2 than given v_1 if m is strictly more persuasive given v_1 than given v_2.*

(5.8) of the above definition says that starting with v_1, and a persuasion rule f_1 which is optimal given v_1, it is possible to find a persuasion rule f_2 which is optimal given

v_2 such that all messages in $\mathbf{M}^*$ get the speaker a weakly higher action under f_2 than under f_1. (5.9) says that starting from f_2 which is optimal under v_2, it is possible to find a persuasion rule f_1 which is optimal under v_1 and such that all messages $m \in \mathbf{M}^*$ which *would have been optimal for any speaker type to use* under v_1, gets a weakly lower action under f_1 than under f_2. In other words, any message whose use might have been observed under f_2 gets a weakly lower action under f_1. An example below will show why this qualification is made in (5.9).

To any persuasion rule f, there corresponds a persuasion rule f' such that (i) f and f' are equivalent in the sense that they implement the same outcome function, and (ii) every message (which is available to some type) is a best response to f' for some type. Thus if we restrict attention to persuasion rules with property (ii), then the qualification in (5.9) can be removed.

Note that (5.10) and (5.11) give similar definitions for individual messages to be *strictly* more persuasive.

The following lemma is useful.

Lemma 5.1. *Suppose that $g \in \mathcal{G}$. Then define a persuasion rule f by:*

$$f(m) := \min\{g(t) : m \in M(t)\}.[7] \tag{5.12}$$

Then $g = g_f$. Moreover, for all f', if $g_{f'} = g$, then for all $m \in \bigcup_{t \in T} M(t)$, $f'(m) \leq f(m)$.

Proof. See Section 5.5.7

[7] The value of f on $m \in \mathbf{M} \setminus \bigcup_{t \in T} M(t)$ is clearly irrelevant.

In fact, for any implementable outcome function g, the function f defined in (5.12) is the persuasion rule implementing g which has the property (ii) described above, or in other words, which makes every message which is available to some type a best response to f for some type. The following theorem uses the preceding lemma to provide a comparative statics result for the persuasiveness of messages under normality.

Theorem 5.6. *Assume that the message structure is normal. Then the messages in* $\mathbf{M}$ *become jointly more persuasive as interests become more aligned.*

Proof. See Section 5.5.8. □

The following example shows why the qualification in (5.9) in the definition of "more persuasive" messages is necessary if one would like to prove a theorem like Theorem 5.6.

Example 5.1. Suppose that $k = 2, T = \{t_1, t_2\}, M(t_1) = \{m_1, m_2\}, M(t_2) = \{m_2\}$. Notice that the message structure is normal. Suppose that $\pi(t) = 1/2$ for $t \in \{t_1, t_2\}$. Moreover, suppose that v_1 and v_2 are given by:

v_1	a_1	a_2
t_1	*1*	*0*
t_2	*0*	*1*

v_2	a_1	a_2
t_1	*0*	*1*
t_2	*0*	*1*

Then interests are more aligned under v_2 than under v_1, but notice that f_2 defined by $f_2(m_1) = a_2, f_2(m_2) = a_1$ is an optimal persuasion rule given v_2. On the other hand, the unique optimal persuasion rule f_1 such that $f_1(m_1) = a_1, f_2(m_2) = a_2$ is the unique optimal persuasion rule given v_1. This means that there does not exist an optimal persuasion rule given v_1 which assigns m_2 a weakly lower action than does f_2. This shows that (5.9) in Definition 5.2 could not have been strengthened to be:

$$\forall f_2 \in \mathcal{F}^*(v_2), \exists f_1 \in \mathcal{F}^*(v_1), \forall m \in \mathbf{M}^*, f_1(m) \leq f_2(m),$$

if we would like to prove Theorem 5.6. Nevertheless, notice that as implied by Theorem 5.6, it would never be optimal for any speaker type to use m_2 given f_2. Moreover, f_2 which assigns the highest action a_2 to all messages is optimal given v_2.

Theorem 5.7. *Assume that the message structure is not normal. Then there exist listener utility functions v_1, v_2 and message m such that:*

(1) *Preferences are more aligned given v_2 than v_1.*

(2) *m is strictly less persuasive given v_2 than v_1.*

Proof. See Section 5.5.9. □

The following informal example shows intuitively why some messages may become less persuasive as interests become more aligned.

Example 5.2. Suppose that the listener must decide on whether to undertake one of two projects, A or B. A third alternative is to stay with a status quo option. The speaker prefers the status quo.[8] The listener would like to base his decision on the testimony of two experts. Each expert has a definite opinion as to which project is better. If both experts agree as to which project is better, then the listener would like to select this better project. However, if the experts disagree, then the listener would prefer to take the status quo option.

[8]It does not matter what the speaker's preference is with respect to the comparison of the projects A and B.

Suppose that only the speaker, and not the listener, has access to the experts' testimonies. However there is only time for the speaker to present the testimony of one expert.[9] Thus the message structure is not normal. There are four possible states, corresponding to the possible pairs of opinions of the two experts. Suppose that each of these states is equiprobable, and that the listener gets a utility of 1 if he takes the right action in a given state, and 0 if he takes the wrong action, so that the listener would like to choose a persuasion rule which minimizes the number of mistakes.

There are two optimal persuasion rules in this example:

(1) Take the status quo action if speaker proves that one expert supports A. Otherwise take B.

(2) Take the status quo action if speaker proves that one expert supports B. Otherwise take A.

Now, suppose that we start with persuasion rule (1). Next suppose that the benefit of project B goes down when B is optimal. This amounts to an increased alignment of interests.[10] But now notice that persuasion rule (1) is no longer optimal. To see this, notice that at an optimal persuasion rule, the listener must not undertake the project at some state of the world where it would be worthwhile. Given that the listener must make such a mistake, he would prefer to make this mistake when the inferior project would be optimal. However under persuasion rule (1), he does just the opposite. In particular, notice that according the the definition of persuasiveness of messages, the message that one of the experts supports B becomes less persuasive.

[9] I assume for definiteness that the speaker can present both the opinion of one expert along with that expert's identity. However, it does not matter for the purpose of the example whether or not the speaker can present the expert's identity.

[10] This can be made rigorous in terms of the relation given in (5.5).

5.4. Symmetry

In this section, I present results which are relevant to endogenous interpretation of messages in a persuasive situation. Part of the motivation of Glazer and Rubinstein (2003, 2004, 2006) comes from the field of pragmatics. Pragmatics studies conversational meaning, or in other words, the meaning that words and sentences acquire in conversation over and above their literal or semantic meanings. Glazer and Rubinstein point out that the classical account of conversational implicature proposed by Grice (1989) assumes cooperative behavior on the part of the participants to the conversation. On the other hand, conversational implicature is possible also in strategic settings, so Grice's account cannot be complete. Glazer and Rubinstein propose a game theoretic account which is designed to handle conversational implicature in a simple strategic interaction.

In this section, I use the tools developed previously to further analyze what is necessary for conversational implicature in strategic interactions. I do so by means of the concept of symmetry. Roughly, two messages are symmetric if from a structural point of view, they cannot be told apart. In other words, the messages can only be distinguished if we have labels, but are structurally indistinguishable. I treat a pair of symmetric messages as having the same literal meaning. The question then becomes: under what conditions must messages with the same literal meanings be treated differently in an optimal persuasion rule. Glazer and Rubinstein have found examples in which the optimal rule treats what I here refer to as symmetric messages differently. The contribution here is threefold; first, I formally define the notion of symmetry. Secondly, I show that under normality there always exists a symmetric optimal persuasion rule. Thirdly, I show that at least within a certain class of examples, when normality fails, it is possible to find persuasion situations in which *all* optimal rules are asymmetric. Once the definition of symmetry is presented it is not hard to come to the third point from an examination of some of Glazer and Rubinstein's examples, and in fact this can be understood to be related to the main point of a slightly different model in Glazer and Rubinstein (2003). Nevertheless, Glazer and Rubinstein (2003) only presents an example and they do not present the general formal definition of symmetry found here, nor do they identify the role of normality. The second contribution is important because it shows that time constraints or more generally some departure from the case of normality is necessary to generate conversational meaning, or in other words, essentially different treatment of messages with the same literal meaning in settings such as those studied here.

Definition 5.3. *A pair of bijections* (φ, ψ) *where* $\varphi : T \to T$ *and* $\psi : \mathbf{M} \to \mathbf{M}$ *is a* ***symmetry*** *if:*

$$M(\varphi(t)) = \{\psi(m) : m \in M(t)\} \tag{5.13}$$

$$v(a_i, \varphi(t)) = v(a_i, t) \tag{5.14}$$

$$\pi(\varphi(t)) = \pi(t) \tag{5.15}$$

Let Φ *be the set of all symmetries.*

Thus a symmetry is a pair of mappings, one from types to types, and the other from messages to messages, which preserve all structural features. If $m' = \psi(m)$ for some symmetry, then the messages m and m' cannot be distinguished without labels.

Not surprisingly, the set of symmetries is a group:

Lemma 5.2. Φ *is a group under the operation* $\bullet$ *defined by:*

$$(\varphi_1, \psi_1) \bullet (\varphi_2, \psi_2) := (\varphi_1 \circ \varphi_2, \psi_1 \circ \psi_2)$$

with inverse

$$(\varphi, \psi)^{-1} = (\varphi^{-1}, \psi^{-1})$$

and identity $(\varphi_{\mathbf{I}}, \psi_{\mathbf{I}})$ *with* $\varphi_{\mathbf{I}}(t) = t, \psi_{\mathbf{I}}(m) = m$.

Proof. See Section 5.5.10.

The following theorem states some useful properties of symmetries.

Theorem 5.8. *Suppose that* (φ, ψ) *is a symmetry.*

(1) $g_f \circ \varphi = g_{f \circ \psi}$

(2) $V(g_f; v, \pi) = V(g_{f \circ \psi}; v, \pi)$

(3) *If* f *is an optimal persuasion rule, then so is* $f \circ \psi$.

Proof. See Section 5.5.11. □

(1) says that if (φ, ψ) is a symmetry, and f is a persuasion rule, then composing the outcome function implemented by f with φ leads to the same outcome function as composing f with ψ to arrive at a new persuasion rule, and then taking the outcome function implemented by this new persuasion rule. As a consequence, if g is any implementable outcome function, so is $g \circ \varphi$, as one would expect. (2) says that the listener's payoffs to f and $f \circ \psi$ are the same. (3) is an immediate consequence of (3), and says that the set of optimal persuasion rules is closed under composition with ψ for any symmetry (φ, ψ).

The next definition uses the notion of a symmetry defined above to define a notion of symmetry which applies to persuasion rules.

Definition 5.4. *A persuasion rule f is* ***symmetric*** *if for every symmetry (φ, ψ):*

$$f = f \circ \psi.$$

Theorem 5.9. *For any symmetric persuasion rule f and any symmetry (φ, ψ),*

$$g_f = g_f \circ \varphi. \tag{5.16}$$

Proof. Choose any symmetry (φ, ψ). Since f is symmetric, $f = f \circ \psi$. (5.16) then follows from (1) of Theorem 5.8. □

The next theorem shows that pragmatic phenomena cannot arise under normality. In other words, under normality, there is always an optimal persuasion rules which treats every pair of symmetric messages in the same way. Thus persuasiveness of such messages can be interpreted as being a function only of their content.

Theorem 5.10. *Suppose that the message structure is normal. Then there is a symmetric optimal persuasion rule.*

Proof. Let f be any optimal persuasion rule. Then consider the rule:

$$f^*(m) = \max\{f \circ \psi(m) : (\varphi, \psi) \in \Phi\} \tag{5.17}$$

It follows from part (3) of Theorem 5.8 that $g_{f \circ \psi}$ is an optimal implementable outcome function for each $(\varphi, \psi) \in \Phi$, and moreover by Theorem 4.6,

$$g_{f^*} = \bigvee\{g_{f\circ\psi} : (\varphi, \psi) \in \Phi.\}$$

It follows from the normality of the message structure and Theorem 5.3 that g_{f^*} is an optimal implementable outcome function.

It remains only to show that f^* is a symmetric persuasion rule. For for any $(\varphi', \psi') \in \Phi$ and any $m \in \mathbf{M}$:

$$\begin{aligned}
f^* \circ \psi'(m) &= \max\{f \circ \psi(\psi'(m)) : (\varphi, \psi) \in \Phi\} && (5.18)\\
&= \max\{f \circ \psi''(m) : (\varphi'', \psi'') \in \Phi \bullet (\varphi', \psi')\} && (5.19)\\
&= \max\{f \circ \psi''(m) : (\varphi'', \psi'') \in \Phi\} && (5.20)\\
&= f^*(m) && (5.21)
\end{aligned}$$

where (5.18) and (5.21) follow from (5.17), and (5.20) follows from the fact that from Lemma 5.2, Φ is a group under $\bullet$, so that $\Phi \bullet (\varphi', \psi') = \Phi$. $\Box$

Example 5.3. Assume that $T = [0, 1]^n$. Suppose that any type $t = (t_1, \ldots, t_n)$ can show any h (or fewer) components of t. More formally, for any $J \subseteq \{1, \ldots, n\}$ and $t \in T$, define:

$$t_J = \{(t_j, j) : j \in J\},$$

and suppose that $M(t) := \{t_J : |J| \leq h\}$. Thus, t can reveal both the value and the index of at most h components, where moreover, it is assumed that $1 \leq h < n$. This means that t can show at least 1 but not all of his components. Note that all types face the same bound h on the number of components that they can show. Suppose that there are two actions $\mathbf{A} = \{a_1, a_2\}$, and the listener's utility function takes the following form:

$$\begin{aligned} v(a_2, t) &= w(\sum_i t_i) \\ v(a_1, t) &= 0 \end{aligned}$$

where w is increasing and for some ℓ with $h < \ell \leq n$:

$$w(\ell - 1) < 0 < w(\ell).$$

In other words, there is some critical number ℓ such that if the speaker has at least ℓ components equal to 1, then the listener would like to take the speaker's preferred action a_2, and otherwise, the listener would like to take action a_1.

Claim 5.1. *In the above non-normal persuasion problem, if it is not optimal to reject every message, then every optimal persuasion rule is asymmetric.*

Proof. Assume that f is a symmetric optimal rule which accepts some message $m \in \bigcup_{t \in T} M(t)$; that is, $f(m) = a_2$. By optimality, there must be some t^* with $\sum_i t_i^* =: p \geq \ell$ and $J \subseteq \{1, \ldots, n\}$, $|J| \leq h$ with $f(t_J^*) = a_2$. In other words, optimality implies that if f accepts some message $m \in \bigcup_{t \in T} M(t)$, there must be some type who has that message who the listener would like to accept. Note that for every q with $h \leq q < p$, there exists $s^q \in T$ with $\sum_i s_i^q = q$ and $s_J^q = t_J^*$. By symmetry, f accepts some message from every type such that $h \leq \sum_i t_i \leq p$. It follows that the optimal rule f is unique, and accepts exactly messages of the form $\{(1, j) : j \in J\}$ with $|J| = h$. But the rule f' which accepts every message accepted by f except $\{(1,j) : 1 \leq j \leq h\}$ would do better, contradiction. $\square$

One might want to be able to show that whenever the message structure is not normal, then one can specify the utility functions in such a way that there does not exist a symmetric optimal rule. However, this might not be possible for a rather trivial reason, namely, it might be that the only symmetry is $(\varphi_{\mathbf{I}}, \psi_{\mathbf{I}})$, that is the pair of identity maps. Then there will exist a symmetric optimal rule regardless of the specification of the listener's utility and regardless of whether the message structure is normal. The following theorem tries to generalize the observations relating to symmetry to message structures in which the only symmetry is the trivial one. In the counter-examples to the existence of symmetric optimal rules when the message structure is not normal, there are collections of messages $\{m_1, m_2, \ldots, m_n\}$ such that for every pair of messages m_i, m_j, there exists a symmetry $(\varphi^{i,j}, \psi^{i,j})$ with $\psi^{i,j}(m_i) = m_j$ but such that no optimal rule treats them all in the same way. Suppose for example that every optimal rule treats m_1 and m_2 differently, in that $f(m_1) = a_1$ and $f(m_2) = a_2$. Of course, it is possible to find another optimal persuasion rule f' such that $f'(m_1) = a_2$ and $f'(m_2) = a_1$. But there may be no optimal rule which assigns either a_1 or a_2 to both m_1 and m_2. Thus the treatment of m_1 and m_2 is interdependent at the optimal rule. Such interdependence is impossible for any pair of messages, regardless of whether they are symmetric or not, under normality. However, one can always specify the listener's utility function so that such interdependence emerges for some pair of messages whenever normality fails:

Theorem 5.11. *The following conditions are equivalent:*

(1) *The message structure is not normal.*

(2) *There exists listener utility function v and probability distribution π, and a pair of messages m_1, m_2 such that the following conditions hold* ***non-vacuously***

$$\forall f \in \mathcal{F}(v,\pi), f(m_1) = a_k \;\; \Rightarrow \;\; f(m_2) < a_k \tag{5.22}$$

$$\forall f \in \mathcal{F}(v,\pi), f(m_2) = a_k \;\; \Rightarrow \;\; f(m_1) < a_k \tag{5.23}$$

Proof. In the course of the proof of Theorem 5.3 in Section 5.5.4, it was shown that whenever the message structure is not normal, it is possible to construct a utility function v and probability distribution π such that for some $S \subseteq T$ with $|S| \geq 2$:

$$g \in B(v,\pi) \Rightarrow \exists I \in \mathcal{I}, |I \cap S| = 1, \text{ and } g = g^I,$$

and moreover,

$$S \subseteq \{I : g^I \in B(v,\pi)\}$$

So for $t_1, t_2 \in S$ with $t_1 \neq t_2$, there exists $m_1 \in M(t_1), m_2 \in M(t_2), f_1, f_2 \in \mathcal{F}(v)$ such that $f_1(m_1) = a_k, f_2(m_2) = a_k$. (5.22) and (5.23) hold non-vacuously for m_1 and m_2. □

To summarize, normal persuasion problems differ from non-normal persuasion problems in that within non-normal problems, the persuasiveness of messages may be interdependent within the class of optimal rules, whereas in normal problems, this is impossible. In particular, in the case of non-normal problems there may be interdependence between symmetric messages, causing them to be treated differently. Thus all optimal persuasion rules in non-normal problems may be asymmetric.

5.5. Proofs

5.5.1. Preliminaries

The following notation will be very useful in the proofs that follow. For any $S \subseteq T$, and $j = 2, \ldots, k$ define the $g_j^S \in \mathbf{A}^T$ by:

$$g_j^S(t) := \begin{cases} a_j, & \text{if } t \in S; \\ a_{j-1}, & \text{otherwise.} \end{cases} \tag{5.24}$$

I write g^S for g_k^S. The rest of this section is phrased in terms of g^S, but all facts will continue to hold if g^S were replaced by g_j^S. Let $\mathcal{I}$ be the family of acceptance sets and $\mathcal{G}$ be the family of implementable outcome functions. Then for any $B \in \mathcal{I}$ $(T, \ldots, T, B) \in \mathbf{C}(n, k)$, and $(T, \ldots, T, B)$ corresponds to g^B. This implies that:

$$B \in \mathcal{I} \Leftrightarrow g^B \in \mathcal{G}. \tag{5.25}$$

Note also that for any $B, C \in \mathcal{I}$:

$$g^B \wedge g^C = g^{B \wedge C} \tag{5.26}$$

$$g^B \vee g^C = g^{B \vee C} = g^{B \cup C}, \tag{5.27}$$

where on the left hand side, $\wedge$ and $\vee$ are evaluated within $\mathcal{G}$ and on the right hand side, $\vee$ and $\wedge$ are evaluated within $\mathcal{I}$.

5.5.2. Proof of Theorem 5.1

First I prove that (1) implies (2). Observe that $V(g; v, \pi) := \sum_{t \in T} v(g(t), t)\pi(t)$ is modular on the lattice $\mathbf{A}^T$. It follows from (i) of Theorem 4.4 that normality implies that $\mathcal{G}$ is a sublattice of $\mathbf{A}^T$. This implies that modularity is preserved moving from $\mathbf{A}^T$ to $\mathcal{G}$.

Notice next it is immediate that 2 implies 3 because modularity is a stronger property than quasi-supermodularity.

I will complete the proof by showing that when normality fails, it is possible to find v and π such that $V(\cdot; v, \pi)$ is not quasi-supermodular on $\mathcal{G}$.

By Theorem 4.1, if $(T, \mathbf{M}, M(\cdot))$ is not normal, then $\mathcal{I}$ is not closed under intersection. It then follows from the finiteness of T that there exist $B, C \in \mathcal{I}$ such that $B \cap C \notin \mathcal{I}$. This implies that $B \nsubseteq C$ and $C \nsubseteq B$. In particular, there exists $t_1 \in B \setminus C$. Given that $B \cap C \notin \mathcal{I}$, $B \wedge C$ must be a proper subset of $B \cap C$. This means that there exists $t_0 \in (B \cap C) \setminus (B \wedge C)$. Since $t_0 \in C, t_1 \notin C$, it follows that $t_0 \neq t_1$.

Next, choose $\pi \in \Pi$ with $\pi(t) > 0$ for all $t \in T$. Assume, next that $v(a_{k-1}, t) = 0$ for all $t \in T$. Next, assume that $v(a_k, t_1) < 0, v(a_k, t_0)\pi(t_0) > -v(a_k, t_1)\pi(t_1)$, and for all $t \in T \setminus \{t_0, t_1\}, v(a_k, t) = 0$. Notice that all of these assumptions on v are consistent with v belonging to (5.4) for any $u \in \mathbb{S}$.

By (5.25), $g^B, g^C \in \mathcal{G}$. So:

$$
\begin{aligned}
&(5.28) \qquad V(g^B \wedge g^C; v, \pi) - V(g^B; v, \pi) &&= V(g^{B\wedge C}; v, \pi) - V(g^B; v, \pi) \\
&(5.29) &&= \sum_{t \in B \setminus (B \wedge C)} (v(a_{k-1}, t) - v(a_k, t))\pi(t) \\
&(5.30) &&= -\sum_{t \in B \setminus (B \wedge C)} v(a_k, t)\pi(t) \\
&(5.31) &&= -[v(a_k, t_0)\pi(t_0) + v(a_k, t_1)\pi(t_1)] \\
&(5.32) &&< 0
\end{aligned}
$$

where (5.28) follows from (5.26), (5.29) follows from the fact that g^B and $g^{B\wedge C}$ agree everywhere except on $B \setminus (B \wedge C)$, (5.30)-(5.32) follow from the assumptions on v and π.

On the other hand,

$$
\begin{aligned}
(5.33)\qquad V(g^C; v, \pi) - V(g^B \vee g^C; v, \pi) &= V(g^C; v, \pi) - V(g^{B\cup C}; v, \pi) \\
(5.34)\qquad &= \sum_{t \in B\setminus C} (v(a_{k-1}, t) - v(a_k, t))\pi(t) \\
(5.35)\qquad &= -v(a_k, t_1)\pi(t_1) \\
(5.36)\qquad &> 0,
\end{aligned}
$$

where (5.33) follows from (5.27), (5.34) follows from the fact that g^C and $g^{B\cup C}$ only differ on $B \setminus C$, and (5.35)-(5.36) follow from the assumptions on v and π.

Finally, note that (5.28)-(5.32) and (5.33)-(5.36) are together inconsistent with quasi-supermodularity of $V(\cdot; v, \pi)$. $\square$

5.5.3. Proof of Theorem 5.2

Consider $t \in T$, and $m_1, m_2 \in \mathbf{M}$ such that $A_{\{m_1\}} \cap A_{\{m_2\}} = \{t\}$. Let $I^1 = A_{\{m_1\}}, I^2 = A_{\{m_2\}}$. Suppressing v, π in $V(\cdot; v, \pi)$, notice that for any j such that $2 \leq j \leq k$, it is easy to see that $\mu_j : 2^T \to \mathbb{R}$ defined by

$$(5.37)\qquad \mu_j(S) = V(g_j^S) - V(g_j^{\emptyset}) = \sum_{t \in S} [v(a_j, t) - v(a_{j-1}, t)]\pi(t)$$

is a signed measure on 2^T. So:

$$(5.38)\qquad \mu_j(I^1) + \mu_j(I^2) = \mu_j(I^1 \cup I^2) + \mu_j(I^1 \cap I^2).$$

Invoking (5.26) and (5.27), supermodularity of V implies:

$$V(g_j^{I^1}) + V(g_j^{I^2}) \leq V(g_j^{I^1 \cup I^2}) + V(g_j^{I^1 \wedge I^2}), \tag{5.39}$$

Subtracting $2V(g_j^{\emptyset})$ from both sides of (5.39), we arrive at:

$$\mu_j(I^1) + \mu_j(I^2) \leq \mu_j(I^1 \cup I^2) + \mu_j(I^1 \wedge I^2).$$

Now, using (5.38) this is equivalent to:

$$\mu_j(I^1 \cap I^2) \leq \mu_j(I^1 \wedge I^2),$$

But by the assumptions of the theorem $I^1 \cap I^2 = \{t\}, I^1 \wedge I^2 = \emptyset$, so this is equivalent to $\mu_j(\{t\}) \leq 0$, and using the definition of μ_j, (5.37),

$$\sum_{t \in T} [v(a_j, t) - v(a_{j-1}, t)]\pi(t) \leq 0.$$

Since t and j were arbitrary (provided $j \geq 2$), it follows that the persuasion rule which assigns the lowest action to every message is optimal. □

5.5.4. Proof of Theorem 5.3

That (1) implies (2) follows from Theorem 5.1 and the fact that the set of maximizers of a supermodular function on a lattice is itself a sublattice. (3) and (4) both follow from (2) because every finite lattice has a greatest and least element.

To complete the proof, it is sufficient to show that if $(T, \mathbf{M}, M(\cdot))$ is not normal, then there exist v and π such that the induced set of maximizers has neither a greatest nor a least element.

Lemma 5.3. *Let X be a finite set and let $\mathcal{Y} \subseteq 2^X \setminus \emptyset$ be an antichain with $|\mathcal{Y}| \geq 2$. Then there exists $Z \subseteq X$ such that for all $Y \in \mathcal{Y}$, $Z \cap Y \neq \emptyset$, and there exist $Y_1, Y_2 \in \mathcal{Y}$ such that $Y_1 \cap Y_2 \cap Z = \emptyset$ and $|Y_1 \cap Z| = |Y_2 \cap Z| = 1$.*

Proof. Start with a set $Z_0 = \emptyset$, and $\mathcal{Y}_0 = \mathcal{Y}$, and repeatedly perform the following steps, which transform sets Z_i and $\mathcal{Y}_i$ into Z_{i+1} and $\mathcal{Y}_{i+1}$.

(1) If $|\mathcal{Y}_i| > 1$, then choose some $x_i \in \bigcup \mathcal{Y}_i \setminus \bigcap \mathcal{Y}_i$, and set $Z_{i+1} = Z_i \cup \{x_i\}, \mathcal{Y}_{i+1} = \mathcal{Y}_i \setminus \{Y \in \mathcal{Y}_i : x_i \in Y\}$.

(2) If $|\mathcal{Y}_i| = 1$, stop.

To see that this procedure is well-defined, notice that if $|\mathcal{Y}_i| > 1$, then $|\mathcal{Y}_{i+1}| \geq 1$. Clearly, this procedure terminates in finite time. Suppose that the procedure terminates at step ℓ. Notice that both $\mathcal{Y}_{\ell-1} \setminus \mathcal{Y}_\ell$ and $\mathcal{Y}_\ell$ are nonempty. $\mathcal{Y}_\ell$ contains a single element Y_1. Choose any $Y_2 \in \mathcal{Y}_{\ell-1} \setminus \mathcal{Y}_\ell$. Since $\mathcal{Y}$ is an antichain, then there exists $x_\ell \in Y_1 \setminus Y_2$. Now set $Z = Z_\ell \cup \{x_\ell\}$. Notice that $Y_1 \cap Z = \{x_\ell\}, Y_2 \cap Z = \{x_{\ell-1}\}$, and $x_\ell \neq x_{\ell-1}$, so Y_1, Y_2, and Z satisfy all the properties required by the lemma. $\square$

If $(T, \mathbf{M}, M(\cdot))$ is not weakly normal, by Theorem 4.3, there exists $B, C \in \mathcal{I}$ such that $B \not\subseteq C, C \not\subseteq B, B \vee C = B \cup C$, and $B \wedge C$ is a strict subset of $B \cap C$. Let:

$$\mathcal{A} := \{A \in \mathcal{I} : B \cap C \subseteq A\}$$

Notice that since $B \cap C \notin \mathcal{I}$, $B \cap C \notin \mathcal{A}$. Moreover, since $B \in \mathcal{A}$, $\mathcal{A}$ is not empty. Therefore, since $\mathcal{I}$, and hence $\mathcal{A}$ is finite, there must be a nonempty subset $\mathcal{D}$ of $\mathcal{A}$ of minimal elements of $\mathcal{A}$. Notice that since $B, C \in \mathcal{A}$, there must be $D, E \in \mathcal{D}$ such that $D \subseteq B$ and $E \subseteq C$. Moreover, any set A in $\mathcal{I}$ which is a subset of both B and C is a proper subset of $B \cap C$ because $B \cap C \notin \mathcal{I}$. It follows that since $D, E \in \mathcal{A}$ and hence are supersets of $B \cap C$, $D \neq E$. Therefore, $\mathcal{D}$ contains at least two elements. Moreover, notice that $\mathcal{D}$ is an anti-chain and that every element of $\mathcal{D}$ contains $B \cap C$ as a proper subset. It follows that the set:

$$\mathcal{F} := \{A \setminus (B \cap C) : A \in \mathcal{D}\}$$

is an anti-chain containing at least two elements. Therefore, by Lemma 5.3, there is a set $Z \subseteq \bigcup \mathcal{F}$ such that for all $A \in \mathcal{F}, A \cap Z \neq \emptyset$, and there exist two distinct sets $F, G \in \mathcal{F}$ such that $F \cap Z = \{t_1\}, G \cap Z = \{t_2\}$ where $t_1 \neq t_2$.

Letting $|T| = n$, choose π such that $\pi(t) = 1/n$ for all $t \in T$, and v such that for all $t \in T$, and all $j < k-1$, $v(a_j, t) < \min\{v(a_{k-1}, t), v(a_k, t)\}$. Suppose that for all $t \in T$, $v(a_{k-1}, t) = 0$. For each $t \in Z$, define $v(a_k, t) = -1$. Notice that $B \cap C \neq \emptyset$, since otherwise it would belong to $\mathcal{I}$. For every $t \in B \cap C$, assume $v(a_k, t) > |Z|$. This is well defined because by construction of Z, $B \cap C \cap Z = \emptyset$. For all $t \in T \setminus (Z \cup (B \cap C))$, set $v(a_k, t) = 0$. These assumptions are consistent with v belonging to (5.4) for any $u \in \mathbb{S}$.

Next, note that it follows from the definitions of F and G that $F \cup (B \cap C), G \cup (B \cap C) \in \mathcal{I}$. So by (5.25), $g^{F \cup (B \cap C)}, g^{G \cup (B \cap C)} \in \mathcal{G}$. Moreover, the properties of v imply that $g^{F \cup (B \cap C)}, g^{G \cup (B \cap C)} \in B(v, \pi)$; to see this, notice that the payoffs are such that the listener's top priority is to assign action a_k to every type in $B \cap C$, but it is not possible to assign action a_k to all these types without assigning a_k to at least one t in Z. $g^{F \cup (B \cap C)}$ and $g^{G \cup (B \cap C)}$ assign a_k to just one $t \in Z$. Notice any that any $g \in \mathcal{G}$ which is greater than both $g^{F\cup(B\cap C)}$ and $g^{G\cup(B\cap C)}$ must assign a_k to at least two types in Z (t_1 and t_2), and hence cannot be optimal. Nor can any $g \in \mathcal{G}$ which is less than both $g^{F\cup(B\cap C)}$ and $g^{G\cup(B\cap C)}$ be optimal, because any such g cannot assign a_k to all types in $(B \cap C)$. It follows that $B(v, \pi)$ has neither a greatest nor a least element. $\square$

5.5.5. Proof of Theorem 5.4

(i): Assume that g_1 and g_2 satisfy the assumptions of part (i) of the theorem, so that in particular, $g_2 \leq g_1$. Optimality of g_1 relative to v_1 implies:

$$\begin{aligned}
(5.40) \qquad 0 &\leq V(g_1; v_1, \pi) - V(g_2; v_1, \pi) \\
&= \sum_{t\in T}[v_1(g_1(t), t) - v_1(g_2(t), t)]\pi(t) \\
(5.41) \qquad &\leq \sum_{t\in T}[v_2(g_1(t), t) - v_2(g_2(t), t)]\pi(t) \\
(5.42) \qquad &= V(g_1; v_2, \pi) - V(g_2; v_2, \pi)
\end{aligned}$$

where (5.41) follows from (5.5). Since g_2 is optimal relative to v_2, this implies that g_1 is optimal relative to v_2. A similar argument implies that g_2 is optimal relative to v_1.

(ii): Suppose that the listener's utility function must be drawn from $\{v_1, v_2\}$, which can also be thought of as a parameter set. Suppose, moreover, that $\{v_1, v_2\}$ is an ordered set, where the ordering $\prec$ satisfies $v_1 \prec v_2$. Then, where $v \in \{v_1, v_2\}$, algebra similar to that in (5.40)-(5.42) shows that $V(g; v, \pi)$ has increasing differences in g and v. On the other hand, Theorem 5.1 implies that $V(g; v, \pi)$ is supermodular in g for all v. Theorem 2.8.1 in Topkis (1998), which says that a parameterized collection of supermodular functions on a lattice which have increasing differences in the parameter has optimal solutions which are increasing (under the strong set order) in the parameter implies (ii).

(iii): If normality fails, then following, in a more compact and slightly different form, the construction in Theorem 5.3, we perform the following steps (see the proof of Theorem 5.3 in Section 5.5.4 for justification): it is possible to find sets $B, C \in \mathcal{I}$ such that $B \cap C \notin \mathcal{I}$, let $\mathcal{D}$ be the $\subseteq$-minimal sets in $\{A \in \mathcal{I} : B \cap C \subseteq A\}$. Then, it is possible to find a set Z such that $Z \cap B \cap C = \emptyset$, but such that Z intersects every element of $\mathcal{D}$. It is possible to choose sets F and G in $\mathcal{D}$[11] and types t_1, t_2 such that $t_1 \neq t_2$ and $F \cap Z = \{t_1\}, G \cap Z = \{t_2\}$. For every $t \in T$, assume $v_1(a_{k-1}, t) = v_2(a_{k-1}, t) = 0$. For every $t \in B \cap C$ (which must be nonempty), assume $v_1(a_k, t) = v_2(a_k, t) > 2|Z|$. For every $t \in Z \setminus \{t_1, t_2\}$, assume $v_1(a_k, t) = v_2(a_k, t) = -2$. For some small ϵ assume:

$$
\begin{aligned}
v_1(a_k, t_1) &= -1 \\
v_1(a_k, t_2) &= -(1-\epsilon) \\
v_2(a_k, t_1) &= -(1-3\epsilon) \\
v_2(a_k, t_2) &= -(1-2\epsilon),
\end{aligned}
$$

and for every $t \notin (B \cap C) \cup Z$, $v_1(a_k, t) = v_2(a_k, t) = 0$.

Next note that for all $u \in \mathbb{S}$, it is possible to satisfy

$$v_1(a_j, t) = v_2(a_j, t) < \min\{v_1(a_{k-1}, t), v_1(a_k, t), v_2(a_{k-1}, t), v_2(a_k, t)\}$$

for all $1 \leq j < k-1$ and $t \in T$ consistently with v_1 and v_2 belonging to the set (5.4) defined in terms of u.

[11] These sets are slightly different but related to the sets F and G in the proof of Theorem 5.3 since they belong to $\mathcal{D}$ rather than $\mathcal{F}$.

Notice, moreover that v_1 and v_2 satisfy the relation given in (5.5). Set $\pi(t) = 1/n$ with $n = |T|$ for all $t \in T$. On the other hand, using reasoning similar to that found at the end of the proof of Theorem 5.3 in Section 5.5.4, one can see that every $g_1 \in B(v_1, \pi)$ will be such that $g_1 = g^H$ for some $H \in \mathcal{I}$ such $B \cap C \subseteq H$ and $H \cap Z = \{t_1\}$. Likewise for every $g_2 \in B(v_2, \pi)$ will be such that $g_2 = g^I$ for some $I \in \mathcal{I}$ such that $B \cap C \subseteq I$ and $I \cap Z = \{t_2\}$. As $t_1 \neq t_2$, the set of optimal satisfy the condition in (iii). $\square$

5.5.6. Proof of Theorem 5.5

Lemma 5.4. *Choose any $B, C \in \mathcal{I}$. If $t \in (B \cap C) \setminus (B \wedge C)$, then t cannot summarize.*

Proof. Assume that t *can* summarize. It follows from (5.6) that:

$$\forall m \in M(t), A_{\{m(t)\}} \subseteq A_{\{m\}}. \tag{5.43}$$

Since $B, C \in \mathcal{I}$, there exist questions Q and R, such that $B = A_Q, C = A_R$. If $t \in A_Q \cap A_R$, then there exists $m_Q \in Q, m_R \in R$ such that $m_Q, m_R \in M(t)$. But then by (5.43):

$$\mathcal{I} \ni A_{\{m(t)\}} \subseteq A_{\{m_Q\}} \cap A_{\{m_R\}} \subseteq A_Q \cap A_R.$$

It follows that $A_{\{m(t)\}} \subseteq A_Q \wedge A_R = B \wedge C$. So if t can summarize then if $t \in B \cap C$, then $t \in B \wedge C$, which implies that $t \notin (B \cap C) \setminus (B \wedge C)$. $\square$

We now strengthen the preceding lemma:

Lemma 5.5. *Choose any $g_1, g_2 \in \mathcal{G}$. Then if $[\min(g_1, g_2)](t) \neq [g_1 \wedge g_2](t)$, then t cannot summarize.*

Proof. Let $(I_2^1, \ldots, I_k^1)$ and $(I_2^2, \ldots, I_k^2)$ be the k-chains corresponding respectively to the outcome functions g_1 and g_2. Then $(I_2^1 \cap I_2^2, \ldots, I_k^1 \cap I_k^2)$ and $(I_2^1 \wedge I_2^2, \ldots, I_k^1 \wedge I_k^2)$ correspond to $\min(g_1, g_2)$ and $g_1 \wedge g_2$ respectively. If $[\min(g_1, g_2)](t) \neq [g_1 \wedge g_2](t)$, then $[g_1 \wedge g_2](t) < a_j = [\min(g_1, g_2)](t)$ for some $j \in \{2, \ldots, k\}$. So $t \in (I_j^1 \cap I_j^2) \setminus (I_j^1 \wedge I_j^2)$. So by Lemma 5.4, t cannot summarize. $\square$

Assume the hypotheses of the theorem. Now choose $g_1, g_2 \in \mathcal{G}$. Let us omit v in π in $U(\cdot; v, \pi)$, and instead write $U(\cdot)$. Then modularity of U on $\mathbf{A}^T$ along with (4.3) in Theorem 4.5) implies:

$$U(g_1) + U(g_2) = U(g_1 \vee g_2) + U(\min(g_1, g_2))$$

It follows that U is supermodular if and only if for all $g_1, g_2 \in \mathcal{G}$:

$$U(\min(g_1, g_2)) \leq U(g_1 \wedge g_2) \tag{5.44}$$

Let $S := \{t \in T : [\min(g_1, g_2)](t) \neq [g_1 \wedge g_2](t)\}$. Then

$$U(\min(g_1, g_2)) - U(g_1 \wedge g_2) = \sum_{t \in S} [v([\min(g_1, g_2)](t), t) - v([g_1 \wedge g_2](t), t)]\pi(t) \leq 0, \tag{5.45}$$

where the inequality follows from the fact that by Lemma 5.5, every t in S cannot summarize, and (5.7). (5.45) implies (5.44), and so completes the proof. $\square$

5.5.7. Proof of Lemma 5.1

Since $g \in \mathcal{G}$, there must exist persuasion rule f'' such that $g = g_{f''}$. Now choose any $t \in T$. Then there must be some $m_t \in T$ such that $f''(m_t) = g(t)$. This means that for all $s \in T$, such that $m_t \in M(s)$, $g(s) \geq g(t)$. So

$$g(t) = \min\{g(s) : m_t \in M(s)\} = f(m_t) \leq g_f(t).$$

On the other hand note that every message $m \in M(t)$ must be such that $f(m) = \min\{g(s) : m \in M(s)\} \leq g(t)$. So t can get at most $g(t)$ under f, or in other words, $g_f(t) \leq g(t)$. So $g(t) = g_f(t)$.

Next consider any f' such that $g_{f'} = g$. Then choose any $m \in \bigcup_{t \in T} M(t)$. Then for all $t \in T$ such that $m \in M(t)$, $f'(m) \leq g(t)$. So $f'(m) \leq \min\{g(t) : m \in M(t)\} = f(m)$. $\square$

5.5.8. Proof of Theorem 5.6

Suppose that interests are more aligned under v_2 than under v_1. Let $f_1 \in \mathcal{F}^*(v_1)$. Then $g_{f_1} \in B(v_1, \pi)$. By part (ii) of Theorem 5.4, there exists $g \in B(v_2, \pi)$ such that $g_{f_1} \leq g$. Define f_2 by:

$$f_2(m) = \min\{g(t) : m \in M(t)\}$$

for all $m \in \bigcup_{t \in T} M(t)$.[12] By Lemma 5.1, $g = g_{f_2}$, which implies that $f_2 \in \mathcal{F}^*(v_2)$. Now choose $m \in \bigcup_{t \in T} M(t)$. Then $f_1(m) \leq \min\{g_{f_1}(t) : m \in M(t)\} \leq \min\{g(t) : m \in M(t)\} = f_2(m)$. This establishes (5.8).

Next choose $f_2' \in \mathcal{F}(v_2)$. This implies that $g_{f_2'} \in B(v_2, \pi)$. It follows from part (ii) of Theorem 5.4 that there exists $g' \in B(v_1, \pi)$ such that $g' \leq g_{f_2'}$. Now, define:

$$f_1'(m) := \min\{g'(t) : m \in M(t)\}.$$

[12] We may assume that f_2 assigns the highest action to every $m \in \mathbf{M} \setminus \bigcup_{t \in T} M(t)$.

for any $m \in \bigcup_{t\in T} M(t)$.[13] Then by Lemma 5.1, $g' = g_{f_1'}$. Now choose any $m \in \bigcup_{t\in T} M(t)$ such that $f_2'(m) \in \{g_{f_2'}(t) : m \in M(t)\}$. Then choose $t \in T$ such that $f_2'(m) = g_{f_2'}(t)$. Then:

$$f_2'(m) = g_{f_2'}(t) = g'(t) \geq g_{f_1'}(t) \geq f_1'(m),$$

where the first inequality follows from the fact that $g' \geq g_{f_1'}$, and the second inequality follows from $m \in M(t)$. $\square$

5.5.9. Proof of Theorem 5.7

In the course of the proof of part (iii) of Theorem 5.4 in Section 5.5.5, it was shown that whenever the message structure is not normal, it is possible to construct utility functions v_1 and v_2 such that for some type t^*:

$$g_1 \in B(v_1, \pi) \;\;\Rightarrow\;\; (\exists H \in \mathcal{I}, t^* \in H \text{ and } g_1 = g^H)$$
$$g_2 \in B(v_2, \pi) \;\;\Rightarrow\;\; (\exists I \in \mathcal{I}, t^* \notin I, \text{ and } g_2 = g^I)$$

It follows that there exists $f_1 \in \mathcal{F}(v_1)$ and $m^* \in M(t^*)$ such that $f_1(m^*) = a_k$. On the other hand, for all $f_2 \in \mathcal{F}_2(v_2)$, $f_2(m^*) < a_k$. So (5.10) with the roles of v_1 and v_2 reversed holds for m^* with a strict inequality. Moreover, for any $g_1 \in B(v_1, \pi)$, $g_1(t^*) = a_k$, so for any $f_1 \in \mathcal{F}(v_1)$ if $f_1(m^*) = g_{f_1}(t)$, then $f_1(m^*) = a_k$, the highest possible action. So (5.11) holds with the roles of v_1 and v_2 reversed. $\square$

5.5.10. Proof of Theorem 5.2

First I prove that if $(\varphi_1, \psi_1), (\varphi_2, \psi_2) \in \Phi$, then $(\varphi_1, \psi_1) \bullet (\varphi_2, \psi_2) \in \Phi$. Notice first that

[13] We may assume that f_1' assigns the lowest action to any $m \in \mathbf{M} \setminus \bigcup_{t\in T} M(t)$.

$\varphi_1 \circ \varphi_2$ is a bijection from T to T and that $\psi_1 \circ \psi_2$ is a bijection from $\mathbf{M}$ to $\mathbf{M}$. Next observe, using (5.13) for (φ_1, ψ_1) and (φ_2, ψ_2), we have that:

$$\begin{aligned} M(\varphi_1 \circ \varphi_2(t)) &= \{\psi_1(m) : m \in M(\varphi_2(t))\} \\ &= \{\psi_1(m) : m \in \{\psi_2(m) : m \in M(t)\}\} \\ &= \{\psi_1 \circ \psi_2(m) : m \in M(t)\} \end{aligned}$$

, implying (5.13) for $(\varphi_1, \psi_1) \bullet (\varphi_2, \psi_2)$.

Likewise, using (5.14) and (5.15) for (φ_1, ψ_1) and (φ_2, ψ_2):

$$v(a_i, t) = v(a_i, \varphi_2(t)) = v(a_i, \varphi_1 \circ \varphi_2(t))$$

$$\pi(t) = \pi(\varphi_2(t)) = \pi(\varphi_1 \circ \varphi_2(t)),$$

implying (5.14) and (5.15) for $(\varphi_1, \psi_1) \bullet (\varphi_2, \psi_2)$. This establishes that Φ is closed under $\bullet$.

Next notice that for any $(\varphi, \psi) \in \Phi$,

$$(\varphi_{\mathbf{I}}, \psi_{\mathbf{I}}) \bullet (\varphi, \psi) = (\varphi, \psi) \bullet (\varphi_{\mathbf{I}}, \psi_{\mathbf{I}}) = (\varphi, \psi).$$

So $(\varphi_{\mathbf{I}}, \psi_{\mathbf{I}})$ is the identity element of Φ.

Next I prove that if $(\varphi, \psi) \in \Phi$, then $(\varphi, \psi)^{-1} = (\varphi^{-1}, \psi^{-1}) \in \Phi$. So assume $(\varphi, \psi) \in \Phi$. First notice that φ^{-1} and ψ^{-1} are bijections. Choose $t \in T$. Then there must be some $s \in T$ such that $\varphi(s) = t$. This means that $\varphi^{-1}(t) = s$. We know that $M(t) = M(\varphi(s)) = \{\psi(m) : m \in M(s)\}$. Since ψ is a bijection from $\mathbf{M}$ to $\mathbf{M}$, this implies that ψ is a bijection from $M(s)$ to $M(t)$. So ψ^{-1} is a bijection from $M(t)$ to $M(s)$. So $M(\varphi^{-1}(t)) = M(s) = \{\psi^{-1}(m) : m \in M(t)\}$. So $(\varphi^{-1}, \psi^{-1})$ satisfies (5.13) . Next, note that $v(a_i, \varphi^{-1}(t)) = v(a_i, \varphi(\varphi^{-1}(t))) = v(a_i, t)$, where the first equality follows from the fact that (φ, ψ) satisfy (5.14), implying that $(\varphi^{-1}, \psi^{-1})$ satisfy (5.14). Likewise $\pi(\varphi^{-1}(t)) = \pi(\varphi(\varphi^{-1}(t))) = \pi(t)$, implying that $(\varphi^{-1}, \psi^{-1})$ satisfies (5.15), and hence that $(\varphi, \psi)^{-1} \in \Phi$.

Observing that:

$$(\varphi, \psi)^{-1} \bullet (\varphi, \psi) = (\varphi, \psi) \bullet (\varphi, \psi)^{-1} = (\varphi_{\mathbf{I}}, \psi_{\mathbf{I}}),$$

it follows that $(\varphi, \psi)^{-1}$ is the inverse of (φ, ψ) in Φ.

Finally, notice that $\bullet$ is obviously associative. This completes the proof that Φ is a group under $\bullet$. $\square$

5.5.11. Proof of Theorem 5.8

(1): Note that for each $t \in T$, there exists $m_t \in M(t)$ such that $f \circ \psi(m_t) = g_{f\circ\psi}(t)$, and for all $m \in M(t)$, $f \circ \psi(m) \leq g_{f\circ\psi}(t)$. But invoking (5.13) and the fact that (φ, ψ) is a symmetry, this implies that given persuasion rule f, $\psi(m_t)$ attains a higher action for type $\varphi(t)$ than any other message in $M(\varphi(t))$. So $g_f(\varphi(t)) = f(\psi(m_t)) = g_{f\circ\psi}(t)$.

(2):

$$V(g_f; v, \pi) = \sum_{t \in T} v(g_f(\varphi(t)), \varphi(t))\pi(\varphi(t)) \tag{5.46}$$

$$= \sum_{t \in T} v(g_{f\circ\psi}(t), \varphi(t))\pi(\varphi(t)) \tag{5.47}$$

$$= \sum_{t \in T} v(g_{f \circ \psi}(t), t)\pi(t) \tag{5.48}$$

$$= V(g_{f \circ \psi}; v, \pi), \tag{5.49}$$

where (5.46) follows from the fact that φ is a bijection from T to T, (5.47) follows from part (5.13) of the theorem, and (5.48) follows from the fact that (φ, ψ) is a symmetry.

(3) is an immediate consequence of (2). $\square$

CHAPTER 6

Relation to the L-principle

In this chapter, I will compare the approach taken in the previous chapters to the approach taken by Glazer and Rubinstein (2006). The previous chapter illustrates the advantages of the current approach. We were able to find properties of the listener's optimization problem which hold exactly when the speaker can summarize information. We derived comparative statics under the assumption that speaker can summarize information, and also under the assumption that if the speaker cannot summarize his information, then the listener views his request for a high action negatively. We were also able to show the significance of normality for symmetry of the optimal rule. Previously, we studied extensively the properties of the listener's choice set as well as the structure of messages. However, ultimately, the approach taken here is complementary with the approach taken by Glazer and Rubinstein, as this section will demonstrate.

First I will present a formulation of the listener's problem presented by Glazer and Rubinstein. Then in Section 6.1, I will generalize this formulation to k actions and relate it to the approach to taken here. Finally, in Section 6.2, I will show how under normality and with two actions, the program studied by Glazer and Rubinstein specializes to a well known linear program known as the "maximal closure problem."

Glazer and Rubinstein (2006) find a integer programming formulation of the listener's problem when there are two actions. They consider the case where T can be partitioned into two disjoint sets $\mathbb{A}$ and $\mathbb{R}$ where $\mathbb{A}$ contains the "good types" and $\mathbb{R}$ contains the "bad types". In this case, the utility to rejection a_1 can be normalized to be zero regardless of the type. The utility of acceptance a_2 is 1 for types in $\mathbb{A}$ and -1 for types in $\mathbb{R}$. It is not difficult to see–as the authors point out–that when $k = 2$, nothing essential is lost when restricting to this special case.

In order to further attack this problem, the authors employ a constraint called the "L-principle" which relates to an idea originally introduced in the context of a different but related model in Glazer and Rubinstein (2004). A pair (t, S) where $t \in \mathbb{A}$ and $S \subseteq \mathbb{R}$ is said to be an L if for any $m \in M(t)$, there is an $s \in S$ such that $m \in M(s)$. If $t \in \mathbb{A}$, this means that when the speaker's type is t, the listener would like to accept the speaker's request. However when the type belongs to $S \subseteq \mathbb{R}$, the listener would like to reject the speaker's request. If (t, S) is an L, this means that any message which can be sent by a good type t, can also be sent by some bad type in S. An L, (t, S) is *minimal* if there does not exist a S' which is a proper subset of S such that (t, S') is also an L. Glazer and Rubinstein (2006) show that the solution to the listener's problem has the same value as the program:

$$
\begin{aligned}
& \min_{\{\mu_t\}_{t \in T}} \sum_{t \in T} \pi(t) \mu_t \\
\text{s.t.} \quad & \mu_t \in \{0, 1\} \text{ for all } t \in T \\
& \sum_{s \in \{t\} \cup S} \mu_s \geq 1 \text{ for every minimal } L, (t, S).
\end{aligned}
\tag{6.1}
$$

The basic idea is that μ_t is the probability of error conditional on type t, that is, the probability of accepting t if $t \in \mathbb{R}$ or rejecting t if $t \in \mathbb{A}$. The listener attempts to minimize error subject to the constraint that the sum of errors over every minimal L is at least 1.

6.1. Relation of the L-principle to Essential Messages and Generalization to k actions

In this section, I will use the representation I have been developed above to write down an integer programming formulation of the listener's problem for k actions. I will then specialize to $k = 2$ and show that the program reduces to (6.1), so that the constraints of this program generalize the L-principle. Notice that it is not clear how to extend the L-principle to k actions, since this principle relied on the distinction between good and bad types which no longer exists with k actions. Moreover, the objective in (6.1) is to minimize the listener's probability of error, and the L-constraints impose conditions on the error probabilities. The notion of error which is employed–accepting a bad type or rejecting a good type–are particular to the case $k = 2$.

The program and approach presented here differs from the program presented by Glazer and Rubinstein in several other ways. In particular, looking at the L-constraints, it is clear that these constraints rely on information not just about what is incentive-feasible, but also on information about the listener's utility function, since an L, (t, S) must be such that the utility of accepting type t is positive and the utility of accepting types in S is negative.

I will break the argument into two parts. First, I will use the results of the previous sections to write down a set of constraints which correspond exactly to the set of implementable outcome functions as a set of linear constraints. Then I will use properties of the listener's utility function to eliminate constraints which are not binding. For the latter exercise, I will use the assumption that that the listener's utility function is concave in the action for all types. A similar argument would apply if I assumed that the listener's utility was a concave transformation of some monotone function of the actions. Interestingly, this was essentially the same assumption which was necessary to generalize Glazer and Rubinstein's result about optimal deterministic credible persuasion rules to k actions. Splitting the exercise into two steps has the value of separating what one can deduce from considering the speaker's incentives from what one can deduce from considering properties of the listener's utility function. This could be useful if one wanted to perform a related exercise in neighboring models. For example, suppose that one assumed that the listener was limited in some way with respect to the persuasion rules which he could choose.

Another difference between the approach taken here and the approach in Glazer and Rubinstein is that I will not appeal to messages directly but rather will appeal to the representation of the family of implementable outcome functions which was developed in Chapter 3. Given the results in that chapter, one could easily translate this back into the language of messages. The advantage of the current approach is that–despite not explicitly mentioning messages–it relates the constraints in a more detailed way to the structure of the speaker's messages. In particular, the approach shows how in writing down the constraints, it is sufficient to restrict attention to essential messages, and is more detailed with respect to the question of which messages are relevant for which constraints.

So consider a vector $\nu = \{\nu_t^j\}_{t \in T}^{j=1,\dots k} \in \{0,1\}^{T \times k}$. ν is to be interpreted as a binary vector which encodes which action is assigned to which type. For example $\nu_t^j = 1$ means that type t gets action a_j and $\nu_t^j = 0$ means that type t does not get action a_j. Vectors ν such that for some t, and $i \neq j$, $\nu_t^i = \nu_t^j = 1$ are impossible to interpret, but the constraints presented below will rule them out. Now start with a message structure and form the family $\mathcal{I}$ of acceptance sets corresponding to this structure, as described in Chapter 3. Define:

$$\mathbb{I} = \{\nu \in \{0,1\}^{T\times k} : \exists (I_2, \ldots, I_k) \in \mathbf{C}(\mathcal{I}, k), \forall j = 1, \ldots, k, \forall t, \nu_t^j = 1 \Leftrightarrow t \in I_j \setminus I_{j+1}\},$$

where as usual $I_1 = T, I_{k+1} = \emptyset$. Using Theorem 3.1, one can see that $\nu \in \mathbb{I}$ if and only if ν is a binary vector which represents some implementable outcome function.

Now define $\mathcal{N}(t)$ to be the set of $\subseteq$-minimal elements of $\mathcal{I}$ containing t. Thus, every element of F of $\mathcal{N}(t)$ is such that $F = A_{\{m\}}$ for some message m which is maximally informative for type t. Next define $\mathcal{N}^*(t) = \{F \setminus \{t\} : F \in \mathcal{N}(t)\}$. Thus, $\mathcal{N}^*(t)$ is the set that results from removing t from every element of $\mathcal{N}(t)$. There is one special case to keep in mind here, although it does not introduce any complications. Notice that $\mathcal{N}(t) = \{\{t\}\} \Leftrightarrow \{t\} \in \mathcal{I}$. In other words, t has a message which distinguishes him from all other types if and only if his unique (up to equivalence) maximally informative message is the one that distinguishes him from all other types. If $\mathcal{N}(t) = \{\{t\}\}$, then $\mathcal{N}^*(t) = \{\emptyset\}$. I point out this case because it is qualitatively a bit different from other cases, but it does not present any problems for the approach taken here.

A *blocking set* for a family $\mathcal{S}$ of sets is a set B which intersects every member of $\mathcal{S}$. A blocking set B is *minimal* if no proper subset is a blocking set. Let $\mathcal{B}(t)$ be the collection of minimal blocking sets for $\mathcal{N}^*(t)$. The following Lemma expresses implementablity in terms of a set of linear constraints:

Lemma 6.1. *Assume* $\nu \in \{0,1\}^{T\times k}$*.* $\nu \in \mathbb{I}$ *is equivalent to the joint satisfaction of the following two conditions:*

$$\forall t \in T, \sum_{j=1}^{k} \nu_t^j = 1 \tag{6.2}$$

$$\forall t \in T, \forall j = 2, \ldots, k, \forall B \in \mathcal{B}(t), \sum_{s\in B}\sum_{i=j}^{k} \nu_s^i \geq \nu_t^j \tag{6.3}$$

Proof. See Section 6.3.1. □

The final step is to use information from the listener's utility function to eliminate many non-binding constraints. Then we will arrive at a formulation which generalizes (6.1).

I make the simplifying assumption that for each $t \in T$, there is a unique action a_i which maximizes $v(a_i, t)$, and define $\jmath(t)$ to be the index of this action. Below, I will discuss the minor changes necessary without this assumption. Define:

$$\mathcal{B}_j(t) := \{B \in \mathcal{B}(t) : \forall s \in B, \jmath(s) < j\}$$

Thus $\mathcal{B}_j(t)$ contains only blocking sets B such that for every type $t \in B$, a_j is suboptimal from the listener's perspective, and moreover it is optimal for the listener to assign every type in B an action lower than a_j. These types are the analog of bad types in the case $k = 2$, but it is now necessary to define bad types locally relative to an action a_j. Finally, let $\eta(t) = \min(\{j : v(a_j, t) < v(a_1, t)\} \cup \{k+1\})$. Thus $\eta(t)$ is the action with lowest index among those which are worse than the lowest action. Note that constraint (6.7) below is vacuous if $\eta(t) - 1 = 1$.

Theorem 6.1. *Assume that $v(a_j, t)$ is concave in a_j and for all $t \in T$, $\pi(t) \neq 0$. Then the listener's problem is equivalent to the program:*

$$\max_{\nu} \sum_{t \in T} \sum_{j=1}^{k} v(a_j, t)\pi(t)\nu_t^j \tag{6.4}$$

$$\text{s.t.} \quad \forall t \in T, \forall j = 1, \ldots, k, \nu_t^j \in \{0, 1\} \tag{6.5}$$

$$\forall t \in T, \sum_{j=1}^{k} \nu_t^j = 1 \tag{6.6}$$

$$\forall t \in T, \forall j = 2, \ldots, \eta(t) - 1, \forall B \in \mathcal{B}_j(t), \sum_{s \in B} \sum_{i=j}^{k} \nu_s^i \geq \nu_t^j \tag{6.7}$$

in the sense that the value of this program is the value of the listener's problem, and $\widehat{g}$ is an optimal implementable outcome function if and only if there exists a solution $\widehat{\nu}$ to this program such that for all t and j, $\widehat{g}(t) = a_j \Leftrightarrow \widehat{\nu}_t^j = 1$.

Proof. See Section 6.3.2. □

Notice that the last constraint differs from (6.3) in that $\mathcal{B}(t)$ has been replaced by $\mathcal{B}_j(t)$, and also in the sense that one only considers constraints corresponding to $j \leq \eta(t)$. This may involve the elimination of many constraints relative to (6.3). If I had not assumed that $v(a_i, t)$ had a unique maximizer, then $\jmath(t)$ could have been defined as either the smallest or largest index of an maximizer of $v(a_i, t)$. More constraints would have been eliminated in (6.7) if the largest were chosen, but given this choice, while the the problem would still have had the value of the listener's problem, the set of optimal implementable outcome functions could only have been ensured to be a subset of the set of solutions to the program.

Consider the case considered by Glazer and Rubinstein (2006) in which $k = 2$, and in which for all $t, v(a_1, t) = 0$, for $t \in \mathbb{A}, v(a_2, t) = 1$, and for $t \in \mathbb{R}, v(a_2, t) = -1$. Then the error probability μ_t can be defined as $1 - \nu_t^2$ if $t \in \mathbb{A}$ and ν_t^2 if $t \in \mathbb{R}$. If $t \in \mathbb{R}$, then $\eta(t) = 2$, and hence there is no constraint corresponding to (6.7). On the other hand it $t \in \mathbb{A}$, then $\eta(t) = 3$, so there generally is a set of such constraints. Note finally that for $t \in \mathbb{A}$, (t, S) is a minimal L if and only if $S \in \mathcal{B}_2(t)$, and then:

$$\sum_{s \in S \cup \{t\}} \mu_s \geq 1 \Leftrightarrow \sum_{s \in S} \nu_s^2 \geq \nu_t^2,$$

and there is only one summation rather than two on the left hand side because $k = 2$. Thus the program (6.1) is derived as a special case.

6.2. Normality and the Maximal Closure Problem

In this section, I consider what happens to the program introduced in the previous section when normality holds and $k = 2$. This is a special case of (6.1). I make the further simplifying assumption that if $t_1 \neq t_2$, then $M(t_1) \neq M(t_2)$. In other words, no pair of types has exactly the same set of messages. Then the relation $\preccurlyeq^{\mathcal{I}}$ is not only a quasi-order, but also a partial order (reflexive, transitive, and antisymmetric).

Under normality, $\mathcal{N}(t)$ contains only one element; in particular, this is the element $N(t) := \bigcap\{I \in \mathcal{I} : t \in I\}$. In Theorem 4.2, we saw that under normality, the set of implementable outcome functions is the set of outcome functions which are monotone with respect to $\preccurlyeq^{\mathcal{I}}$. It is easy to confirm that for types t_1, t_2, $t_1 \preccurlyeq^{\mathcal{I}} t_2 \Leftrightarrow t_2 \in N(t_1)$. Moreover, $\mathcal{B}(t) = \{\{s\} : s \in N(t) \setminus \{t\}\}$. In the case $k = 2$, we can normalize $v(a_1, t) = 0$ for all $t \in T$, and then write $w(t) := v(a_2, t)\pi(t)$, and define $\nu_t := \nu_t^2$. Then, given the assumption $\nu_t \in \{0, 1\}$, constraint (6.2) becomes redundant, and maximization of (6.4) subject to (6.2) and (6.3) is equivalent to:

$$\begin{aligned} &\max_{\nu} \sum_{t \in T} w(t)\nu_t \\ \text{s.t.} \quad &\forall t \in T, \nu_t \in \{0,1\} \\ &\forall t, s \in T, t \preccurlyeq^{\mathcal{I}} s \Rightarrow \nu_t \leq^* \nu_s. \end{aligned}$$

It follows from Lemma 6.1 that this program corresponds to the listener's problem in this case. This is a well-known linear program studied by Picard (1976) and others known as the maximal closure problem. For a recent discussion of this problem, see Hochbaum (2001). The problem has the interpretation of choosing a set of nodes on a directed graph so as to maximize some weight function (taking both positive and negative values), and subject to the constraint that if a node x_1 is chosen and a directed edge points from x_1 to x_2, then x_2 must be chosen as well. Here the nodes correspond to the types, and an edge points from type t_1 to type t_2 if t_2 is in immediate successor of t_1 relative to $\preccurlyeq^{\mathcal{I}}$.

The maximal closure problem has been extensively studied in settings very different from the one studied here. The problem can be viewed as a special case of the max-flow min-cut problem, and the comparative statics result in (ii) of Theorem 5.4 when specialized to the case $k = 2$ coincides with the well-known comparative statics result for the minimum cut problem. (See Theorems 3.7.2 and 3.7.4 in Topkis (1998)).

6.3. Proofs

6.3.1. Proof of Lemma 6.1

In order to prove Lemma 6.1, I first have to prove Lemma 6.2.

Lemma 6.2. *$\nu \in \mathbb{I}$ is equivalent to the joint satisfaction of the following two conditions:*

$$\forall t \in T, \sum_{j=1}^{k} \nu_t^j = 1 \tag{6.8}$$

$$\forall t \in T, \nu_t^j = 1 \Rightarrow \exists I^t \in \mathcal{N}(t), \forall s \in I^t \setminus \{t\}, \exists i_s \geq j, \nu_s^{i_s} = 1 \tag{6.9}$$

Proof. (6.8) follows from $\nu \in \mathbb{I}$ because there is exactly one j such that $t \in I_j \setminus I_{j+1}$. Next choose $\nu \in \mathbb{I}$, and let $(I_2, \ldots, I_k)$ be the corresponding element of $\mathbf{C}(\mathcal{I}, k)$. Choose $t \in T$ and suppose $\nu_t^j = 1$. Then $t \in I_j$, and there must be some $I^t \subseteq \mathcal{N}(t)$ such that $t \in I^t \subseteq I_j$. It follows that for each $s \in I^t \setminus \{t\}$, $s \in I_j$, so the unique i_s such that $s \in I_{i_s} \setminus I_{i_s+1}$ must be at least as large as j, establishing (6.9).

Next assume (6.8) and (6.9). Notice that (6.8) implies that for each $t \in T$, there is a unique j such that $\nu_t^j = 1$ and for all $i \neq j$, $\nu_t^i = 0$. To establish $\nu \in \mathbb{I}$, we must find $(I_2, \ldots, I_k) \in \mathbf{C}(\mathcal{I}, k)$ corresponding to ν. Define $I_j = \{t \in T : \forall i, \nu_t^i = 1 \Rightarrow i \geq j\}$. Then clearly the sequence $(I_2, \ldots, I_k)$ is decreasing according to $\subseteq$. It remains only to show that each I_j belongs to $\mathcal{I}$. Since $\mathcal{I}$ is closed under union, and $I^t \in \mathcal{N}(t) \subseteq \mathcal{I}$, it is sufficient to show that:

$$I_j = \bigcup_{t \in I_j} I^t =: J. \tag{6.10}$$

Since for all t, $t \in I^t$, $I_j \subseteq J$. To prove that $J \subseteq I_j$, it is sufficient to show that for all $t \in I_j, I^t \subseteq I_j$. So choose $t \in I_j$. Then there exists $i \geq j$ such that $\nu_t^i = 1$. So by (6.9), $I^t \subseteq I_i \subseteq I_j$, where the last inclusion follows from $i \geq j$ and the fact that $(I_2, \ldots, I_k)$ is a decreasing sequence, establishing (6.10) and hence completing the proof. $\square$

I now use Lemma 6.2 to prove Lemma 6.1. It follows from Lemma 6.2 that it is sufficient to prove that (6.9) and (6.3) are equivalent given (6.2). So first assume (6.9). Choose t and j. If $\nu_t^j = 0$, then (6.3) is automatically satisfied. So suppose $\nu_t^j = 1$. Then by (6.9), there exists $I^t \in \mathcal{N}(t)$ such that for all $s \in I^t \setminus \{t\}$, there exists $i_s \geq j$ such that $\nu_t^{i_s} = 1$. Then by the definition of $\mathcal{B}(t)$, every $B \in \mathcal{B}(t)$ must contain at least one $s \in I^t \setminus \{t\}$. It follows that $\sum_{s\in B}\sum_{i=j}^{k} \nu_s^i \geq 1$, and hence (6.3) is satisfied.

Next assume (6.3). Now assume for contradiction that (6.9) is violated from some t. It follows that for some j, $\nu_t^j = 1$, and for every $F \in \mathcal{N}^*(t)$, there is some $s_F \in F$ such that for all $i \geq j$, $\nu_i^{s_F} = 0$. Notice that $\{s_F : F \in \mathcal{N}(t)\}$ is a blocking set for $\mathcal{N}^*(t)$, so there is some $B \subseteq \{s_F : F \in \mathcal{N}^*(t)\}$ with $B \in \mathcal{B}(t)$, and $\sum_{s\in B}\sum_{j=1}^{k} \nu_s^i = 0$, contradicting (6.3). $\square$

6.3.2. Proof of Theorem 6.1

Proof of Theorem 6.1. It follows from Theorem 3.1 that the listener's problem is equivalent to maximizing the objective in (6.4) subject to the constraint that $\nu \in \mathbb{I}$ and (6.5). It then follows from Lemma 6.1 that the listener's problem is equivalent to maximizing the objective in (6.4) subject to (6.5), (6.6) and (6.3). Thus, I would like to show that there is a solution to the program (6.4)-(6.7) that will satisfy (6.3).

Since (6.3) can differ from (6.7) in two ways: the index j is only required to vary from 2 to $\eta(t)$, rather than from 2 to k, and $\mathcal{B}(t)$ is replaced by $\mathcal{B}_j(t)$, for clarity, I break the argument into two steps. In the first step, I deal with the $\mathcal{B}_j(t)$, and in the second with $\eta(t)$. So in place of (6.3), I first consider:

$$\forall t \in T, \forall j = 2, \ldots, k, \forall B \in \mathcal{B}_j(t), \sum_{s \in B} \sum_{i=j}^{k} \nu_s^i \geq \nu_t^j, \tag{6.11}$$

So I would like to show that if we take any ν which is feasible with respect to the constraints (6.5), (6.6), and (6.11) but which violates some constraint in (6.3), then one can find $\widehat{\nu}$ which satisfied constraints (6.5), (6.6), and (6.11), and more of the constraints in (6.3), and which attains a *strictly* higher utility for the listener. Henceforth, I will omit mention of (6.5) and (6.6) because it will be obvious that these constraints will always remain satisfied. In what follows the notation (6.11-(t, B, j)) as the constraint of the form (6.7) for which ν_t^j is on the left hand side of the inequality and summation is over the blocking set B, and similar notation will be used for constraints in (6.3) and (6.7).

So consider ν satisfying all constraints in (6.11) but violating some constraint (6.3-(t_0, B_0, j_0)). First I argue that:

$$\exists F^* \in \mathcal{N}^*(t_0), \forall t \in T : (t \in F^* \text{ and } \jmath(t) < j_0) \Rightarrow \sum_{i=j_0}^{k} \nu_t^i = 1. \tag{6.12}$$

Assume for contradiction that (6.12) is false. Then for all $F \in \mathcal{N}^*(t), \exists s_F \in F$ such that $\jmath(s_F) < j_0$ and $\sum_{i=j_0}^{k} \nu_{s_F}^i = 0$. But then there must exist $B \subseteq \{s_F : F \in \mathcal{N}^*(t)\}$ such that $B \in \mathcal{B}_j(t_0)$, and it follows that constraint (6.11-(t_0, B, j_0)) is violated, a contradiction.

So choose F^* satisfying (6.12), and define $\widehat{\nu}$ by:

$$\widehat{\nu}_t^i = \begin{cases} 1, & \text{if } i = j_0, t \in F^*, \jmath(t) \geq j_0, \sum_{i=j_0}^{k} \nu_t^i = 0; \\ 0, & \text{if } i \neq j_0, t \in F^*, \jmath(t) \geq j_0, \sum_{i=j_0}^{k} \nu_t^i = 0; \\ \nu_t^i, & \text{otherwise.} \end{cases}$$

I argue that $\widehat{\nu}$ still satisfies all constraints of the form (6.3-(t, B, j)) which were satisfied by ν (including *all* constraints of the form (6.11)-(t, B, j)). Since ν only differs from $\widehat{\nu}$ by assigning some types higher actions, if $\nu_t^j \geq \widehat{\nu}_t^j$, then the fact that ν satisfied (6.3-(t, B, j)) implies that $\widehat{\nu}$ satisfies this constraint as well. Therefore, the only case left to consider is when $\widehat{\nu}_t^j = 1$, $\nu_t^j = 0$. But by the definition of $\widehat{\nu}$, this only happens when $t \in F^*$ and $j = j_0$. This implies that there is some $H \in \mathcal{N}^*(t)$ such that $H \subseteq F^* \cup \{t_0\}$. Also, by the definition of $\widehat{\nu}$, (6.12), and the fact that $\nu_{t_0}^{j_0} = 1$, every $s \in F^* \cup \{t_0\}$ is such that $\sum_{i=j_0}^{k} \widehat{\nu}_s^i = 1$. Since every $B \in \mathcal{B}_j(t)$ must contain some $s \in H \subseteq F^* \cup \{t_0\}$, this implies that (6.3-(t, B, j)) is satisfied. A similar–but slightly simpler–argument shows that constraint (6.3-(t_0, B_0, j_0))–which was not satisfied by ν, is satisfied by $\widehat{\nu}$. Finally, note that the concavity of $v(a_i, t)$ in a_i and the fact that $\pi(t) \neq 0$ for all $t \in T$, and the way that $\widehat{\nu}$ differs from ν, implies that the listener is strictly better off under $\widehat{\nu}$ than under ν. To summarize, we have shown how starting from ν we can find $\widehat{\nu}$ which satisfies strictly more constraints and makes the speaker strictly better off.

The preceding argument establishes that any solution to (6.4)-(6.6) and (6.11) is a solution to (6.4)-(6.6) and (6.3). Next, I argue that any solution to (6.4)-(6.7) is a solution to (6.4)-(6.6) and (6.11). Again, I neglect mention of (6.5) and (6.6), as they will be satisfied throughout. So choose some ν satisfying (6.7) but not (6.11). Let j_0 be the greatest index $j \geq 2$ such that ν violates some constraint of the form (6.11-(t, B, j)), and let S be the set of types s such that (6.11-(s, B, j_0)) is violated. Notice that since constraints of the form (6.7) are all satisfied, then for all $s \in S$, $\eta(s) < j_0$, and that for all $s \in S$, $\nu_s^{j_0} = 1$. Let $|S| =: p$. Choose some $t_0 \in S$, and define $\ell = \max\{\max K(t_0, F) :$

$F \in \mathcal{N}^*(t_0)\}$, where

$$K(t_0, F) = \{j : \forall t \in F, \jmath(t) < j \Rightarrow \sum_{i=j}^{k} \nu_t^i = 1\}.$$

Note that ℓ is well defined because for all $t, \sum_{i=1}^{k} \nu_t^i = 1$. Assume for contradiction that $\ell \geq j_0$. Then choose $F \in \mathcal{N}^*(t_0)$ so that $\max K(t_0, F) = \ell$. Note that any $B \in \mathcal{B}_j(t_0)$ must contain some element from $\{t \in F : \jmath(t) < j_0\} \subseteq \{t \in F : \jmath(t) < \ell\}$, which then implies by the definition of ℓ and the way F was chosen that all constraints of the form (6.11-(t_0, B, j_0)) are satisfied, a contradiction. So $\ell < j$.

Let $F^* \in \mathcal{N}(t)$ be such that $\max K(t_0, F^*) = \ell$. Define

$$\widehat{\nu}_t^i := \begin{cases} 1, & \text{if } t = t_0 \text{ and } i = \ell; \\ 0, & \text{if } t = t_0 \text{ and } i \neq \ell. \\ \nu_t^i, & \text{otherwise.} \end{cases}$$

Then note that since for every $t \in F^*, \jmath(t) < \ell \Rightarrow \sum_{i=\ell}^{k} \nu_t^i = 1$, then every constraint of the form (6.7-(t_0, B, j)) is satisfied. Now assume for contradiction that there is some t_1 and some constraint of the form (6.7-(t_1, B_1, j_1)) which is violated by $\widehat{\nu}$. Since this constraint was satisfied by ν, it must be violated now because t_0 is assigned a lower action and so $\nu_{t_1}^h = 1$ for some $h > \ell$. Moreover, for the same reason, there must be some $H \in \mathcal{N}^*(t_1)$ such that $t_0 \in H$ and for all $s \in H \setminus \{t_0\}, \jmath(t) < h \Rightarrow \sum_{i=1}^{k} \nu_t^i = \sum_{i=1}^{k} \widehat{\nu}_t^i = 1$. But then there exists some $D \in \mathcal{N}^*(t_0)$ such that $D \subseteq H \cup \{t_1\}$, which implies that for all $t \in D, \jmath(t) < h \Rightarrow \sum_{i=h}^{k} \nu_t^i = 1$. So $K(D, t_0) \geq h$, contradiction. Notice that $\widehat{\nu}$ is such that the number of types s for whom some constraint of the form (6.7-(s, B, j_0)) is violated is now $p - 1$. We iterate the procedure until all constraints of the form (6.7-(s, B, j))

are satisfied. Notice that in the process the listener's utility only changes in types which were initially assigned actions higher than $\eta(t)$ are now assigned lower actions, so that by concavity of $v(a_i, t)$ in a_i and the fact that $\pi(t) \neq 0$ for all $t \in T$, the listener is strictly better off. $\square$

References

Bull, J., and J. Watson (2004): "Evidence Disclosure and Verifiability," *Journal of Economic Theory*, 118.

——— (2007): "Hard Evidence and Mechanism Design," *Games and Economic Behavior*, 58, 75–93.

Deneckere, R., and S. Severinov (2001): "Mechanism Design and Communication Costs," Unpublished Manuscript, University of Wisconsin.

Engelking, R. (1989): *General Topology*. Helderman Verlag.

Forges, F., and F. Koessler (2005): "Communication Equilibria with Partially Verifiable Types," *Journal of Mathematical Economics*, 41, 793–811.

Glazer, J., and A. Rubinstein (2003): "Debates and Decisions: On a Rationale of Argumentation Rules," *Games and Economic Behavior*, 36, 158–173.

——— (2004): "On the Optimal Rules of Persuasion," *Econometrica*, 72, 1715–1736.

——— (2006): "A Study in the Pragmatics of Persuasion: A Game Theoretical Approach," *Theoretical Economics*, 1, 395–410.

Green, J., and J. Laffont (1986): "Partially Verifiable Information and Mechanism Design," *Review of Economic Studies*, 53, 447–456.

Greene, C., and D. Kleitman (1978): "Proof Techniques in the Theory of Finite Sets," in *Studies in Combinatorics*, ed. by G. Rota, vol. 17, pp. 22–79. The Mathematical Association of America.

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

Hochbaum, D. (2001): "A New-Old Algorithm for Minimum-cut and Maximum-flow in Closure Graphs," *Networks*, 37, 171–193.

KLEITMAN, D. (1976): "Extremal Properties of Collections of Subsets Containing No Two Sets and their Union," *Journal of Combinatorial Theory (A)*, 20, 390–392.

LIPMAN, B., AND D. 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 Information of Interested Parties," *Rand Journal of Economics*, 17, 18–32.

PICARD, J.-C. (1976): "Maximal Closure of a Graph and Applications to Combinatorial Problems," *Management Science*, 22, 1268–1272.

SHIN, H. (1994a): "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 64, 253–264.

——— (1994b): "News Management and the Value of Firms," *Rand Journal of Economics*, 25, 58–71.

——— (2003): "Disclosure and Asset Returns," *Econometrica*, 71, 105–133.

SINGH, N., AND D. WITTMAN (2001): "Implementation with Partial Verification," *Review of Economic Design*, 6, 63–84.

SPERNER, E. (1928): "Ein Satz über Untermengen einer endlichen Menge," *Math. Zeitschrift*, 27, 544–548.

TOPKIS, D. (1998): *Supermodularity and Complementarity*. Princeton University Press.