Coarse Communication and Institution Design *

Qinggong Wu †

September 8, 2017

Abstract

Many institutions aggregate information for a common objective via coarse communication. Coarseness gives rise to interesting institution design problems which would otherwise be trivial. The paper first elaborates on this point with an analysis of the optimal binary voting systems for the Condorcet Jury Problem, then proposes a unified framework for modeling a general class of information-aggregating institutions. Within this class, it is shown that institution A outperforms institution B for *any* common objective if and only if the underlying communication infrastructure of A can be obtained from that of B by a sequence of elementary operations. Each operation either removes redundant communication instruments from B or introduces effective ones to it. The general analysis is applied to two specific problems. In the first application, it is shown that an optimal generalized voting system has a sequential procedure and a dictatorship-like rule. In the second application, it is shown that data overload can be avoided for an organization with limited data-processing capacity.

1 Introduction

This paper studies information-aggregating institutions in which precise internal communication is not possible.

*I am indebted to Tilman Börgers, Kfir Eliaz, Stephan Lauermann and David Miller for constant support and encouragement throughout this project. I am also grateful to Miguel Ballester, Pierpaolo Battigalli, Jimmy Chan, Mehmet Ekmekci, Shaowei Ke, Daniel Krähmer, Anqi Li, Bart Lipman, Heng Liu, Taejun Oh, Efe Ok, Scott Page, John Quah, Anne-Katrin Roesler, Ariel Rubinstein, Todd Sarver, Ran Shorrer, Philipp Strack, Alexander White, Chenyu Yang and Fudong Zhang for great conversations and feedback.

†Department of Decision Sciences and Managerial Economics, Chinese University of Hong Kong. *Email:* `wqg@baf.cuhk.edu.hk`

It is often very difficult for an informed person to convey his information precisely to an uninformed person. In the first place, the communication instrument available to the informed person, such as words, gestures, or some other medium, may not be sufficient to express the complexity of his information. In addition, upon receiving the message that carries the information, the uninformed person may fail to digest the message to its utmost precision. For example, imagine an expert who wishes to report an important number to a manager. The decimal expansion of this number is very long. To save time, the expert reports this number only to its fourth decimal place. To save memory, the manager memorizes the report only to its second decimal place. In this example, time and memory impose constraints on communication and render it coarse.

Many institutions are established to aggregate information coarsely from a group of people working on a common objective when precise communication within the group is impossible or costly due to restrictions on time, memory, or other resources. For example, a voting system is established to aggregate information coarsely from the public via votes, because gathering each individual's precise opinion about the candidates is costly. Similar institutions include surveys, polls and ratings. As another example, a firm's hierarchical briefing system is established to aggregate information coarsely from different divisions of the firm via briefs, because collecting and processing detailed reports in a centralized fashion is costly. As a third example, when it is difficult for a person to meticulously remember everything he has in mind, he often uses the "mnemonic institution" of taking short notes or forming crude impressions to remind his future self of what he knows in the present.

The infeasibility of precise communication within institutions gives rise to interesting design problems that would otherwise be trivial. Indeed, given common interest induced by the common objective, first-best information aggregation can be easily achieved in equilibrium by an institution permitting precise communication. It is plain to see that, regardless of institutional details, every such institution is reducible to the direct mechanism that implements the efficient outcome. On the other hand, an institution that does not permit precise communication cannot be reduced to a direct mechanism, and consequently institutional details become important to determine its performance.

To elaborate on the design problem regarding institutions not permitting precise communication, this paper first analyzes the optimal binary voting system in a common values environment. In this problem, which is known as the Condorcet Jury Problem, a group of privately informed jurors have to arrive at a verdict to find a defendant guilty or innocent. The jury must use a binary voting system to solve the problem. A binary voting system requires that each juror is given one opportunity to cast a vote of either "guilty" or "innocent". Every binary voting system is characterized by two institutional components: the procedure and the rule. The procedure specifies the number of stages the voting has and who votes in each stage. The procedure may imply simultaneous voting, sequential voting, or a mixture of both. The rule specifies the verdict for each vote profile. The rule may correspond to a simple majority rule, a unanimous rule, or even some non-standard rule.

Three design problems are considered. In the first problem, given that there is no restriction on the number of stages of voting, among the optimal binary voting systems there is a system that has a sequential procedure and a rule under which the last voter serves as a dictator. In the second problem, given that voting has to finish in one stage, among the optimal one-stage systems there is a system that has a weighted majority rule. In the third problem, given that the procedure is fixed and has two stages, it turns out that the optimal system may have a counterintuitive rule: If a voter unilaterally changes his vote from "guilty" to "innocent", the verdict may change from "innocent" to "guilty".

The paper then proposes a framework for modeling a general class of information-aggregating institutions including typical voting systems, hierarchical organizations, and institutions of other sorts. An institution within this class is built on a communication infrastructure which provides each participant with (1) a set of messages for him to convey information to other participants, and (2) a set of units called perceptions for him to receive and process messages from other participants. Communication constraints on the sending side are captured by the fact that the set of messages available to a participant may be smaller than the set of all pieces of information he may wish to convey. Communication constraints on the receiving side are captured by the fact that a participant may be unable to distinguish between distinct message profiles if they correspond to the same perception.

A robust Pareto order is introduced to compare institutions within this class: Institution A is said to dominate institution B if, for *any* common objective, the best equilibrium under A generates a weakly higher common expected payoff than the best equilibrium under B. The purpose of focusing on the dominance order is threefold. First, for a more specific design problem, the designer can use the dominance order to eliminate dominated institutions without having to know the environment parameters that determine the common objective. Second, understanding why an institution dominates another provides insight to understanding the advantages or disadvantages of specific institutional details. Third, from a theoretical perspective, the comparison of institutions in terms of dominance is parallel to the comparison of experiments analyzed in Blackwell (1951). This parallelism is discussed in more detail in Section 2.

The paper provides two characterizations of the dominance order. First, institution A dominates institution B if and only if A can induce weakly more social choice functions in pure strategies than B. Therefore, "better" and "more versatile" turn out to be equivalent regarding institutions. Second, if the set of all pieces of information each player may have is sufficiently rich, then institution A dominates institution B if and only if the underlying communication infrastructure of A can be obtained from that of B by a sequence of operations, each operation either removes redundant messages or perceptions from B, or introduces effective messages or perceptions to B. The sufficiency part is relatively straightforward, although identifying redundant messages and perceptions calls for care, in particular if the institution is complicated. Necessity is more difficult to establish. The argument is based on the observation that there is a social choice function inducible by institution B in pure strategies such that any other institution that can induce the same social choice function in pure strategies must embed the effective part of B, that is, the underlying communication infrastructure of B with redundant messages and perceptions removed. Thus it is possible to construct a sequence of redundancy-reducing operations that transforms B to its effective part, and then there is a sequence of complementing operations that transforms the effective part of B to A

The general analysis is applied to two specific problems. The first application investigates the design problem regarding generalized voting systems. It is shown that among the optimal generalized voting systems there is one that has a sequential procedure, a full disclosure policy, and a rule under which the last voter is always pivotal. The second application analyzes the marginal benefit of perceptions and messages within a hierarchical organization. It is shown that under mild conditions the marginal benefit of either perceptions or messages is always strictly positive. In particular, the result implies that even if an organization has very limited message-processing capacity, it will still strictly benefit from having more available messages.

In the rest of the paper, Section 2 discusses the relevant literature. Section 3 analyzes the Condorcet Jury Problem. Section 4 introduces the general model. Section 5 presents analysis of the dominance order. Section 6 applies the general analysis to two problems. Section 7 concludes. The Appendices include the proofs and some technical details.

2 Literature

Two literatures within the field of mechanism design and implementation have explicitly considered communication constraints. One of them investigates the minimal amount of communication necessary to implement a given social choice function. Within this literature, Nisan and Segal (2006) and Segal (2007) show that for a class of social choice functions the minimal communication mechanisms are generalized price mechanisms, assuming sincere players. Fadel and Segal (2009) and Segal (2010) consider similar implementation questions with strategic players. In spirit this literature may be seen as solving the dual problem to the problem studied in the present paper. The present paper asks for the efficiency maximizing mechanism subject to communication constraints, whereas papers in the above literature ask for the communication minimizing mechanism that achieves certain efficiency level.

The other literature considers mechanism design problems subject to given communication constraints. Recent papers in this literature include Blumrosen et al. (2007), Van Zandt (2007), Kos (2012), Blumrosen and Feldman (2013), Kos (2014), and Mookherjee and Tsumagari (2014). All of the papers assume the presence of conflicts of interest among the players, and a main emphasis of the literature is the interplay between communication constraints and incentive constraints. The present paper, by considering an environment with common interest, removes the tension between information aggregation and preference aggregation, and thus allows a sharper focus on implications of communication constraints alone.

The quest for the optimal voting system for the Condorcet Jury Problem may be viewed as a specific mechanism design exercise regarding institutions subject to communication constraints. The formulation of the Condorcet Jury Problem is historically attributed to Condorcet (1785). Papers discussing information aggregation efficiency of various voting systems include Feddersen and Pesendorfer (1998), McLennan (1998), and Duggan and Martinelli (2001), all of which emphasize asymptotic efficiency as the jury size grows large. For a fixed jury, Costinot and Kartik (2007) show that the optimal voting system is invariant to the possibility of boundedly rational voters. Dekel and Piccione (2000) show that in a symmetric environment, any equilibrium under simultaneous voting stays an equilibrium under sequential voting. They conclude with a note that sequentiality does not bring improvement in terms of information aggregation. The present paper shows that, if a voting system is evaluated by its best equilibrium, instead of by its worst one as is implicitly assumed in Dekel and Piccione (2000), then sequentiality may bring improvement.

As mentioned in the Introduction, the comparison of institutions studied in the present paper is analogous to the comparison of experiments studied in Blackwell (1951). Indeed, similar to an experiment *à la* Blackwell, an institution may be viewed as a device that generates signals (messages) to facilitate decision-making based on the true state (the dispersed information). On the other hand, unlike an experiment, an institution has a multi-player dynamic structure and the signals (messages) are generated endogenously by strategic players. In Blackwell (1951), an experiment dominates another if and only the former can be obtained from the latter by a well-defined transformation. This corresponds in spirit to the finding of the present paper that an institution dominates another if and only if the former can be obtained from the latter by a sequence of well-defined operations.

Chapter 8 of Marschak and Radner (1972) extends Blackwell's single-player model to multiple players. In their model, players move sequentially, and later players make decisions based on their own information and noisy realization of messages from earlier players. The model in the present paper has a similar sequential procedure. In two important aspects the two models are different. First, in the present model players are strategic, whereas they are non-strategic in the model of Marschak and Radner (1972). Second, imperfection of communication in the present model is due to coarseness, whereas it is due to noise in the model of Marschak and Radner (1972). The analysis of Marschak and Radner (1972) is restricted to specific examples and has a different focus.

3 The Condorcet Jury Problem

In this section we view the classical Condorcet Jury Problem from an institution design perspective.

A jury $J = \{1, ..., n\}$ has to reach a verdict on a defendant. The verdict is either "guilty" (G) or "innocent" (I). Let $\omega = G$ denote the event that the defendant is in fact guilty, and $\omega = I$ that the defendant is in fact innocent. $\omega = G$ with probability π where $0 < \pi < 1$, and $\omega = I$ with probability $1 - \pi$. Every juror receives a payoff of 1 if the verdict matches the fact ω, or 0 otherwise.

Each juror i has a private signal x_i that carries some information about ω. Specifically, x_i is independently drawn from a finite, but possibly very large, subset X_i of $\mathbb{R}$ with probability $f^i_\omega(x_i) > 0$ conditional on ω. [1] Assume that for any $i \in J$,

(MLRP) $\quad \frac{f^i_G(x_i)}{f^i_I(x_i)} > \frac{f^i_G(x'_i)}{f^i_I(x'_i)}$ if $x_i > x'_i$.

Assumption MLRP implies that a higher signal carries a stronger evidence for $\omega = G$.

Suppose that, due to a shortage of resources necessary for precise communication among the jurors to fully reveal their information, the jury has to use a **binary voting system** to reach the verdict. In a binary voting system, each juror can send a message once, and the set of messages available to him is $\{G, I\}$. The messages may be interpreted as votes, and the action of sending a message may be interpreted as voting.

Each binary voting system is characterized by two institutional components. The first is the **procedure** that specifies the order in which the jurors vote. The procedure is formulated as a function $r : J \to \mathbb{N}$, with the interpretation that $r(i) > r(j)$ means i votes after j, and $r(i) = r(j)$ means i and j vote

[1] The finiteness assumption is to simplify analysis. See also Footnote **??**.

simultaneously. Assume that a juror can see all previously casted votes. [2] The second institutional component is the **rule** that specifies a verdict for each vote profile. The rule is formulated as a function $d : \{G, I\}^n \to \{G, I\}$ such that $d(z_1, ..., z_n)$ is the verdict given vote profile $(z_1, ..., z_n)$. A binary voting system with procedure r and rule d is denoted as (r, d).

Let $\Sigma_V(r)$ denote the set of all strategy profiles of the game induced by voting system (r, d). (Note that $\Sigma_V(r)$ does not depend on d.) Given the common payoff function and the common prior, all voters' *ex ante* preferences over $\Sigma_V(r)$ can be represented by the same expected payoff function $u(\cdot|r, d)$. The **value** of (r, d) is defined as the highest common expected payoff achieved by any perfect Bayesian equilibrium of the game induced by (r, d). Let $U(r, d)$ denote the value of (r, d). We can Pareto-rank voting systems by their values, and say that the optimal system is the one with the highest value.

Calculating the value of a voting system from definition can be computationally heavy, because it is often tedious to determine the set of equilibria of a game. Fortunately, as asserted by the following lemma, the common interest environment allows us to simplify the calculation by circumventing that step. The lemma, which follows immediately from the proof of Proposition 3 to appear later, generalizes Theorem 1 in McLennan (1998).

Lemma 1. $U(r, d) = \max_{\sigma \in \Sigma_V(r)} u(\sigma|r, d)$.

Voting takes place in multiple stages under a sequential or partially sequential procedure, and may require a long time to finish if there are many stages. It is thus reasonable to consider situations in which the jury can only use a voting system whose procedure has a limited number of stages. Below we study the optimal voting systems in three cases. In the first case, there is no restriction on the number of stages of voting. In the second case, voting has to take place in one stage. In the third case, a particular procedure with two stages is used.

3.1 Unlimited Number of Stages

When there is no restriction on the number of stages, the following proposition asserts that, to find an optimal voting system, it suffices to focus on ones that have a sequential procedure and a rule that depends only on the last voter's vote.

Proposition 1. *Among the optimal binary voting systems there is* (r^*, d^*) *such that:*

1. r^* *is a sequential procedure.*

[2] Here, to keep the example simple, we do not consider partial disclosure of previously casted votes. Partial disclosure policies are discussed in Section 6.

2. d^* *depends only on the last voter's vote.*

The proof is based on the observation that any pure strategy profile under (r, d) can be "replicated" by an outcome-equivalent pure strategy profile under (r^*, d^*). It follows that the best strategy profile under (r^*, d^*) must be no worse than the best strategy profile under (r, d), which then implies that the value of (r^*, d^*) must be no less than that of (r, d) by Lemma 1.

Section A.1 in the Appendices gives a direct proof of this proposition. It will become clear that the proposition can also be derived as a corollary of Proposition 5, which is proved using the machinery to be introduced in Section 5. Moreover, the proof of Proposition 5 gives an alternative explanation for the superiority of (r^*, d^*): The system with a sequential procedure and a dictatorial rule makes full use of the communication instruments allowed by binary voting, whereas any other binary voting system does not.

3.2 One Stage

Suppose that voting must take place in a single stage, that is, the procedure must be simultaneous. Finding an optimal voting system in this case is equivalent to finding an optimal rule for simultaneous voting. The following proposition asserts that, for this quest, it suffices to focus on weighted majority rules. Rule d is a weighted majority rule if there is a vector of non-negative weights $(w_1, ..., w_n)$ and a threshold k such that $d(z_1, ..., z_n) = G$ if and only if $\sum_{i \in J} w_i \mathbf{1}(z_i) \geq k$, where $\mathbf{1}(z_i) = 1$ if $z_i = G$ or $\mathbf{1}(z_i) = 0$ otherwise.

Proposition 2. *Among the optimal rules for simultaneous voting there is a weighted majority rule.*

The proof is built on the observation that, given any vote profile $(z_1, ..., z_n)$, an optimal rule d^* produces verdict G if and only if $\omega = G$ is more likely than $\omega = I$ conditional on players following the best strategy profile s^* under d^* and the realized vote profile being $(z_1, ..., z_n)$. In other words, if the rule is replaced with an uninformed decision maker whose interest is aligned with those of the players, the decision maker would also choose verdict G given $(z_1, ..., z_n)$ based on Bayesian updating. Proposition 2 then follows from the fact that the probability of $\omega = G$ conditional on s^* and $(z_1, ..., z_n)$ is log-linear in the number of G-votes contained in $(z_1, ..., z_n)$.

3.3 Two Stages: Non-Monotonicity

In this subsection we present an example showing that, if the procedure is partially sequential, the optimal voting systems may display counterintuitive properties.

There are three jurors. Suppose the procedure is fixed: Juror 1 votes in the first stage, Jurors 2 and 3 vote simultaneously in the second stage. Again we look for the optimal rule for this particular procedure. The environment parameters are as follows. $\pi = 1 - \pi = 0.5$. Tables 1 and 2 respectively show the conditional probabilities of x_1 and x_2. The conditional probabilities of x_3 are the same as x_2.

Table 1: $f^1_\omega(x_1)$

	$\omega = G$	$\omega = I$
$x_1 = 1$	0.6	0.4
$x_1 = 0$	0.4	0.6

Table 2: $f^2_\omega(x_2)$

	$\omega = G$	$\omega = I$
$x_1 = 1$	0.4	0.1
$x_1 = 1/2$	0.5	0.5
$x_1 = 0$	0.1	0.4

Using Lemma 1 we find multiple optimal rules. The unique best equilibrium under each optimal rule is in cutoff strategies. There is only one optimal rule d^* under which the best equilibrium is "truthful", in the sense that each juror votes G if and only if his signal is above the cutoff. Recall that Assumption MLRP implies higher signals are more indicative of $\omega = G$. Thus in a truthful equilibrium jurors always use a G-vote to express a stronger evidence supporting $\omega = G$, whereas in an equilibrium that is not truthful, a juror sometimes uses an I-vote to express a stronger evidence supporting $\omega = G$. It is reasonable to consider d^* as superior to the other optimal rules, because presumably the jurors can more easily coordinate on a truthful equilibrium.

However, d^* displays a counterintuitive property: It is not monotone in the vote profile. In particular, it is the case that

$$d^*(G, G, I) = I, \qquad d^*(I, G, I) = G,$$

that is, if Juror 1 unilaterally changes his vote from G to I, the verdict changes in the opposite direction from I to G. This property is unexpected because a G vote from Juror 1 carries stronger evidence supporting $\omega = G$ than an I vote.

To explain the phenomenon, notice that in the best equilibrium under d^*, the aggregate evidence contained in the (G, I) vote combination from Jurors 2 and 3 depends on the vote from Juror 1. If Juror 1's vote is G then Juror 2 votes G if $x_2 \in \{1/2, 1\}$, whereas if Juror 1's vote is I then Juror 2 votes G if $x_2 = 1$. Juror 3's strategy is the same as Juror 2's. It is easy to verify that the aggregate evidence contained in the (G, I) vote combination from Jurors 2 and 3 is against, and overrules, the evidence contained in the vote from Juror 1.

4 The Model

In this section we introduce a framework for modeling a general class of institutions that is broad enough to capture many specific real life institutions, including binary voting systems analyzed in the previous section.

4.1 The Common Objective

A group $\mathscr{N} = \{1, ..., N\}$ of players work on a common objective. Each player $i \in \mathscr{N}$ contributes by choosing an action a_i from a finite set A_i. A_i could be a singleton $\{null\}$, that is, the common objective does involve any actual action from player i. We shall see the player can still contribute by providing information. A vector of actions $\mathbf{a} = (a_1, ..., a_N)$ is called an **outcome**. $A = A_1 \times ... \times A_N$ is the set of all feasible outcomes.

The value of an outcome depends on the **state of the world** (hereby **state**), which is an $N-$tuple $\mathbf{x} = (x_1, ..., x_N)$ from a finite set $X = X_1 \times ... \times X_N$. Each player i only observes x_i but not other dimensions of the state. The common prior distirbution of $\mathbf{x}$ is denoted as F. Every player receives the *same* payoff $\phi(\mathbf{a}, \mathbf{x})$ if the outcome is $\mathbf{a}$ and the state is $\mathbf{x}$.

The group may be concerned with multiple common objectives: It could be that they co-operate many times on different common objectives, or it could be that they co-operate only once but are uncertain of which common objective they will face. Assume A and X are the same across common objectives but ϕ and F can vary. We use objective-specific parameters (ϕ, F) to denote a particular common objective.

4.2 The Institution

Within the model, an **institution** refers to the indirect mechanism described as follows. Players move sequentially according to their indices: Player 1 moves first, Player 2 moves second, etc. Every player moves only once. When it is player i's turn to move, he chooses action a_i from A_i, and he also chooses a message m_i from some set M_i to send to players who have not moved yet. M_i is the set of messages provided to player i by the institution. Player N does not send any message. For Player N to have a non-trivial role in the institution, we assume that $|A_N| > 1$.

Players do not observe actions chosen by the other players. A player imperfectly observes the messages he has received. Let $T_i = M_1 \times ... \times M_{i-1}$ denote the set of message profiles that player i may receive. i's observation of received messages is determined by a partition P_i of T_i. Each element p of P_i is called a **perception** of i. i can distinguish between two message profiles if and only if they are in different perceptions of his. Therefore it can be said that player i perceives all message profiles in the same perception $p \in P_i$ as if they are the same.

Define $T = \cup_{i \in \mathcal{N}} T_i$, $P = \cup_{i \in \mathcal{N}} P_i$, and $M = \cup_{i<N} M_i$. The institution is denoted by the tuple (T, P, M).

Below we show various real life institutions that the general model captures.

Example 1. Voting

Consider a voting system that generalizes the binary voting system analyzed in Section 3. A group J of voters have to collectively choose from a set Y of candidates. Each voter $i \in J$ receives a private signal x_i. The value of each candidate is the same to all voters and depends on the vector of private signals $(x_i)_{i \in J}$.

Each voter has to cast one vote from the set of votes Z. [3] To avoid trivial cases assume $|Y| > 1$ and $|Z| > 1$. The voting system consists of three institutional components: the procedure $r : J \to \mathbb{N}$ that assigns each voter i to the $r(i)$th stage of the voting, the rule $d : Z^{|J|} \to Y$ that elects candidate $d(\mathbf{z})$ given vote profile $\mathbf{z}$, and a disclosure policy t which specifies how past votes are disclosed to those who have not voted yet. Examples of disclosure policies include full disclosure, disclosing the votes but not the voters' indices (anonymous voting), etc.

To capture voting system (r, d, t) using the general institution model introduced above, index the voters as $1, ..., |J|$ such that $i > j$ if $r(i) > r(j)$. Let $\mathcal{N} = \{1, ..., |J|, |J|+1\}$ where player $|J|+1$ represents the rule d. The action set A_i is the singleton $\{\overline{a}\}$ for every $i \leq |J|$, whereas $A_{|J|+1} = Y$. Note that only player $|J|+1$ has a non-singleton action set, because he represents d and thus him alone determines the real outcome, that is, the chosen candidate from Y.

For each player $i \leq |J|$, X_i is the set of private signals that voter i may observe. $X_{|J|+1} = \{\overline{x}\}$, implying that the voter representing the rule always receives the uninformative signal $\overline{x}$.

For any outcome $\mathbf{a} = (\overline{a}, ..., \overline{a}, y)$ and vector of private signals $\mathbf{x} = (x_1, ..., x_{|J|}, \overline{x})$, $\phi(\mathbf{a}, \mathbf{x})$ is the expected value of candidate y conditional on each voter i observing private signal x_i.

Voting system (r, d, t) is represented by institution (T, P, M) as follows:

1. For each $i \leq |J|$, M_i is equal to Z, that is, a vote is interpreted as a message provided by the institution.

2. For each $i \leq |J|$, P_i is consistent with voter i's observation of votes from voters $1, ..., i-1$ under procedure r and disclosure policy t. For example, if two vote profiles $(z_1, ..., z_{i-1})$ and $(z'_1, ..., z'_{i-1})$ differ only at

[3]Typically $Z = Y$. The present formulation allows other voting protocols, for example the inclusion of blank votes or abstention.

votes from those who vote simultaneously with i, that is, if $z_j = z'_j$ for every j where $r(j) < r(i)$, then the two vote profiles are in the same perception of i. As another example, if according to the disclosure policy voter i may only observe the number of votes already casted for one particular candidate y, then $(z_1, ..., z_{i-1})$ and $(z'_1, ..., z'_{i-1})$ are in the same perception of i if the number of votes for y from those who voted before i are the same in both vote profiles.

3. For player $|J| + 1$, vote profiles $(z_1, ..., z_{|J|})$ and $(z'_1, ..., z'_{|J|})$ are in the same perception if and only if $d(z_1, ..., z_{|J|}) = d(z'_1, ..., z'_{|j|})$.

It is plain to see that the game induced by voting system (r, d, t) is essentially the same game as that induced by institution (T, P, M) in which player $|J|+1$ is committed to choose $d(z_1, ..., z_{|J|})$ given his perception that contains vote profile $(z_1, ..., z_{|J|})$. [4]

Example 2. Reporting

If we replace the rule in the voting model with an uninformative decision maker, the consequent model captures the situation in which a group of consultants advise an uninformed boss on choosing a project from the set of alternatives Y. The voters are reinterpreted as consultants, Player $|J| + 1$ as the boss, the votes as internal reports, the procedure and the disclosure policy as a protocol that organizes reporting.

Example 3. Organization

An organization has N levels. The official of level i gathers intelligence $x_i \in X_i$, and has to take immediate action $a_i \in A_i$ in response. Moreover, he also sends a message $m_i \in M_i$ to inform the officials of levels above him. The organization is modeled by institution (T, P, M). M_i is interpreted as internal codes available to the official of level i. P_i can either be used to describe how the official of level i interprets the codes he has received, or the code-processing protocol of the organization.

Example 4. Memory

Institution (T, P, M) can also model the dynamic optimization problem of a single person with imperfect recall. There is only one man, who on each day i learns something $x_i \in X_i$, and does something $a_i \in A_i$. Moreover, to remind himself of what he has learnt or done, he takes down a note, or forms a crude impression, in the form of $m_i \in M_i$. P_i represents how the person digests past notes or recalls past impressions. For example, if distinct $(m_1, ..., m_{i-1})$ and $(m'_1, ..., m'_{i-1})$ belong to the same $p \in P_i$, it may reflect the fact that on the ith day impressions $(m_1, ..., m_{i-1})$ and $(m'_1, ..., m'_{i-1})$ strike the person as the same, or the fact that the person does not read his notes

[4] Apart from the modeling artifact that, in the game induced by institution (T, P, M), the voters take the dummy action $\bar{a}$ along the way, and player $|J| + 1$ receives an uninformative signal $\bar{x}$.

very carefully. The decision maker with limited memory studied in Wilson (2014), for example, can be reformulated, with small modifications, as an infinite-horizon extension of the present model in which the decision maker can only recall the note he took down in the previous period. □

Institution (T, P, M) can be schematically depicted as a rooted tree equipped with a partition. Each node of the tree represents a message profile, and each edge represents a message. The root of the tree is the empty message profile. Two nodes are linked if one is extended from the other by one message, and the message is the edge that links the two nodes. The partitioning of the nodes at level i, which are $i-1$ degrees from the root, agrees with the partitioning of T_i by P_i. Figure 1 shows the graph representing the voting system with two voters, the simultaneous procedure, and the rule that chooses verdict G only if both votes are G.

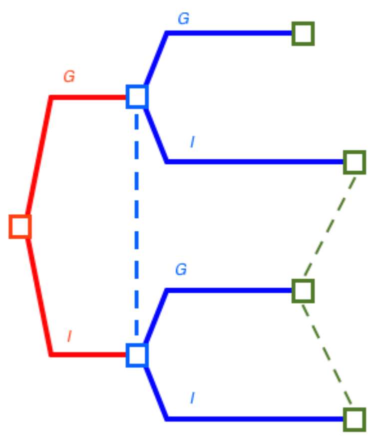

Figure 1: Simultaneous voting, unanimous rule

The graph of an institution visually resembles that of an extensive form game in which histories correspond to message profiles and information sets correspond to perceptions. It should be noted that the graph of an institution does not fully depict the game induced by the institution, because for each player i the choice of action a_i and the acquisition of information x_i are not reflected in the graph.

5 Comparing Institutions

The common interest environment provides us with a natural measure to evaluate an institution.

Definition. The **value** of institution (T, P, M) for common objective (ϕ, F) is the highest common expected payoff achievable by any perfect Bayesian equilibrium of the game induced by (T, P, M) and (ϕ, F). [5]

For a given common objective, all institutions are totally Pareto-ordered by their values. Finding an optimal institution among a set of alternatives becomes a standard optimization problem. However, it is often the case that when an institution is to be established, the objective-specific parameters are unknown to the designer. Moreover, the same institution may be used repeatedly for variable common objectives. Due to these concerns, we would like to Pareto-order institutions without knowledge of the objective-specific parameters. It is thus natural to consider the parameter-free **dominance order** defined as follows.

Definition. Institution (T', P', M') **dominates** another institution (T, P, M) if the value of (T', P', M') is weakly higher than the value of (T, P, M) for *any* (ϕ, F) where ϕ is a real valued function on $A \times X$ and F is a probability function on X.

Clearly the dominance order is reflexive and transitive. It may not be complete, though.

Remark: When an institution is defined, it is assumed that players move sequentially according to their indices. Hence when comparing (T', P', M') and (T, P, M) we implicitly assume that players move in the same order under both institutions. A natural question that follows is whether it is possible to compare two mechanisms, both describable by the general model if players are indexed appropriately, but under which the players *de facto* move in different orders. This indeed is possible, because, after all, if the common objective is the same, then any two mechanisms can be compared by their values. Thus, the dominance order can naturally be extended to the set of all mechanisms without restriction on the order of moves. In this paper, as a first step of the research agenda, we focus only on institutions with the same order of moves.

A pure strategy profile s under institution (T, P, M) induces a **social choice function** $\boldsymbol{\alpha}(\cdot|s) : X \to A$, such that if the players follow s, then in state x the outcome is $\boldsymbol{\alpha}(\mathrm{x}|s)$. Let $C(T, P, M)$ denote the set of all social choice functions inducible by a pure strategy profile under (T, P, M). The following proposition gives a simple characterization of the dominance order.

Proposition 3. *Institution (T', P', M') dominates another institution (T, P, M) if and only if $C(T, P, M) \subset C(T', P', M')$.*

The proposition is a corollary of a more general result, Proposition 8 in Section A.4, that asserts an analogous statement for any two finite mechanisms. Within the present context, the proof is based on the observation that among the strategy profiles that maximize the common expected payoff there is a pure strategy equilibrium. Therefore (T', P', M') must have a

[5]The value exists because the induced game is a finite.

(weakly) higher value than (T,P,M) for any common objective because it can induce more social choice functions in pure strategies. On the other hand, for any pure strategy profile s under (T,P,M) there is a common objective for which $\boldsymbol{\alpha}(\cdot|s)$ is efficient, therefore if (T',P',M') has a weakly higher value for this common objective, (T',P',M') must also be able to induce $\boldsymbol{\alpha}(\cdot|s)$ in pure strategies.

The reader is reminded of the main message of Blackwell (1951) that an experiment is more valuable if and only if it is more informative. The more informative experiment *à la* Blackwell allows more state-to-action mappings for the concerned single-player decision problem. Proposition 3 thus strikes a similar note. Indeed, because of common interest, the institution can be interpreted as a dynamic decision situation that a single player faces, as Example 4 in Section 4 suggests. It should be noted, however, that the information structure induced by an institution is endogenously generated via messages, whereas that induced by an experiment is exogenously generated via noisy signals.

In Blackwell (1951) the "more informative" order of experiments has a very simple structural characterization, which can be verified by examining the distribution functions representing the experiments. The rest of the section is dedicated towards a goal in the same spirit, and gives us a method to compare institutions by examining their structures.

To preview, the structural characterization of the dominance order takes the following form: Institution (T',P',M') dominates institution (T,P,M) if and only if (T',P',M') can be obtained from (T,P,M) by a sequence of operations, each being of one of the following five types: (1) expanding, (2) refining, (3) trimming, (4) relabeling, and (5) merging. Each type of operation involves adding or removing messages, perceptions, or both. Before elaborating on the operations, we first introduce a useful auxiliary concept, the improper institution, which is the "intermediate product" produced in the process of obtaining one institution from another using the operations.

Improper Institutions

An **improper institutions** has a similar message-perception backbone as a (proper) institution introduced in Section 4. The only difference is that the set of messages available to a player in an improper institution may depend on which message profile he has received. Note that the improper institution is a generalization of the (proper) institution. We extend the system of notation denoting components of a (proper) institution to denote the components of an improper institution: For each $i \in \mathscr{N}$, T_i is the set of message profiles player i may receive. P_i is a partition of T_i that represents i's perception. If $i < N$ then for each $h \in T_i$, $M(h)$ is the set of messages available to i if the message profile he receives is h. Denote $T = \cup_{i\in\mathscr{N}} T_i$ and

$P = \cup_{i\in\mathcal{N}} P_i$. The tuple (T, P, M), where M now denotes the correspondence that determines the set of available messages for each received message profile, is now used to represent an improper institution. Clearly, if (T, P, M) is a proper institution then $M(h) = M_i$ for each $h \in T_i$.

For a generic message profile h, let h_j denote the jth component of h, let $h(j)$ denote the first j components of h, and let $|h|$ denote the length of h. Thus for any $h \in T_i$ we have $h = (h_1, ..., h_{i-1})$, $h(j) = (h_1, ..., h_j)$, and $|h| = i - 1$.

Given $h \in T_i$, message profile $g \in T$ is said to be a **descendant** of h, and h is said to be an **ancestor** of g, if g is an extension of h. Clearly, if g is a descendant of h then $h = g(|h|)$. Moreover, if g is a one-component extension of h, then h is said to be the **parent** of g, and g is said to be a **child** of h. Clearly, if h is the parent of g then $h = g(|g| - 1)$. If h is the parent of g and the last component of g is m, we sometimes denote g as $h \times m$.

If distinct g and g' are descendants of h such that (1) $|g| = |g'|$, and (2) $g_j = g'_j$ for any $j \neq |h| + 1$, then g and g' are said to be $h-$**cousins**. Note that two $h-$cousins only differ at the $|h|$th component. [6]

Let $P(h)$ denote the perception $p \in P$ that contains message profile h.

We require that an improper institution (T, P, M) must satisfy the following regularity conditions:

- C1 T_i is nonempty for any $i \in \mathcal{N}$.
- C2 If $h \in T$ then every ancestor of h is in T.
- C3 If $h \in T$ then $h \times m \in T$ if and only if $m \in M(h)$.
- C4 If $P(h) = P(h')$ then $M(h) = M(h')$.

C1 and C2 imply that the graph representing (T, P, M) is a rooted tree with N levels. C3 implies that the tree is generated by M. C4 implies that the sets of available messages given different message profiles in the same perception are the same.

Note that given C4, conditional on perception $p \in P_i$, player i cannot acquire additional information on which particular message profile in p is the one he has received by examining the set of currently available messages. Therefore it causes no ambiguity to define $M(p)$ as the set of messages available to player i conditional on p.

Figure 2 shows the graph of an improper institution.

Given C1-C4, an improper institution induces a well-defined dynamic game in which players sequentially take actions and send messages where the set of available messages might be perception-dependent. The concepts of value and dominance are naturally extended to improper institutions. Proposition 3 extends to improper institutions as well.

[6] Two message profiles with the same parent are also cousins by this definition.

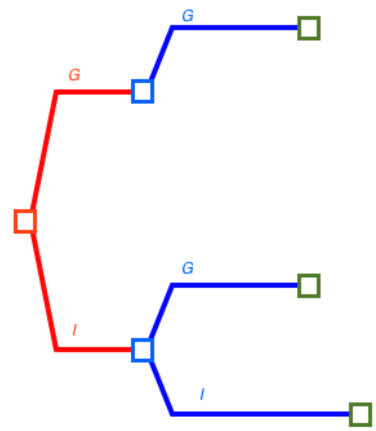

Figure 2

Proposition 4. *Improper institution* (T', P', M') *dominates another improper institution* (T, P, M) *if and only if* $C(T, P, M) \subset C(T', P', M')$.

For the rest of the section we will stop distinguishing between improper institutions and (proper) institutions, as all results apply to improper institutions and thus to proper institutions as special cases. The word “institution” will be used to refer to either.

Expanding

Expanding is the operation of creating a new institution (T', P', M') by adding more messages to an existing institution (T, P, M), while maintaining the perceptibility of the existing message profiles, that is, for any two message profiles h and g in T, they are in the same perception under P' if and only if they are in the same perception under P. The operation is defined formally as follows.

Definition. (T', P', M') is obtained from (T, P, M) by **expanding** if:

E1 $T \subset T'$.

E2 For any $h, g \in T$, $P(h) = P(g)$ if and only if $P'(h) = P'(g)$.

We say that (T, P, M) is a **sub-institution** of (T', P', M') if the latter is obtained from the former by expanding. For example, the institution depicted in Figure 4 is obtained from that depicted in Figure 3 by expanding, in particular, by making the additional message I available to Player 2.

The following lemma states that expansion creates a better institution.

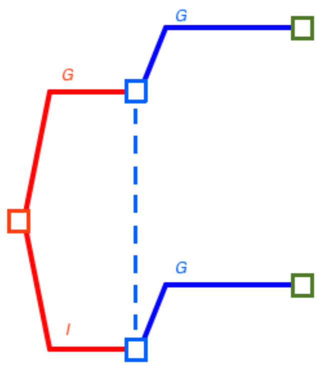

Figure 3

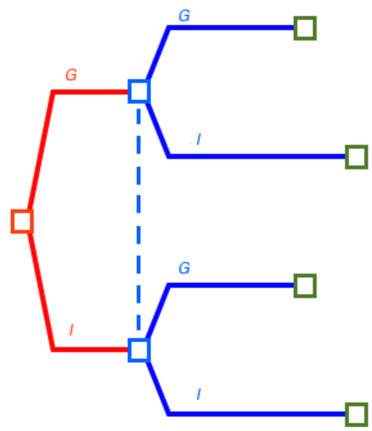

Figure 4

Lemma 2. *If* (T', P', M') *is obtained from* (T, P, M) *by expanding then* (T', P', M') *dominates* (T, P, M).

Refining

Refining is the operation that improves the players' observation of received messages. Refining is formally defined as follows.

Definition. (T', P', M') is obtained from (T, P, M) by **refining** if $T' = T$, $M' = M$, and P'_i is a weak refinement of P_i for each $i \in \mathcal{N}$.

Thus under (T', P', M') a player perceives received messages (weakly) more accurately than he does under (T, P, M). For example, the institution depicted in Figure 6 is obtained from that depicted in Figure 5 by refining, in particular Player 2's observation of received messages is strictly improved.

It is plain to see that any strategy profile of (T, P, M) can be "replicated" in (T', P', M') to produce the same outcome. Therefore (T', P', M') dominates (T, P, M) by Proposition 4. This observation is formally asserted in the following lemma. The proof is omitted.

Lemma 3. *If* (T', P', M') *is obtained from* (T, P, M) *by refining then* (T', P', M') *dominates* (T, P, M).

Trimming

In real life, two words have the same communicative function if they are synonymous. Therefore one of the synonymous words may be viewed as

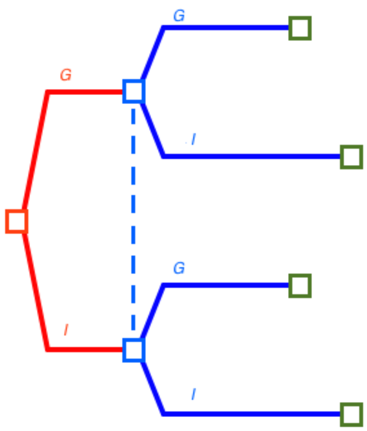

Figure 5

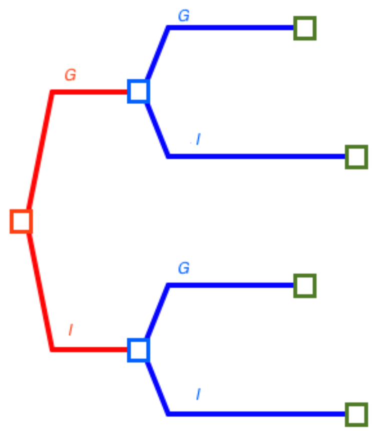

Figure 6

functionally redundant. Excluding the redundant word from the vocabulary does not compromise communication. Trimming is the analogous operation of excluding a redundant message from an institution.

Before elaborating on trimming, it is helpful to first understand synonymity and redundancy within an institution. Whether two words in real life are synonymous or not often depends on the context. In one context they are synonymous; in another they have different meanings. Within an institution (T, P, M), whether two messages available to player i are synonymous or not also depends on the context, and the context is the message profile $h \in T_i$ that player i has received. Given $h \in T_i$, messages $m_i \in M(h)$ and $m_i' \in M(h)$ are considered to be synonymous if:

- Player $i+1$ cannot distinguish between $h \times m_i$ and $h \times m_i'$.
- Regardless of what message m_{i+1} that player $i+1$ sends, player $i+2$ cannot distinguish between $h \times m_i \times m_{i+1}$ and $h \times m_i' \times m_{i+1}$.
- Regardless of what message m_{i+2} that player $i+2$ sends, player $i+3$ cannot distinguish between $h \times m_i \times m_{i+1} \times m_{i+2}$ and $h \times m_i' \times m_{i+1} \times m_{i+2}$.
- And so on for every player who moves after.

Therefore, keeping the messages from everyone else fixed, if player i unilaterally deviates from sending m_i to sending m_i', no other player would perceive the difference. The formal definition is given as follows.

Definition. Fix institution (T, P, M). For any $h \in T_i$ where $i < N$, messages $m_i \in M(h)$ and $m_i' \in M(h)$ are **synonymous given** h **within** (T, P, M) if $P(g) = P(g')$ for any $h-$cousins g and g' where $g, g' \in T$, $g_i = m_i$ and $g_i' = m_i'$.

Within the institution depicted in Figure 7, Player 2's messages G and I are synonymous given message profile (G), because Player 3 cannot distinguish between (G, G) and (G, I). Similarly G and I are synonymous given message profile (L) as well.

If m_i and m_i' are synonymous given every $h \in p$ within (T, P, M) for some $p \in P_i$ then we say m_i and m_i' are **synonymous given** p **within** (T, P, M). Message m_i may be considered as redundant given p. Trimming is the operation that excludes the redundant message m_i from the set of available messages given p. The formal definition is as follows.

Definition. (T', P', M') is obtained from (T, P, M) by **trimming** if there exist $i < N$, $p \in P_i$ and $m_i \in M(p)$ such that:

- T1 (T', P', M') is a sub-institution of (T, P, M). Moreover $h \in T \backslash T'$ implies $h(i-1) \in p$ and $h_i = m_i$.
- T2 m_i is synonymous to some $m_i' \in M(p)$ given p within (T, P, M).

By T1, (T', P', M') is the institution corresponding to player i not provided with message m_i given perception p. T2 emphasizes that m_i is indeed redundant given p.

The institution depicted in Figure 8 is obtained from that depicted in Figure 7 by trimming off message I, which is synonymous to G and is therefore redundant given Player 2's only perception.

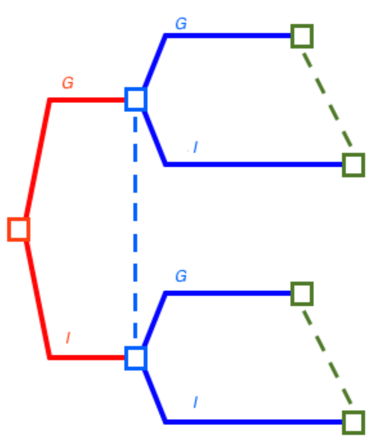

Figure 7

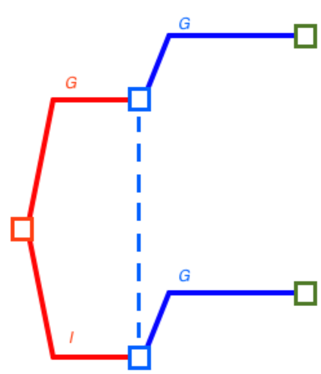

Figure 8

The following lemma asserts that trimming does not change an institution functionally.

Lemma 4. *If* (T', P', M') *is obtained from* (T, P, M) *by trimming then* (T', P', M') *and* (T, P, M) *dominate each other.*

The proof is based on the observation that given perception p, player i cannot use synonymous messages m_i and m_i' to effectively communicate different pieces of information, because if i unilaterally changes from sending m_i to sending m_i' other players cannot perceive the change and hence will not react differently. Therefore, excluding the redundant message m_i from $M(p)$ does not compromise communication.

Relabeling

Relabeling is the operation of changing the message labels of an institution without changing its essential structure. Intuitively, as along as distinct messages have distinct labels, what those labels are should not matter to the use of the messages. In the jury voting case, for example, changing the vote labels from $Guilty$ and $Innocent$ to G and I only changes the institution superficially.

For an institution, the operation of relabeling changes message labels on a perception-by-perception basis. The operation can be thought of as the following process: First relabel the messages available to Player 1 given his only perception. Then relabel the messages available to Player 2 given each of his perceptions. The process continues until Player $N-1$'s messages are relabeled. The following is a formal definition.

Definition. (T', P', M') is obtained from (T, P, M) by **relabeling** if there is a **relabeling function** $\gamma : T \to T'$ such that:

- R1 γ is a bijection.
- R2 γ preserves parent-child relation.
- R3 $P(h) = P(g)$ if and only if $P'(\gamma(h)) = P'(\gamma(g))$.
- R4 For any $i < N$, $p \in P_i$ and $m_i \in M(p)$ there is a message $\kappa(m_i, p)$ such that $\gamma(h \times m_i) = \gamma(h) \times \kappa(m_i, p)$ for any $h \in p$.

The definition is given in terms of the final product instead of the construction. To link the definition to the construction, note that R1, R2 and R3 imply that the graphs of (T, P, M) and (T', P', M') are isomorphic if the edges are label-less. $h \in T$ and $\gamma(h) \in T'$ are "essentially the same" message profile except that the labels of the messages they contain are different. R2 implies that if after relabeling the message profile $(m_1, ..., m_i)$ becomes $(m_1', ..., m_i')$, then for any message profile h whose first i components are $(m_1, ..., m_i)$, the first i components of the relabeled counterpart $\gamma(h)$ are $(m_1', ..., m_i')$. R4 is related to the perception-by-perception basis on which relabeling is conducted: Every edge with label m_i issued from perception p is given the same new label $\kappa(m_i, p)$.

Intuitively, relabeling should be invertible, that is, we should be able to retrieve (T, P, M) from (T', P', M') by "labelling back", where "labelling back" itself an operation of relabeling. The following lemma confirms this intuition.

Lemma 5. *If* (T', P', M') *is obtained from* (T, P, M) *by relabeling with relabeling function* γ *then* (T, P, M) *is obtained from* (T', P', M') *by relabeling with relabeling function* γ^{-1}.

Figure 9 shows an example of relabeling. The institution depicted in Panel (c) is obtained from that depicted in Panel (a) by relabeling, where Panel (b) shows the implied κ described in R4. Note that both edges with the label H issuing from the only perception of Player 2 are relabeled to L, and both edges with label L issuing from the same perception are relabeled to H, as R4 requires that edges with the same label issuing from the same perception are relabeled identically.

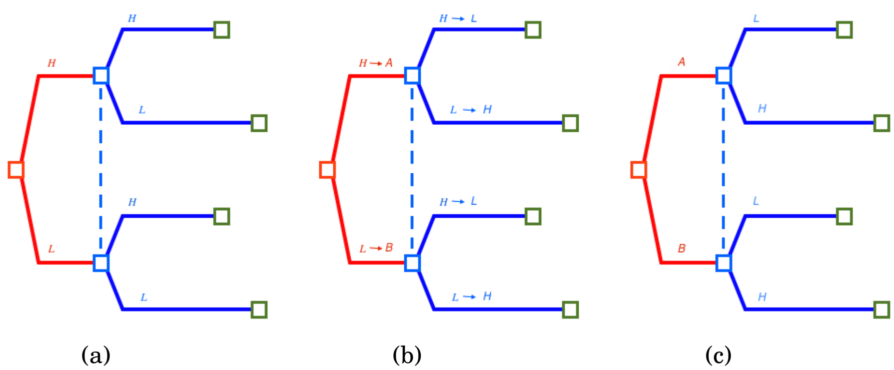

Figure 9

The following lemma asserts that relabeling does not change an institution functionally.

Lemma 6. *If* (T', P', M') *is obtained from* (T, P, M) *by relabeling then* (T', P', M') *and* (T, P, M) *dominate each other.*

Merging

Given institution (T, P, M), the purpose of merging is to produce a new institution in which two perceptions $p \in P_i$ and $q \in P_i$ of some player $i < N$ where $|A_i| = 1$ are combined into one perception $p \cup q$, where the new institution is functionally equivalent to the old institution.

It is useful to first motivate the idea of merging. Suppose a speaker wishes to convey his private information x to a listener using one of the two messages L and R. There are two possible contexts, A and B, which are relevant to the listener's decision problem. If the speaker knows the context, he can use a communication strategy that depends on the context and x. Suppose instead the speaker cannot observe the context but has four available messages $\{m_1, m_2, m_3, m_4\}$. In this situation, by using the following communication strategy the speaker can convey x as precisely as if he knew the context but only had two available messages:

1. Send m_1 if given x he would send L in context A and L in context B.
2. Send m_2 if given x he would send R in context A and R in context B.
3. Send m_3 if given x he would send L in context A and R in context B.
4. Send m_4 if given x he would send R in context A and L in context B.

Merging is based on the same idea. On one hand we make the observation of player i less accurate by combining his perceptions p and q into one perception $p \cup q$. On the other hand we compensate the possible loss of communicative capacity by providing player i with more messages given the combined perception $p \cup q$.

In practice, merging involves the following four steps:

Step 1 If $|M(p)| < |M(q)|$ then the institution is expanded by making more messages available to player i given p, so that the new sets of available messages $\hat{M}(p)$ and $\hat{M}(q)$ are equal in size. Each of the new messages is set to be synonymous to some existing message given p.

Step 2 Relabel the messages in $\hat{M}(p)$ and $\hat{M}(q)$ so that the new sets of messages $\tilde{M}(p)$ and $\tilde{M}(q)$ are the same, not only in terms of size but also in terms of labels.

Step 3 For each pair of distinct messages $(m_i, m_i') \in \tilde{M}(p) \times \tilde{M}(p)$, expand the institution by making a new message $n(m_i, m_i')$ available to player i given both p and q. $n(m_i, m_i')$ is set to be synonymous to m_i given p, and to m_i' given q.

Step 4 Combine p and q into one perception $p \cup q$.

Steps 1 and 2 make the technical preparation so that after combining p and q the resulting structure is an institution. Step 3 introduces redundant messages. However, some of the redundant messages will no longer be redundant after the merge of p and q. Given the less accurate perception $p \cup q$, player i can use message $n(m_i, m_i')$ to convey the private information that he would use m_i to convey if he knew the perception was p and would use m_i' to convey if he knew the perception was q.

The formal definition of merging is complicated, and is thus relegated to Section A.9 in the Appendices.

Figure 10 shows merging in those steps. Panel (a) depicts the original institution. The two perceptions $\{H\}$ and $\{L\}$ of Player 2 are to be merged. Panel (b) depicts the end product of Step 1, that is, an additional message

A, set to be synonymous to B given perception $\{L\}$, is introduced so that the number of messages available to Player 2 given both perceptions are the same. Panel (c) depicts the end product of Step 2, that is, after relabeling, messages available to Player 2 given both perceptions have the same labels. Panel (d) depicts the end product of Step 3, that is, introducing additional redundant messages to both perceptions, where, for example, $n(H, L)$ is set to be synonymous to H given perception $\{H\}$, and to L given perception $\{L\}$. Panel (d) depicts the end product of Step 4, that is, combining the two perceptions.

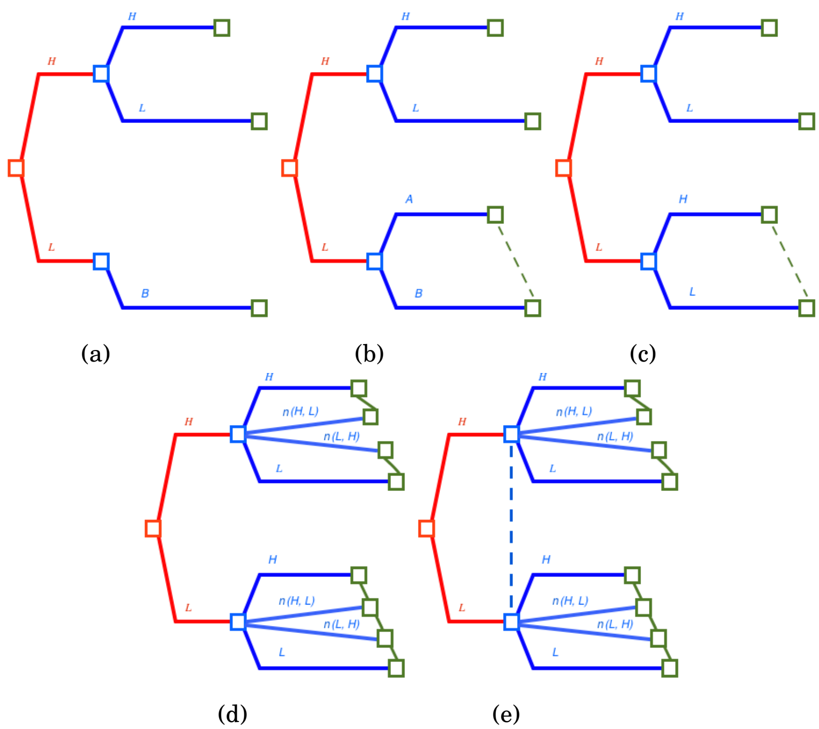

Figure 10

Steps 1, 2 and 3 involve either relabeling or introducing redundant messages (the inverse operation of trimming), and therefore by Lemmas 6 and 4 do not change the institution functionally. Following the intuition of the earlier example, going from Step 3 to Step 4 should not change the institution functionally either. The following lemma confirms the intuition.

Lemma 7. *If* (T', P', M') *is obtained from* (T, P, M) *by merging then* (T', P', M') *and* (T, P, M) *dominate each other.*

Now we are ready to state the structural characterization of dominance.

Theorem 1. *Given two institutions* (T, P, M) *and* (T', P', M')*:*

1. *(T', P', M') dominates (T, P, M) if (T', P', M') is obtained from (T, P, M) by a sequence of operations of expanding, refining, trimming, relabeling or merging.*
2. *For any $i \in \mathscr{N}$ there is $t_i \in \mathbb{N}$ such that if $|X_i| \geq t_i$ then (T', P', M') dominates (T, P, M) only if (T', P', M') can be obtained from (T, P, M) by a sequence of operations of expanding, refining, trimming, relabeling or merging.*

Part 1, the "if" direction, of the theorem is an immediate consequence of Lemmas 2, 3, 4, 6 and 7.

The proof of Part 2, the "only if" direction, can be broken into the following steps. First, it is clear that after applying merging operations to (T, P, M) for finitely many times we can obtain some $(\hat{T}, \hat{P}, \hat{M})$ such that $|\hat{T}_i| = 1$ for any $i < N$ where $|A_i| = 1$. If $|X_i|$ is sufficiently large for each i, there is a strategy profile $\hat{s} \in S(\hat{T}, \hat{P}, \hat{M})$ such that every non-redundant message in $(\hat{T}, \hat{P}, \hat{M})$ is utilized. Moreover, any institution (T', P', M') that can induce $\boldsymbol{\alpha}(\cdot|\hat{s})$ in pure strategies must embed the relabeled version of $(\hat{T}, \hat{P}, \hat{M})$'s non-redundant backbone (T^*, P^*, M^*) as its sub-institution. (T^*, P^*, M^*) is shown to be obtained from $(\hat{T}, \hat{P}, \hat{M})$ by refining and trimming. Since the relabeled version of (T^*, P^*, M^*) is embedded in (T', P', M'), (T', P', M') can be obtained from (T^*, P^*, M^*) by relabeling and expanding.

If the state space is not rich enough, that is, if $|X_i| < t_i$ for some $i \in \mathscr{N}$, then Part 2 of Theorem 1 need not be true. Indeed, if $|X_i|$ is sufficiently small for every $i \in \mathscr{N}$ then (T, P, M) and (T', P', M') may both accommodate *precise* communication, despite that one may not be obtainable from the other by a sequence of the five types of operations. It is worth noting that the value of t_i depends on the dominated institution (T, P, M) only.

6 Applications

6.1 Voting Revisited

In Example 1 of Section 4 we have described how to model voting system (r, d, t) in terms of an institution. Because voting systems are finite mechanisms for a common objective, the definition of value extends to them, and the dominance order can also be extended to compare them. Let $C_V(r, d, t)$ denote the set of all mappings from $X_1 \times ... \times X_{|J|}$ to Y inducible in pure strategies under voting system (r, d, t). A result analogous to Proposition 3 holds for voting systems.

Lemma 8. *Voting system (r', d', t') dominates another voting system (r, d, t) if and only if $C_V(r, d, t) \subset C_V(r', d', t')$.*

Like Proposition 3, Lemma 8 is also a corollary of Proposition 8, because Proposition 8 applies to any pair of finite mechanisms.

Let (T, P, M) be the institution representing (r, d, t). Recall that player $|J|+1$ represents the rule and $P(\mathbf{z}) = P(\mathbf{z}')$ if and only if $d(\mathbf{z}) = d(\mathbf{z}')$ for any $\mathbf{z}, \mathbf{z}' \in T_{|J|+1}$, where $T_{|J|+1}$ is the set of all complete vote profiles. Let $S(T, P, M|d)$ denote the set of pure strategy profiles of (T, P, M) satisfying the following condition:

V1 For any $\mathbf{z} \in T_{|J|+1}$, player $|J|+1$ chooses the action $d(\mathbf{z})$ given perception $P(\mathbf{z})$.

V1 requires player $|J| + 1$ to exactly follow the choice rule determined by d. It is straightforward to see that for any pure strategy profile s under voting system (r, d, t) there is a corresponding strategy profile $s' \in S(T, P, M|d)$ under institution (T, P, M) such that s and s' lead to the same choice of $y \in Y$ given any vector of private signals $(x_1, ..., x_{|J|})$, and vice versa, because, as discussed in Example 1 in Section 4, the game induced by (r, d, t) and the game induced by (T, P, M) are essentially the same if player $|J| + 1$ has to choose according to d.

Let $C(T, P, M|d)$ denote the set of all social choice functions inducible by any $s \in S(T, P, M|d)$. The following lemma results immediately from the observation in the previous paragraph. The proof is omitted.

Lemma 9. *Voting system (r', d', t') dominates another voting system (r, d, t) if and only if $C(T, P, M|d) \subset C(T', P', M'|d')$ where (T, P, M) and (T', P', M') are institutions respectively representing (r, d, t) and (r', d', t').*

It is possible to compare voting systems by analyzing the institutions representing them. However, it should be noted that one institution dominating another institution is usually not sufficient for the voting system that one represents to dominate the voting system that the other represents, due to the additional constraint imposed on the strategy of player $|J| + 1$. Despite this caveat, the machinery developed in Section 5 still provides tools for us to conclude the following results.

Proposition 5. *For any voting system (r, d, t):*

1. *(r, d, t) is dominated by voting system (r, d, t') where t' is the full disclosure policy.*
2. *If t is the full disclosure policy, then (r, d, t) is dominated by voting system (r', d, t) where r' is a sequential procedure.*
3. *If t is the full disclosure policy, then (r, d, t) is dominated by voting system (r, d', t) where d' is a rule under which the collective choice is not determined before voting in the last stage (according to r) has taken place.*

Parts 1 and 2 are based on the observation that allowing full disclosure or making the procedure sequential corresponds to refining the institution that represents the voting system. Moreover, since these modifications do not change player $|J|+1$'s perception, player $|J|+1$ can still choose according to the rule d.

Part 3 is based on the observation that, if the collective choice is determined before voting in the last stage has taken place, then votes in the last stage become redundant, because they cannot effectively carry private information from those who vote in the last stage to affect the collective choice. Changing the rule to one that allows consideration of those last stage votes renders them useful and thus leads to improvement.

An immediate implication of Proposition 5 is that, to find the optimal voting system for any collective choice problem, it is sufficient to focus on ones that have a sequential procedure, full disclosure policy, and a rule under which the last voter is always pivotal.

It should be noted, however, that the last voter being always pivotal does not mean that the rule depends entirely on his vote. For example, earlier votes may effectively determine the set of candidates that the last voter can choose from.

6.2 The Benefit of Complexity

An institution offers two kinds of instruments that facilitate communication: messages and perceptions. A more complex institution has more messages and perceptions. Lemmas 2 and 3 imply that complex institutions weakly outperform less complex ones. In this section we investigate whether the benefit of additional complexity is always strictly positive.

The complexity of institution (T, P, M) has two dimensions, one that concerns the messages and the other the perceptions. The message-complexity of (T, P, M) is measured by the vector $(|M_1|, ..., |M_{N-1}|)$. The perception-complexity of (T, P, M) is measured by the vector $(|P_1|, ..., |P_N|)$.

We ask two questions:

1. Whether increasing the perception-complexity of an institution while keeping the message-complexity fixed leads to a strictly better institution in terms of dominance.

2. Whether increasing the message-complexity of an institution while keeping the perception-complexity fixed leads to a strictly better institution in terms of dominance.

The first question can be thought of as concerning the situation in which messages are costly to provide, whereas perceptions are relatively cheap, so that it is worthwhile to increase the number of perceptions as long as it strictly improves the institution. The second question can be thought of as concerning the situation in which perceptions are costly but messages are cheap.

For the rest of the subsection we assume that $|X_i|$ is very large for each $i \in \mathscr{N}$, that is, it is arbitrarily close to ∞, so that the lower bound requirement on $|X_i|$ in Part 2 of Theorem 1 is satisfied for any institution we are going to consider. Moreover, assume that every player has a non-singleton action set. The assumptions are important to the results of the present subsection.

The answer to the first question is affirmative.

Proposition 6. *Any institution (T, P, M) where $|P_i| < |T_i|$ for some $i \in \mathscr{N}$ is strictly dominated by some institution (T', P', M') where*

1. $|M'_j| = |M_j|$ *for any* $j < N$.
2. $|P'_j| \geq |P_j|$ *for any* $j \in \mathscr{N}$.

The proof is based on the observation that if $|P'_i| > |P_i|$ for some i then (T, P', M) is not dominated by (T, P, M), because no operation of expanding, refining, trimming, relabeling or merging can decrease the number of perceptions of i (merging does not, because it only merges perceptions of players with singleton action sets). If player i's observation of received message profiles is not perfect under P_i (implied by $|P_i| < |T_i|$), then refining (T, P, M) by strictly refining P_i strictly improves the institution.

The answer to the second question is also affirmative if the institution is "mildly" complex.

Proposition 7. *Suppose $N \geq 3$. Any institution (T, P, M) where $|P_{i+1}| \geq 2$ for some $i \geq 2$ is strictly dominated by some institution (T', P', M') where*

1. $|P'_j| = |P_j|$ *for any* $j \in \mathscr{N}$.
2. $|M'_j| \geq |M_j|$ *for any* $j < N$.

The result may not seem surprising at first sight, but let us illustrate a concern which would suggest that additional messages might be of no additional value at all. Suppose there are only two players: the speaker and the listener. The listener has two perceptions. Clearly, if there are already two messages available to the speaker, any additional message is going to be redundant because it will be synonymous to one of the existing messages. This example, which shows that the decision maker (the listener) is not able to make use of more data (messages) because of the constraints on his data-processing capacity (the number of perceptions), reflects a prominent phenomenon, termed as data overload, in many real life situations.

Since an institution may face stringent message-processing constraints due to limited perceptions, it is natural to expect that data overload will eventually occur, in particular when the existing message-complexity is already very high. However, Proposition 7 implies that even if there is only one player that has multiple perceptions, data overload can still be avoided regardless of the message-complexity of the existing institution. That there are more than two players within the institution is crucial for this result. Indeed, if there are only two players then data overload will eventually occur, as in the speaker-listener example. However, if there are more than two players, it is possible to simultaneously enlarge the message sets for multiple players and carefully arrange how other players perceive message profiles containing these newly introduced messages, so that no additional redundancy is created by the modification.

7 Concluding Remarks

This paper proposes a framework for modeling a general class of information-aggregating institutions, introduces a robust Pareto order on institutions thus modeled, and derives two characterizations of this order.

It is not difficult to extend the model to capture more complex institutions, for example, those in which the players engage in conversation-like interactive communication, those in which actions are observable to certain degree, or those in which the players take actions after the communication phase is over. In fact, any mechanism that tackles a common objective by information aggregation can be captured by a straightforward extension of the present model, because the essential part of the model is no more than a partial structure of the extensive form game induced by the corresponding mechanism. As a generalization of Proposition 3, Proposition 8 in the Appendices provides a characterization of the dominance order on all finite mechanisms. It is natural to ask if Theorem 1 can also be generalized so that other mechanisms can be compared structurally in a similar way. As a first step in this direction, it is worthwhile investigating whether institutions that only differ in the order in which players move may be compared structurally.

To extend the analysis in a different direction, we can consider mildly relaxing the common interest assumption to the extent that institutions can still be Pareto-ordered in a non-trivial way. One possibility, for example, is that every player's payoff only depends on his own action. Along with some extension to the model, we may compare pure information sharing systems, for example social networks, in which there is no need to coordinate actions.

References

Blackwell, David (1951), "Comparison of experiments." In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 93–102, University of California Press, Berkeley, Calif.

Blumrosen, Liad and Michal Feldman (2013), "Mechanism design with a restricted action space." *Games and Economic Behavior*, 82, 424 – 443.

Blumrosen, Liad, Noam Nisan, and Ilya Segal (2007), "Auctions with severely bounded communication." *J. Artif. Int. Res.*, 28, 233–266.

Condorcet, Marquis de (1785), *Essai sur l'application de l'analyse á la probabilité des décisions rendues á la pluralité des voix*. Paris.

Costinot, Arnaud and Navin Kartik (2007), "On optimal voting rules under homogeneous preferences."

Dekel, Eddie and Michele Piccione (2000), "Sequential voting procedures in symmetric binary elections." *Journal of Political Economy*, 108, pp. 34–55.

Duggan, John and Cesar Martinelli (2001), "A bayesian model of voting in juries." *Games and Economic Behavior*, 37, 259 – 294.

Fadel, Ronald and Ilya Segal (2009), "The communication cost of selfishness." *Journal of Economic Theory*, 144, 1895 – 1920.

Feddersen, Timothy and Wolfgang Pesendorfer (1998), "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting." *The American Political Science Review*, 92, pp. 23–35.

Kos, Nenad (2012), "Communication and efficiency in auctions." *Games and Economic Behavior*, 75, 233 – 249.

Kos, Nenad (2014), "Asking questions." *Games and Economic Behavior*, 87, 642 – 650.

Marschak, Jacob and Roy Radner (1972), *Economic theory of teams*. Cowles Foundation for Research in Economics at Yale University, Yale University Press, New Haven, London.

McLennan, Andrew (1998), "Consequences of the condorcet jury theorem for beneficial information aggregation by rational agents." *The American Political Science Review*, 92, pp. 413–418.

Mookherjee, Dilip and Masatoshi Tsumagari (2014), "Mechanism Design with Communication Constraints." *Journal of Political Economy*, 122, 1094 – 1129.

Nisan, Noam and Ilya Segal (2006), "The communication requirements of efficient allocations and supporting prices." *Journal of Economic Theory*, 129, 192 – 224.

Segal, Ilya (2007), "The communication requirements of social choice rules and supporting budget sets." *Journal of Economic Theory*, 136, 341 – 378.

Segal, Ilya R. (2010), "Nash implementation with little communication." *Theoretical Economics*, 5, 51–71.

Van Zandt, Timothy (2007), "Communication complexity and mechanism design." *Journal of the European Economic Association*, 5, pp. 543–553.

Wilson, Andrea (2014), "Bounded memory and biases in information processing." *Econometrica*, 82, 2257–2294.

A Appendices

A.1 Proof of Proposition 1

Proof. Choose any voting system (r, d) and $s \in \Sigma_V(r)$. There is a sequential procedure r^* such that $r^*(i) > r^*(j)$ for any $i, j \in J$ such that $r(i) \geq r(j)$. Thus if j's vote is observable to i under r then it is also observable to i under r^*. Let $s_i(x_i, (z_j)_{r(j)<r(i)})$ denote i's strategy given past votes $(z_j)_{r(j)<r(i)}$ and signal x_i. Construct $s' \in \Sigma_V(r^*)$ such that for every $i \in J$, $s'_i(x_i, (z'_j)_{r^*(j)<r^*(i)}) = s_i(x_i, (z_j)_{r(j)<r(i)})$ if $z'_j = z_j$ for any j such that $r(j) < r(i)$. Obviously s and s' are outcome equivalent, implying $u(s|r, d) = u(s'|r^*, d)$. Thus $\max_{\sigma \in \Sigma_V(r)} u(\sigma|r, d) \leq \max_{\sigma \in \Sigma_V(r^*)} u(\sigma|r^*, d)$.

Let k be the last voter according to r^*. Construct d^* such that $d^*(z_1, ..., z_n) = z_k$. Choose any $\hat{s} \in \Sigma_V(r^*)$. Construct $\tilde{s} \in \Sigma_V(r^*)$ such that $\tilde{s}$ and $\hat{s}$ agree for every player $i \neq k$, and for any signal x_k and past votes $\mathbf{z}_{-k}$,

$$\tilde{s}_k(x_k, \mathbf{z}_{-k}) = \begin{cases} d(\mathbf{z}_{-k}, G) & \text{if } d(\mathbf{z}_{-k}, G) = d(\mathbf{z}_{-k}, I) \\ \hat{s}_k(x_i, \mathbf{z}_{-k}) & \text{if } d(\mathbf{z}_{-k}, G) \neq d(\mathbf{z}_{-k}, I) \ . \end{cases}$$

It is straightforward to verify that $\tilde{s}$ and $\hat{s}$ are outcome equivalent, implying $u(\hat{s}|r^*, d) = u(\tilde{s}|r^*, d^*)$. Thus $\max_{\sigma \in \Sigma_V(r^*)} u(\sigma|r^*, d) \leq \max_{\sigma \in \Sigma_V(r^*)} u(\sigma|r^*, d^*)$. It follows that $\max_{\sigma \in \Sigma_V(r)} u(\sigma|r, d) \leq \max_{\sigma \in \Sigma_V(r^*)} u(\sigma|r^*, d^*)$, implying $U(r, d) \leq U(r^*, d^*)$ by Lemma 1. □

A.2 Proof of Proposition 2

For any procedure r and $s \in \Sigma_V(r)$, let $\Pr(G|\mathbf{z}, s)$ denote the probability that $\omega = G$ conditional on the jurors following s and the realized votes are $\mathbf{z}$. The proof is assisted by the following lemma.

Lemma 10. *For any procedure r, if $d^* \in \text{argmax}_d U(r, d)$ then for any $s^* \in \Sigma_V(r)$ such that $u(s^*|r, d^*) = U(r, d^*)$ and vote profile $\mathbf{z} \in \{G, I\}^n$,*

$$d^*(\mathbf{z}) = \begin{cases} G & \textit{if} \quad \Pr(G|\mathbf{z}, s^*) > 0.5, \\ I & \textit{if} \quad \Pr(G|\mathbf{z}, s^*) < 0.5. \end{cases}$$

Proof. Fix procedure r. Suppose there is $d^* \in \text{argmax}_d U(r, d)$ and $s^* \in \Sigma_V(r)$ where $u(s^*|r, d^*) = U(r, d^*)$ such that d^* does not satisfy the condition in the lemma. Define $K = \Big\{\mathbf{z} \in \{G, I\}^n : d^*(\mathbf{z}) = I \text{ and } \Pr(G|\mathbf{z}, s^*) > 0.5\Big\}$ and $L = \Big\{\mathbf{z} \in \{G, I\}^n : d^*(\mathbf{z}) = G \text{ and } \Pr(G|\mathbf{z}, s^*) < 0.5\Big\}$. By assumption $K \cup L \neq \emptyset$. Let $q_\omega(\mathbf{z})$ be the probability that the realized vote profile is $\mathbf{z}$

conditional on ω and the jurors following s^*. Construct rule d' such that for any $\mathbf{z} \in \{G, I\}^n$,

$$d'(\mathbf{z}) = \begin{cases} G & \text{if} \quad \Pr(G|\mathbf{z}, s^*) > 0.5, \\ I & \text{if} \quad \Pr(G|\mathbf{z}, s^*) < 0.5, \\ d^*(\mathbf{z}) & \text{if} \quad \Pr(G|\mathbf{z}, s^*) = 0.5. \end{cases}$$

It follows that $d'(\mathbf{z}) = d^*(\mathbf{z})$ if $\mathbf{z} \notin K \cup L$. We have

$$u(s^*|r, d') - u(s^*|r, d^*) = \sum_{\mathbf{z} \in K} \Big(\pi q_G(\mathbf{z}) - (1-\pi) q_I(\mathbf{z})\Big) + \sum_{\mathbf{z} \in L} \Big((1-\pi) q_I(\mathbf{z}) - \pi q_G(\mathbf{z})\Big).$$

If $\mathbf{z} \in K$ then $\Pr(G|\mathbf{z}, s^*) = \frac{\pi q_G(\mathbf{z})}{\pi q_G(\mathbf{z}) + (1-\pi) q_I(\mathbf{z})} > 0.5$, implying $\pi q_G(\mathbf{z}) - (1 - \pi) q_I(\mathbf{z}) > 0$. Similarly if $\mathbf{z} \in L$ then $(1-\pi) q_I(\mathbf{z}) - \pi q_G(\mathbf{z}) > 0$. It follows that $u(s^*|r, d') - u(s^*|r, d^*) > 0$ because $K \cup L \neq \emptyset$. Therefore $U(r, d') > U(r, d^*)$, contradicting the assumption that $d^* \in \text{argmax}_d U(r, d)$. $\square$

Proof of Proposition 2

Fix the simultaneous procedure r. Choose any $d' \in \text{argmax}_d U(r, d)$ and $s' \in \Sigma_V(r)$ such that $u(s'|r, d') = U(r, d')$. Lemma 10 implies that for any vote profile $\mathbf{z} \in \{G, I\}^n$,

$$d'(\mathbf{z}) = \begin{cases} G & \text{if} \quad \Pr(G|\mathbf{z}, s') > 0.5, \\ I & \text{if} \quad \Pr(G|\mathbf{z}, s') < 0.5. \end{cases}$$

For each $i \in J$ let p^i_ω denote the probability that juror i votes G conditional on ω and s'. Note that since $f^i_\omega(x_i) > 0$ for any $i \in J$, $x_i \in X_i$ and $\omega \in \{G, I\}$, $p^i_I = 0$ if and only if $p^i_G = 0$. Construct s^* that satisfies the following for any $i \in J$:

- If $p^i_G \notin \{0, 1\}$:
  - If $p^i_G / p^i_I \geq 1$ then s^* prescribes the same strategy for i as s'.
  - If $p^i_G / p^i_I < 1$ then given any signal x_i, i votes G with the same probability that he votes I given x_i under s'.
- If $p^i_G \in \{0, 1\}$ then i votes G with probability 0.5 regardless of x_i.

Construct d^* such that $d^*(\mathbf{z}) = G$ if and only if $\Pr(G|\mathbf{z}, s^*) \geq 0.5$. It is straightforward to verify that $u(s'|r, d') = u(s^*|r, d^*)$. Let t^i_ω denote the probability that juror i votes G conditional on ω and s^*. Clearly $0 < t^i_\omega < 1$ for $\omega \in \{G, I\}$. Moreover $t^i_G / t^i_I = p^i_G / p^i_I$ if $p^i_G / p^i_I \geq 1$, $t^i_G / t^i_I = (1 - p^i_G)/(1 - p^i_I)$ if $p^i_G / p^i_I < 1$, and $t^i_G / t^i_I = 1$ if $p^i_G \in \{0, 1\}$. Consequently $t^i_G / t^i_I \geq 1$ for any $i \in J$. Recall that for any vote profile $\mathbf{z} = (z_1, ..., z_n)$, $\mathbf{1}(z_i) = 1$ if $z_i = G$ or 0 otherwise. We have

$$\Pr(G|\mathbf{z}, s^*) = 1\Big/\Big[1 + \frac{(1-\pi)}{\pi} \prod_{i\in J, \mathbf{1}(z_i)=1} \frac{t_I^i}{t_G^i} \prod_{i\in J, \mathbf{1}(z_i)=0} \frac{1-t_I^i}{1-t_G^i}\Big].$$

$\Pr(G|\mathbf{z}, s^*) \geq 0.5$ if and only if

$$\frac{(1-\pi)}{\pi} \prod_{i\in J, \mathbf{1}(z_i)=1} \frac{t_I^i}{t_G^i} \prod_{i\in J, \mathbf{1}(z_i)=0} \frac{1-t_I^i}{1-t_G^i} \leq 1,$$

or equivalently

$$\begin{aligned}
&\log\frac{1-\pi}{\pi} + \sum_{i\in J, \mathbf{1}(z_i)=1} \log\frac{t_I^i}{t_G^i} + \sum_{i\in J, \mathbf{1}(z_i)=0} \log\frac{1-t_I^i}{1-t_G^i} \leq 0 \\
\Longrightarrow\ &\log\frac{1-\pi}{\pi} + \sum_{i\in J} \mathbf{1}(z_i)\log\frac{t_I^i}{t_G^i} + \sum_{i\in J}(1-\mathbf{1}(z_i))\log\frac{1-t_I^i}{1-t_G^i} \leq 0 \\
\Longrightarrow\ &\sum_{i\in J} \mathbf{1}(z_i)\log\frac{1-t_I^i}{1-t_G^i}\frac{t_G^i}{t_I^i} \geq \log\frac{1-\pi}{\pi} + \sum_{i\in J}\log\frac{1-t_I^i}{1-t_G^i}.
\end{aligned}$$

Let $w_i = \log\frac{1-t_I^i}{1-t_G^i}\frac{t_G^i}{t_I^i}$ and $k = \log\frac{1-\pi}{\pi} + \sum_{i\in J}\log\frac{1-t_I^i}{1-t_G^i}$. $w_i \geq 0$ because $\frac{1-t_I^i}{1-t_G^i}\frac{t_G^i}{t_I^i} \geq 1$. Thus $d^*(\mathbf{z}) = G$ if and only if $\sum_{i\in J} w_i\mathbf{1}(z_i) \geq k$. □

A.3 Notation for Proofs of Results in Section 5

Introduce the following notation for the game induced by (T, P, M), where (T, P, M) can either be an institution or an improper institution. The notation will be used throughout the Appendices for the proofs of results in Section 5.

- $S(T, P, M)$: the set of all *pure* strategy profiles of the game induced by (T, P, M).
- (x_i, p): a typical information set of player i, where x_i is to his private information about the state, and p is his perception that contains the message he has received.
- For $s \in S(T, P, M)$, $i \in \mathscr{N}$ and $\mathbf{x} = (x_1, ..., x_N) \in X$,

  $a_i(x_i, p|s)$: i's choice of action under s given (x_i, p).

  $m_i(x_i, p|s)$: (if $i < N$) i's choice of message under s given (x_i, p).

  $\rho_i(\mathbf{x}|s)$: the message profile i receives conditional on x and s.

$\alpha_i(\mathbf{x}|s)$: the action i takes conditional on $\mathbf{x}$ and s.

$\mu_i(\mathbf{x}|s)$: the message i sends conditional on $\mathbf{x}$ and s.

$\rho_j(\mathbf{x}|s, i, h, m_i)$: the message profile that player $j > i$ receives conditional on (1) every player after i follows s, (2) player i receives $h \in T_i$, (3) player i sends message $m_i \in M(h)$.

The following equalities hold by definition:

$$\alpha_i(\mathbf{x}|s) = a_i(x_i, Pt(\rho_i(\mathbf{x}|s))|s),$$
$$\mu_i(\mathbf{x}|s) = m_i(x_i, P(\rho_i(\mathbf{x}|s))|s),$$
$$\rho_{i+1}(\mathbf{x}|s) = \rho_i(\mathbf{x}|s) \times \mu_i(\mathbf{x}|s).$$

A.4 Proof of Proposition 3

We will prove a more general result which implies Proposition 3 as a corollary.

It is straightforward to extend the definition of value to any mechanism Γ whose set of outcomes is a subset of A. Then we can also extend the definition of dominance: Mechanism Γ' dominates mechanism Γ if the value of Γ' is weakly higher than the value of Γ for any common objective.

Fix a mechanism Γ. Let $S(\Gamma)$ denote the set of all pure strategy profiles of Γ. Let $C(\Gamma)$ denote the set of all social choice functions inducible by a pure strategy profile of Γ. For $s \in S(\Gamma)$ let $v(s|\Gamma, \phi, F)$ denote the common expected payoff achieved by s in the game induced by Γ and (ϕ, F). Let $V(\Gamma, \phi, F)$ denote the value of Γ for (ϕ, F).

First we show a lemma.

Lemma 11. $V(\Gamma, \phi, F) = \max_{s \in S(\Gamma)} v(s|\Gamma, \phi, F)$ *for any finite mechanism* Γ *and common objective* (ϕ, F)*.*

Proof: Fix Γ and (ϕ, F). Let $\Sigma(\Gamma)$ denote the set of all strategy profiles of Γ. For $\epsilon \in (0, \bar{\epsilon})$ where $\bar{\epsilon}$ is sufficiently small let $\Gamma(\epsilon)$ denote the perturbed version of Γ such whenever a player chooses a generic action (to be distinguished from the action a_i that player i contributes to the common objective) as a realization of a possibly mixed strategy, his chosen action will realize with probability $1 - (n-1)\epsilon$ where n is the total number of generic actions available at this point, and each of the other $n-1$ generic actions will realize with probability ϵ.

For any $\sigma \in \Sigma(\Gamma)$ let $w(\sigma, \epsilon)$ denote the common expected payoff achieved by σ in $\Gamma(\epsilon)$. Fix $\epsilon \in (0, \bar{\epsilon})$. $\text{argmax}_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$ is nonempty because $\Sigma(\Gamma)$ is compact and $w(\cdot, \epsilon)$ is continuous in its first argument. Choose $\hat{\sigma} \in \text{argmax}_{\sigma \in \Sigma} w(\sigma, \epsilon)$.

Let $\hat{\beta}$ denote the belief system derived from $\hat{\sigma}$ using Bayes' rule in $\Gamma(\epsilon)$. Suppose there is an information set K of player i such that under $\hat{\sigma}$ player i is not best responding. Thus i would find it profitable to deviate to some strategy σ_i' in K. Let σ' denote the strategy profile under which i unilaterally deviates to σ_i' in K. Clearly i's expected payoff under σ' in $\Gamma(\epsilon)$ is strictly higher than that under $\hat{\sigma}$ because K is reached with strictly positive probability, implying $w(\sigma', \epsilon) > w(\hat{\sigma}, \epsilon)$, contradicting the choice of $\hat{\sigma}$. We have thus established that $(\hat{\sigma}, \hat{\beta})$ is a perfect Bayesian equilibrium of $\Gamma(\epsilon)$.

Construct pure strategy profile $\hat{s}$ such that for each player i, $\hat{s}$ prescribes a pure strategy that is in the support of the (possibly mixed) strategy taken by i under $\hat{\sigma}$. That $\hat{\sigma}$ being a perfect Bayesian equilibrium of $\Gamma(\epsilon)$ implies $w(\hat{s}, \epsilon) = w(\hat{\sigma}, \epsilon)$. Thus $w(\hat{s}, \epsilon) = \max_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$, implying $\max_{s \in S(\Gamma)} w(s, \epsilon) = \max_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$.

Note that, for any fixed σ, $w(\sigma, \epsilon)$ is a polynomial function of ϵ of finite degrees. Since $S(\Gamma)$ is finite, for some $\eta > 0$ there is $s^* \in S(\Gamma)$ such that $w(s^*, \epsilon) = \max_{s \in S(\Gamma)} w(s, \epsilon)$ for any $\epsilon < \eta$. It follows that $s^* \in \text{argmax}_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$ if $\epsilon < \eta$. Let $\beta^*(\epsilon)$ be the belief system derived from s^* using Bayes' rule in $\Gamma(\epsilon)$. $s^* \in \text{argmax}_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$ implies $(s^*, \beta^*(\epsilon))$ is a perfect Bayesian equilibrium of $\Gamma(\epsilon)$ by the argument in the second paragraph of the proof. Note that $\beta^*(\epsilon)$ is continuous in ϵ and thus $\beta^* = \lim_{\epsilon \to 0} \beta^*(\epsilon)$ exists. Clearly (s^*, β^*) is a perfect Bayesian equilibrium of the unperturbed game. Suppose there is $\tilde{\sigma} \in \Sigma(\Gamma)$ such that $v(\tilde{\sigma}|\Gamma, \phi, F) > v(s^*|\Gamma, \phi, F)$. Since $\lim_{\epsilon \to 0} w(\sigma, \epsilon) = v(\sigma|\Gamma, \phi, F)$ for any σ, there is some $\tilde{\eta} > 0$ such that $w(\tilde{\sigma}, \epsilon) > w(s^*, \epsilon)$ for any $\epsilon < \tilde{\eta}$, contradicting that $s^* \in \text{argmax}_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$ for any $\epsilon < \eta$. Hence $s^* \in \text{argmax}_{s \in \Sigma(\Gamma)} v(s|\Gamma, \phi, F)$. It follows that $V(\Gamma, \phi, F) = v(s^*|\Gamma, \phi, F) = \max_{s \in S(\Gamma)} v(s|\Gamma, \phi, F)$. $\square$

Proposition 8. *If Γ' and Γ are finite mechanisms then Γ' dominates Γ if and only if $C(\Gamma) \subset C(\Gamma')$.*

Proof: (The "if" direction.) Suppose $C(\Gamma) \subset C(\Gamma')$. Fix (ϕ, F) and choose $s \in \text{argmax}_{S(\Gamma)} v(s|\Gamma, \phi, F)$. By assumption there is $s' \in S(\Gamma')$ such that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$. It follows that

$$\begin{aligned} v(s'|\Gamma', \phi, F) &= \sum_{\mathbf{x} \in X} F(\mathbf{x}) \phi(\boldsymbol{\alpha}(\mathbf{x}|s'), \mathbf{x}) \\ &= \sum_{\mathbf{x} \in X} F(\mathbf{x}) \phi(\boldsymbol{\alpha}(\mathbf{x}|s), \mathbf{x}) = v(s|\Gamma, \phi, F). \end{aligned}$$

By Lemma 11, $V(\Gamma', \phi, F) \geq v(s'|\Gamma', \phi, F) = v(s|\Gamma, \phi, F) = V(\Gamma, \phi, F)$, implying that Γ' dominates Γ because (ϕ, F) is arbitrarily chosen.

(The "only if" direction.) Suppose Γ' dominates Γ. Choose any $s \in S(\Gamma)$. Construct ϕ such that $\phi(\mathbf{a}, \mathbf{x}) = 1$ if $\mathbf{a} = \boldsymbol{\alpha}(\mathbf{x}|s)$ and $\phi(\mathbf{a}, \mathbf{x}) = 0$ otherwise. Choose F that is strictly positive on X. Clearly $v(s|\Gamma, \phi, F) = 1$. By Lemma

11, that Γ' dominates Γ implies

$$\max_{s \in S(\Gamma')} v(s|\Gamma', \phi, F) \geq \max_{s \in S(\Gamma)} v(s|\Gamma, \phi, F) = 1,$$

which in turn implies there is $s' \in S(\Gamma')$ such that $v(s'|\Gamma', \phi, F) \geq 1$. If $\boldsymbol{\alpha}(\mathbf{y}|s') \neq \boldsymbol{\alpha}(\mathbf{y}|s)$ for some $\mathbf{y} \in X$ then

$$v(s'|\Gamma', \phi, F) = \sum_{\mathbf{x} \in X} F(\mathbf{x})\phi(\boldsymbol{\alpha}(\mathbf{x}|s'), \mathbf{x}) \leq 1 - F(\mathbf{y}) < 1,$$

a contradiction. Thus $\boldsymbol{\alpha}(\mathbf{x}|s') = \boldsymbol{\alpha}(\mathbf{x}|s)$ for any $\mathbf{x} \in X$, implying $C(\Gamma) \subset C(\Gamma')$ because s is arbitrarily chosen. □

A.5 Proof of Lemma 2

Proof. Suppose (T', P', M') is obtained from (T, P, M) by expanding. E2 implies that for any $p \in P'$ where $p \cap T \neq \emptyset$ there is $\tau(p) \in P$ such that $\tau(p) = p \cap T$. E1 implies $M(\tau(p)) \subset M'(p)$. Choose $s \in S(T, P, M)$. Construct $s' \in S(T', P', M')$ such that for any $i \in \mathcal{N}$, $x_i \in X_i$ and $p \in P'_i$ where $p \cap T \neq \emptyset$,

$$a_i(x_i, p|s') = a_i(x_i, \tau(p)|s),$$
$$\text{(if } i < N) \quad m_i(x_i, p|s') = m_i(x_i, \tau(p)|s).$$

To verify that s' is properly defined, observe that for any $p \in P'_i$ where $p \cap T \neq \emptyset$ we have $m_i(x_i, p|s') = m_i(x_i, \tau(p)|s) \in M(\tau(p)) \subset M'(p)$. Choose any $\mathbf{x} = (x_1, ..., x_N) \in X$. Clearly $\rho_1(\mathbf{x}|s') = \rho_1(\mathbf{x}|s)$. Suppose $\rho_i(\mathbf{x}|s') = \rho_i(\mathbf{x}|s)$ for any $i \leq k$ for some $k \geq 1$. Thus

$$\begin{aligned}\mu_k(\mathbf{x}|s') &= m_k(x_k, P'(\rho_k(\mathbf{x}|s'))|s') = m_k(x_k, \tau(P'(\rho_k(\mathbf{x}|s')))|s) \\ &= m_k(x_k, P(\rho_k(\mathbf{x}|s'))|s) = m_k(x_k, P(\rho_k(\mathbf{x}|s))|s) = \mu_k(\mathbf{x}|s).\end{aligned}$$

It follows that

$$\rho_{k+1}(\mathbf{x}|s') = \rho_k(\mathbf{x}|s') \times \mu_k(\mathbf{x}|s') = \rho_k(\mathbf{x}|s) \times \mu_k(\mathbf{x}|s) = \rho_{k+1}(\mathbf{x}|s).$$

Thus $\rho_i(\mathbf{x}|s') = \rho_i(\mathbf{x}|s)$ for any $i \in \mathcal{N}$, implying

$$\begin{aligned}\alpha_i(\mathbf{x}|s') &= a_i(x_i, P'(\rho_i(\mathbf{x}|s'))|s') = a_i(x_i, \tau(P'(\rho_i(\mathbf{x}|s')))|s) \\ &= a_i(x_i, P(\rho_i(\mathbf{x}|s'))|s) = a_i(x_i, P(\rho_i(\mathbf{x}|s))|s) = \alpha_i(\mathbf{x}|s).\end{aligned}$$

Thus $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T, P, M) \subset C(T', P', M')$ since s is arbitrarily chosen. The lemma then follows immediately from Proposition 4. □

A.6 Proof of Lemma 4

Two lemmas are used to assist the proof.

Lemma 12. *Fix* (T, P, M)*. For any* $i < N$ *and* $h \in T_i$*, if* $m_i, m'_i \in M(h)$ *are synonymous given* h *within* (T, P, M) *then* $P(\rho_j(\mathbf{x}|s, i, h, m_i)) = P(\rho_j(\mathbf{x}|s, i, h, m'_i))$ *for any* $j > i$*,* $\mathbf{x} \in X$ *and* $s \in S(T, P, M)$*.*

Proof. Choose any $\mathbf{x} = (x_1, ..., x_N) \in X$, $s \in S(T, P, M)$, $i < N$, $h \in T_i$ and $m_i, m'_i \in M(h)$ such that m_i and m'_i are synonymous given h within (T, P, M). For any $j > i$ denote $g^j = \rho_j(\mathbf{x}|s, i, h, m_i)$ and $f^j = \rho_j(\mathbf{x}|s, i, h, m'_i)$.

Clearly $g^{i+1} = h \times m_i$ and $f^{i+1} = h \times m'_i$. Thus g^{i+1} and f^{i+1} are h-cousins where $g^{i+1}_i = m_i$ and $f^{i+1}_i = m'_i$. It follows by synonymity that $P(g^{i+1}) = P(f^{i+1})$. Suppose g^j and f^j are h-cousins, where $g^j_i = m_i$ and $f^j_i = m'_i$, for any $j \leq k$ for some $k \geq i + 1$.

$P(g^k) = P(f^k)$ by synonymity. Note that $g^{k+1} = g^k \times m_k(x_k, P(g^k)|s)$ and $f^{k+1} = f^k \times m_k(x_k, P(f^k)|s)$. $P(g^k) = P(f^k)$ implies the last components of g^{k+1} and f^{k+1} are identical, and hence g^{k+1} and f^{k+1} are h-cousins where $g^{k+1}_i = m_i$ and $f^{k+1}_i = m'_i$ by the inductive hypothesis. Hence $P(g^{k+1}) = P(f^{k+1})$ by synonymity. The present lemma follows by the principle of induction. □

Lemma 13. *If* (T', P', M') *is obtained from* (T, P, M) *by trimming such that there exist* $i < N$*,* $p \in P_i$ *and* $m_i, m'_i \in M(p)$ *satisfying T1 and T2, then for any* $j < N$ *and* $p' \in P'_j$*:*

1. *There is* $\zeta(p') \in P_j$ *such that* $p' \subset \zeta(p')$*.*
2. *If* $\zeta(p') \neq p$ *then* $M'(p') = M(\zeta(p'))$*.*
3. *If* $\zeta(p') = p$ *then* $M'(p') = M(p) \backslash \{m_i\}$*.*

Moreover, $|P'_j| = |P_j|$ *for any* $j \in \mathscr{N}$*.*

Proof. Part 1 follows from T1 immediately.

To show Parts 2 and 3, choose any $j < N$ and $p' \in P'_j$. Note that $M'(p') \subset M(\zeta(p'))$ by T1. If $\zeta(p') \neq p$ then for any $h \in p'$ and $\hat{m}_j \in M(h)$ we have $h \times \hat{m}_j \in T'$ by T1, implying Part 2. If $\zeta(p') = p$ then $j = i$. Note that for any $h \in p'$ and $\hat{m}_i \in M(h)$ we have $h \times \hat{m}_i \in T'$ if and only if $\hat{m}_i \neq m_i$ by T1, implying Part 3.

Pick any $j \in \mathscr{N}$. If $j \leq i$ then $T_j = T'_j$ and thus $P'_j = P_j$ by T1. If $j > i$ then $|P'_j| \leq |P_j|$ by T1. If $|P'_j| < |P_j|$ then there is some $\hat{p} \in P_j$ such that $h \notin T'$ for any $h \in \hat{p}$. Choose any $h \in \hat{p}$. It follows by T1 that $h(i-1) \in p$ and $h_i = m_i$. It then follows by T2 that there exists some $h' \in \hat{p}$ such that $h'(i-1) \in p$ and $h'_i = m'_i$. Hence $h' \in T'$ by T1, a contradiction. Therefore $|P'_j| = |P_j|$ for any $j \in \mathscr{N}$. □

Proof of Lemma 4. Suppose (T', P', M') is obtained from (T, P, M) by trimming such that there exist $i < N$, $p \in P_i$ and $m_i, m_i' \in M(p)$ satisfying T1 and T2. Since (T', P', M') is a sub-institution of (T, P, M), Lemma 2 implies (T, P, M) dominates (T', P', M').

Now show that (T', P', M') dominates (T, P, M). By assumption there are $i < N$, $h \in T_i$, $p \in P_i$ and $m_i, m_i' \in M(p)$ such that m_i and m_i' are synonymous given p within (T, P, M). Choose any $s \in S(T, P, M)$. Construct $\hat{s} \in S(T, P, M)$ that agrees with s except in the following: $m_i(x_i, p|\hat{s}) = m_i'$ for any $x_i \in X_i$ such that $m_i(x_i, p|s) = m_i$.

Choose any $\mathbf{x} = (x_1, ..., x_N) \in X$. If $\rho_i(\mathbf{x}|s) \notin p$ or $m_i(x_i, p|s) \neq m_i$ then clearly $\rho_j(\mathbf{x}|\hat{s}) = \rho_j(\mathbf{x}|s)$ for any $j \in \mathscr{N}$. Suppose $\rho_i(\mathbf{x}|s) \in p$ and $m_i(x_i, p|s) = m_i$. Obviously $\rho_j(\mathbf{x}|\hat{s}) = \rho_j(\mathbf{x}|s)$ for any $j \leq i$. Denote $h = \rho_i(\mathbf{x}|s)$. For any $j > i$ we have

$$\begin{aligned} P(\rho_j(\mathbf{x}|\hat{s})) &= P(\rho_j(\mathbf{x}|\hat{s}, i, h, m_i')) \\ &= P(\rho_j(\mathbf{x}|\hat{s}, i, h, m_i)) \\ &= P(\rho_j(\mathbf{x}|s, i, h, m_i)) \\ &= P(\rho_j(\mathbf{x}|s)) \end{aligned}$$

where the first line is due to $h = \rho_i(\mathbf{x}|\hat{s})$ and $m_i(x_i, p|\hat{s}) = m_i'$, the second line is due to Lemma 12 because m_i and m_i' are synonymous given h within (T, P, M), the third line is due to the fact that $\hat{s}$ and s agree for every player after i, and the fourth line is due to $\rho_i(\mathbf{x}|s) = h$ and $m_i(x_i, p|s) = m_i$. By induction $P(\rho_j(\mathbf{x}|\hat{s})) = P(\rho_j(\mathbf{x}|s))$ for any $j \in \mathscr{N}$. Therefore

$$\alpha_j(\mathbf{x}|\hat{s}) = a_j(x_j, P(\rho_j(\mathbf{x}|\hat{s}))|\hat{s}) = a_j(x_j, P(\rho_j(\mathbf{x}|s))|\hat{s}) = a_j(x_j, P(\rho_j(\mathbf{x}|s))|s)$$

for any $j \in \mathscr{N}$. Hence $\boldsymbol{\alpha}(\cdot|\hat{s}) = \boldsymbol{\alpha}(\cdot|s)$.

By Part 1 of Lemma 13, for each $j \in \mathscr{N}$ and $p' \in P_j'$ there is $\zeta(p') \in P_j$ such that $p' \subset \zeta(p')$. Construct $s' \in S(T', P', M')$ such that for any $j \in \mathscr{N}$, $x_j \in X_j$ and $p' \in P_j'$,

$$\begin{aligned} & a_j(x_j, p'|s') = a_j(x_j, \zeta(p')|\hat{s}), \\ \text{(if } j < N\text{)} \quad & m_j(x_j, p'|s') = m_j(x_j, \zeta(p')|\hat{s}). \end{aligned}$$

s' is properly defined if and only if $m_j(x_j, p'|s') \in M'(p')$ for any $j < N$, $x_j \in X_j$ and $p' \in P_j'$. To verify it, observe that if $\zeta(p') \neq p$ then $m_j(x_j, p'|s') = m_j(x_j, \zeta(p')|\hat{s}) \in M(\zeta(p')) = M'(p')$ by Part 2 of Lemma 13; whereas if $\zeta(p') = p$ then we have $j = i$ and moreover $m_i(x_i, p'|s') = m_i(x_i, p|\hat{s}) \in M(p) \backslash \{m_i\} = M'(p')$ by Part 3 of Lemma 13 because by construction $m_i(x_i, p|\hat{s}) \neq m_i$.

Choose any $\mathbf{x} = (x_1, ..., x_N) \in X$. Clearly $\rho_1(\mathbf{x}|s') = \rho_1(\mathbf{x}|\hat{s})$. Suppose

$\rho_j(\mathbf{x}|s') = \rho_j(\mathbf{x}|\hat{s})$ for any $j \leq k$ for some $k \geq 1$. Thus

$$\begin{aligned}\mu_k(\mathbf{x}|s') &= m_k(x_k, P'(\rho_k(\mathbf{x}|s'))|s') = m_k(x_k, \zeta(P'(\rho_k(\mathbf{x}|s')))|\hat{s}) \\ &= m_k(x_k, P(\rho_k(\mathbf{x}|s'))|\hat{s}) = m_k(x_k, P(\rho_k(\mathbf{x}|\hat{s}))|\hat{s}) = \mu_k(\mathbf{x}|\hat{s}).\end{aligned}$$

It follows that

$$\rho_{k+1}(\mathbf{x}|s') = \rho_k(\mathbf{x}|s') \times \mu_k(\mathbf{x}|s') = \rho_k(\mathbf{x}|\hat{s}) \times \mu_k(\mathbf{x}|\hat{s}) = \rho_{k+1}(\mathbf{x}|\hat{s}).$$

Thus $\rho_j(\mathbf{x}|s') = \rho_i(\mathbf{x}|\hat{s})$ for any $j \in \mathscr{N}$, implying

$$\begin{aligned}\alpha_j(\mathbf{x}|s') &= a_j(x_i, P'(\rho_j(\mathbf{x}|s'))|s') = a_j(x_j, \zeta(P'(\rho_j(\mathbf{x}|s')))|\hat{s}) \\ &= a_j(x_j, P(\rho_j(\mathbf{x}|s'))|\hat{s}) = a_j(x_j, P(\rho_j(\mathbf{x}|\hat{s}))|\hat{s}) = \alpha_j(\mathbf{x}|\hat{s}).\end{aligned}$$

It follows that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|\hat{s})$, implying $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$. Thus $C(T, P, M) \subset C(T', P', M')$ because s is arbitrarily chosen from $S(T, P, M)$. The present lemma then follows from Proposition 4. □

A.7 Proof of Lemma 5

Proof. Suppose (T', P', M') is obtained from (T, P, M) by relabeling with relabeling function γ satisfying R1-R4. R3 implies that for any $p \in P'$ there is $\tau(p) \in P$ such that $h \in \tau(p)$ if and only if $\gamma(h) \in p$. Let γ^{-1} be the inverse mapping of γ. Clearly γ^{-1} satisfies R1-R3.

Now show that γ^{-1} satisfies R4. Define $\kappa'(m_i, p)$ where $p \in P'_i$ and $m_i \in M'(p)$ such that $\kappa'(m_i, p) \in M(\tau(p))$ and $\kappa(\kappa'(m_i, p), \tau(p)) = m_i$. First verify that κ' is defined for every $p \in P'_i$ and $m_i \in M'(p)$. Choose any $p \in P'_i$, $m_i \in M'(p)$ and $h \in \tau(p)$. That γ^{-1} satisfies R1 and R3 implies $\gamma^{-1}(\gamma(h) \times m_i) = h \times \hat{m}_i$ for some $\hat{m}_i \in M(\tau(p))$. By definition of κ we have $\kappa(\hat{m}_i, \tau(p)) = m_i$, confirming that indeed $\kappa'(m_i, p)$ exists.

Choose any $p \in P'_i$, $m_i \in M'(p)$ and $h \in p$. $\gamma^{-1}(h) \in \tau(p)$ by R3. By construction $h \times m_i = h \times \kappa(\kappa'(m_i, p), \tau(p))$. Also note that

$$\begin{aligned}\gamma\Big(\gamma^{-1}(h) \times \kappa'(m_i, p)\Big) &= \gamma(\gamma^{-1}(h)) \times \kappa(\kappa'(m_i, p), \tau(p)) \\ &= h \times \kappa(\kappa'(m_i, p), \tau(p)) = h \times m_i.\end{aligned}$$

Applying γ^{-1} to both sides we have $\gamma^{-1}(h) \times \kappa'(m_i, p) = \gamma^{-1}(h \times m_i)$. Thus γ^{-1} satisfies R4. It follows that (T, P, M) is obtained from (T', P', M') by relabeling with relabeling function γ^{-1}. □

A.8 Proof of Lemma 6

Proof. Suppose (T', P', M') is obtained from (T, P, M) by relabeling with relabeling function γ satisfying R1-R4. R3 implies that for any $p \in P'$ there is $\tau(p) \in P$ such that $h \in \tau(p)$ if and only if $\gamma(h) \in p$.

Choose any $s \in S(T, P, M)$. Construct $s' \in S(T', P', M')$ such that for any $i \in \mathscr{N}$, $x_i \in X_i$ and $p \in P_i'$,

$$\begin{aligned} a_i(x_i, p|s') &= a_i(x_i, \tau(p)|s), \\ \text{(if } i < N\text{)} \quad m_i(x_i, p|s') &= \kappa(m_i(x_i, \tau(p)|s), \tau(p)), \end{aligned}$$

where κ is defined in R4. To verify that s' is properly defined, choose any $h \in \tau(p)$ where $p \in P_i'$ for some $i < N$. Thus $\gamma(h) \in p$. R4 and R1 implies

$$\gamma(h \times m_i(x_i, \tau(p)|s)) = \gamma(h) \times \kappa(m_i(x_i, \tau(p)|s), \tau(p)) = \gamma(h) \times m_i(x_i, p|s') \in T'.$$

Since $\gamma(h) \in p$, it follows that $m_i(x_i, p|s') \in M'(\gamma(h))$, implying $m_i(x_i, p|s') \in M'(p)$ because $\gamma(h) \in p$.

Choose any $\mathbf{x} = (x_1, ..., x_N) \in X$. Clearly $\rho_1(\mathbf{x}|s') = \gamma(\rho_1(\mathbf{x}|s))$. Suppose $\rho_i(\mathbf{x}|s') = \gamma(\rho_i(\mathbf{x}|s))$ for any $i \leq k$ for some $k \geq 1$. Denote $h = \rho_k(\mathbf{x}|s)$, $h' = \rho_k(\mathbf{x}|s')$, $p = P(h)$ and $p' = P'(h')$. By the inductive hypothesis $h' = \gamma(h)$ and thus $p = \tau(p')$. We have

$$\begin{aligned} \rho_{k+1}(\mathbf{x}|s') &= h' \times m_k(x_k, p'|s') = \gamma(h) \times \kappa(m_k(x_k, \tau(p')|s), \tau(p')) \\ &= \gamma(h) \times \kappa(m_k(x_k, p|s), p) = \gamma(h \times m_k(x_k, p|s)) = \gamma(\rho_{k+1}(\mathbf{x}|s)). \end{aligned}$$

Thus $\rho_i(\mathbf{x}|s') = \gamma(\rho_i(\mathbf{x}|s))$ for any $i \in \mathscr{N}$, implying $P(\rho_i(\mathbf{x}|s)) = \tau(P'(\rho_i(\mathbf{x}|s')))$. Therefore

$$\begin{aligned} \alpha_i(\mathbf{x}|s') &= a_i(x_i, P'(\rho_i(\mathbf{x}|s'))|s') = a_i(x_i, \tau(P'(\rho_i(\mathbf{x}|s')))|s) \\ &= a_i(x_i, P(\rho_i(\mathbf{x}|s))|s) = \alpha_i(\mathbf{x}|s). \end{aligned}$$

Thus $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T, P, M) \subset C(T', P', M')$ since s is arbitrarily chosen from $S(T, P, M)$.

Lemma 5 implies (T, P, M) is obtained from (T', P', M') by relabeling. Thus by an analogous argument as above we have $C(T', P', M') \subset C(T, P, M)$. The present lemma then follows by Proposition 4.

□

A.9 The Formal Definition of Merging

To introduce merging formally, it is useful to first define a special kind of expanding. Fix institution (T, P, M). Consider the following construction:

Pick any $i < N$, $p \in P_i$ and $m_i \in M(p)$. Let T^{m_i} denote the set of all $h \in T$ such that $h(i-1) \in p$ and $h_i = m_i$. Let $T^{m_i'}$ denote the set of all message profiles derived from changing the ith component of some $h \in T^{m_i}$ from m_i to $m_i' \notin M(p)$. Thus for each $h \in T^{m_i'}$ there is $f^h \in T^{m_i}$ such that h and f^h differ only at the ith component. Observe that T and $T^{m_i'}$ are disjoint, for if $h \in T \cap T^{m_i'}$ then $h(i-1) \in p$ and $h_i = m_i'$, implying $m_i' \in M(p)$, a contradiction. Let $T' = T \cup T^{m_i'}$. Let P' be a partition of T' such that

- D1 $P'(h) = P'(g)$ if and only if $P(h) = P(g)$ for any $h, g \in T$.
- D2 $P'(h) = P'(f^h)$ for any $h \in T^{m_i'}$.

It is plain to see that T' equipped with partition P' is derived from T equipped with P by putting each additional $h \in T^{m_i'}$ in the same partition cell as f^h. Define M' such that

- D3 $M'(h) = M(h)$ if $h \in T \backslash p$.
- D4 $M'(h) = M(h) \cup \{m_i'\}$ if $h \in p$.
- D5 $M'(h) = M(f^h)$ if $h \in T^{m_i'}$.

Observe that (T', P', M') is an improper institution. To verify it, the only non-obvious part is to show that $M'(h) = M'(g)$ for any $h, g \in T'$ such that $P'(h) = P'(g)$. If $h, g \in T$ then $P'(h) = P'(g)$ implies $P(h) = P(g)$ and consequently $M'(h) = M(h) = M(g) = M'(g)$ if $h, g \notin p$, or $M'(h) = M(h) \cup \{m_i'\} = M(g) \cup \{m_i'\} = M'(g)$ if $h, g \in p$. If $h \in T$ and $g \in T^{m_i'}$ then $P'(h) = P'(g)$ implies by D2 that $P'(h) = P'(f^g)$, and it follows from $f^g \in T$ and D5 that $M'(h) = M'(f^g) = M'(g)$. If $h \in T^{m_i'}$ and $g \in T^{m_i'}$ then $P'(h) = P'(g)$ implies $P'(f^h) = P'(f^g)$ and it follows by D5 that $M'(h) = M'(f^h) = M'(f^g) = M'(g)$. Thus (T', P', M') is indeed an improper institution. We say that (T', P', M') is obtained from (T, P, M) by $(i, p, m_i, m_i')-$**duplication**.

The following lemma suggests that $(i, p, m_i, m_i')-$**duplication** does no more than expanding (T, P, M) by duplicating message m_i for p and giving the cloned message the label m_i'.

Lemma 14. *If* (T', P', M') *is obtained from* (T, P, M) *by* $(i, p, m_i, m_i')-$*duplication then:*

1. $T_j' = T_j$ *and* $P_j' = P_j$ *for any* $j \leq i$.
2. m_i *and* m_i' *are synonymous given* p *within* (T', P', M').
3. (T, P, M) *can be obtained from* (T', P', M') *by trimming.*
4. (T, P, M) *and* (T', P', M') *dominate each other.*

Proof. Inherit the notation $T^{m_i'}$ and f^h from the introduction of duplication.

$T_j = T_j'$ for any $j \leq i$ because $T^{m_i'}$ only contains message profiles of length at least i. Thus D1 implies $P_j' = P_j$, establishing Part 1.

Now show Part 2. Choose any $h \in p$. D4 implies $m_i, m_i' \in M'(h)$. Part 1 implies $P'(h) = p$. Thus $m_i, m_i' \in M'(p)$. Let $g, g' \in T'$ be $h-$cousins such that $g_i = m_i$ and $g_i' = m_i'$. Clearly $g = f^{g'}$. It follows from D2 that $P'(g) = P'(g')$, implying m_i and m_i' are synonymous given h within (T', P', M'). Part 2 follows immediately because h is arbitrarily chosen from p.

Now show Part 3. Clearly (T, P, M) is a sub-institution of (T', P', M'). If $h \in T' \backslash T$ then $h \in T^{m_i'}$ and it follows by the construction of $T^{m_i'}$ that $h(i-1) \in p$ and $h_i = m_i'$, establishing T1. T2 follows from Part 2.

Part 4 follows from Part 3 by Lemma 4. □

Now we introduce merging. The definition is given in terms of the construction. Each step is explained by the accompanying remark.

Definition. (T', P', M') is obtained from (T, P, M) by **merging** if there is $i < N$ where $|A_i| = 1$, $p \in P_i$ and $q \in P_i$ (without loss of generality assume $|M(p)| < |M(q)|$) such that (T', P, M') is the product of the following algorithm:

Step 1 Fix $m_i \in M(p)$. Set $(T^0, P^0, M^0) = (T, P, M)$. Produce (T^k, P^k, M^k) from $(T^{k-1}, P^{k-1}, M^{k-1})$ by $(i, p, m_i, m_i^k)-$duplication [7] where m_i^k is an arbitrary message not in $M^{k-1}(p)$. Stop if $|M^k(p)| = |M^k(q)| = |M(q)|$. [8] Denote the terminal product of this step as $(\tilde{T}, \tilde{P}, \tilde{M})$.

Remark: This step equalizes the number of messages available to player i given p and that given q.

Step 2 Choose any bijections $\lambda : \tilde{M}(p) \to \tilde{M}(q)$. Construct mapping γ on $\hat{T}$ such that

(a) If g is not a descendant of some $h \in p$ then $\gamma(h) = h$.

(b) If g is a descendant of some $h \in p$ then $\gamma(g) = e^g$ where e^g is derived from g by replacing the ith component of g from g_i to $\lambda(g_i)$.

Denote the end product of this step as $(\hat{T}, \hat{P}, \hat{M})$.

Remark: It is easy to verify that γ is a relabeling function satisfying R1-R4. This step relabels the messages available to player i given p, so that now $\hat{M}(p) = \hat{M}(q)$.

Step 3 Arbitrarily index messages in $\hat{M}(p)$ as $m^1, ..., m^K$ where $K = |\hat{M}(p)|$. For each $j \neq k$, abritrarily choose a unique message $n^{j,k}$ where $n^{j,k} \notin \hat{M}(p)$. Let $n^{j,j} = m^j$.

Set $(\hat{T}^{1,1}, \hat{P}^{1,1}, \hat{M}^{1,1}) = (\hat{T}, \hat{P}, \hat{M})$. Iterate through $k = 1 : K$ as follows: For each j and $k \neq j-1$ produce $(T^{j,k+1}, P^{j,k+1}, M^{j,k+1})$ from $(T^{j,k}, P^{j,k}, M^{j,k})$ by the following two substeps:

[7] Part 1 of Lemma 14 implies $p \in P^{k-1}$.
[8] Part 1 of Lemma 14 implies $q \in P^{k-1}$. D3 implies $M^k(q) = M(q)$.

(a) Produce $(T_p^{j,k+1}, P_p^{j,k+1}, M_p^{j,k+1})$ from $(T^{j,k}, P^{j,k}, M^{j,k})$ by $(i, p, m^j, n^{j,k})$-duplication.

(b) Produce $(T^{j,k+1}, P^{j,k+1}, M^{j,k+1})$ from $(T_p^{j,k+1}, P_p^{j,k+1}, M_p^{j,k+1})$ by $(i, q, m^k, n^{j,k+1})$-duplication.

In the case that $k = j-1$, set $(T^{j,j}, P^{j,j}, M^{j,j}) = (T^{j,j-1}, P^{j,j-1}, M^{j,j-1})$.

When $(T^{j,K}, P^{j,K}, M^{j,K})$ is reached, produce $(T^{j+1,1}, P^{j+1,1}, M^{j+1,1})$ from $(T^{j,K}, P^{j,K}, M^{j,K})$ by the following two substeps:

(a) Produce $(T_p^{j+1,1}, P_p^{j+1,1}, M_p^{j+1,1})$ from $(T^{j,K}, P^{j,K}, M^{j,K})$ by $(i, p, m^{j+1}, n^{j+1,1})$-duplication.

(b) Produce $(T^{j+1,1}, P^{j+1,1}, M^{j+1,1})$ from $(T_p^{j+1,1}, P_p^{j+1,1}, M_p^{j+1,1})$ by $(i, q, m^1, n^{j+1,1})$-duplication.

Stop when $(T^{K,K}, P^{K,K}, M^{K,K})$ is reached. Denote $(T^{K,K}, P^{K,K}, M^{K,K})$ as (T^*, P^*, M^*).

Remark: It is straightforward to verify that $M^*(p) = M^*(q) = \{n^{j,k} : (j,k) \in \{1, ..., K\}^2\}$. This step expands $(\hat{T}, \hat{P}, \hat{M})$ in a particular way, so that for each pair j, k where $j \neq k$, the message $n^{j,k}$ is introduced to p and q, where it is synonymous to m^j given p, and to m^k given q.

Step 4 Produce (T', P', M') from (T^*, P^*, M^*) such that $T' = T^*$ and P^* satisfies

(a) $P'(h) = P'(g)$ if and only if $P^*(h) = P^*(g)$ for any $h, g \notin p \cup q$.

(b) $P'(h) = P'(g)$ for any $h, g \in p \cup q$.

Remark: This step combines separate perceptions p and q into one single perception $p \cup q$.

A.10 Proof of Lemma 7

Proof. Inherit the notation introduced the formal definition of merging in Section A.9. (T^*, P^*, M^*) is obtained from (T, P, M) by a sequence of duplication and relabeling operations. Thus (T^*, P^*, M^*) and (T, P, M) dominate each other by Lemmas 6 and 14. Clearly (T^*, P^*, M^*) can be obtained from (T', P', M') by refining. Thus (T^*, M^*, P^*) dominates (T', P', M') by Lemma 3 and hence (T, P, M) dominates (T', P', M').

To establish that (T', P', M') dominates (T, P, M) it suffices to show that (T', P', M') dominates (T^*, M^*, P^*). It is straightforward to verify that $M^*(p) = M^*(q) = \{n^{j,k} : (j,k) \in \{1, ..., K\}^2\}$. Recall that $n^{j,j} = m^j$. Choose any $n^{j,k}$ where $j \neq k$. Since $(T_p^{j,k}, P_p^{j,k}, M_p^{j,k})$ is obtained from $(T^{j,k-1}, P^{j,k-1}, M^{j,k-1})$ (or $(T^{j-1,K}, P^{j-1,K}, M^{j-1,K})$ in the case $k = 1$) by $(i, p, m^j, n^{j,k})$-duplication, it follows by Lemma 14 that m^j and $n^{j,k}$ are synonymous given p within

$(T_p^{j,k}, P_p^{j,k}, M_p^{j,k})$. Choose any $h \in p$ and $h-$cousins $g, g' \in T^*$ such that $g_i = m^j$ and $g'_i = n^{j,k}$. Clearly $g, g' \in (T^{j,k}, P^{j,k}, M^{j,k})$. It follows that $P^{j,k}(g) = P^{j,k}(g')$ by synonymity. Since $(T_p^{j,k}, P_p^{j,k}, M_p^{j,k})$ is a sub-institution of (T^*, M^*, P^*), we have $P^*(g) = P^*(g')$. Moreover since for any $l > i$ we have $P'_l = P^*_l$, it follows that $P'(g) = P'(g')$. Therefore m^j and $n^{j,k}$ are synonymous given any $h \in p$ within (T^*, P^*, M^*) and within (T', P', M'). Similarly m^k and $n^{j,k}$ are synonymous given any $h \in q$ within (T^*, P^*, M^*) and within (T', P', M').

Choose any $s \in S(T^*, P^*, M^*)$. Construct $\hat{s}$ such that $\hat{s}$ agrees with s except in the following:

$$m_i(x_i, p|\hat{s}) = m^j \text{ for any } x_i \text{ such that } m_i(x_i, p|s) = n^{j,k},$$

$$m_i(x_i, q|\hat{s}) = m^k \text{ for any } x_i \text{ such that } m_i(x_i, q|s) = n^{j,k}.$$

Clearly $\rho_l(\mathbf{x}|\hat{s}) = \rho_l(\mathbf{x}|s)$ for any $l \in \mathscr{N}$ if $\rho_i(\mathbf{x}|s) \notin p \cup q$. Fix $\mathbf{x} = (x_1, ..., x_N) \in X$ such that $\rho_i(\mathbf{x}|s) \in p$. $P^*(\rho_l(\mathbf{x}|\hat{s})) = P^*(\rho_l(\mathbf{x}|s))$ for any $l \leq i$ because $\rho_l(\mathbf{x}|\hat{s}) = \rho_l(\mathbf{x}|s))$. Denote $h = \rho_i(\mathbf{x}|\hat{s})$. Note that $m_i(x_i, p|\hat{s}) = m^j$ and $m_i(x_i, p|s) = n^{j,k}$ for some $(j,k) \in \{1, ..., K\}^2$. For any $l > i$ observe that

$$\begin{aligned}
P^*(\rho_l(\mathbf{x}|\hat{s})) &= P^*(\rho_l(\mathbf{x}|\hat{s}, i, h, m^j)) \\
&= P^*(\rho_l(\mathbf{x}|\hat{s}, i, h, n^{j,k})) \\
&= P^*(\rho_l(\mathbf{x}|s, i, h, n^{j,k})) \\
&= P^*(\rho_l(\mathbf{x}|s)).
\end{aligned}$$

The first line is due to $h = \rho_i(\mathbf{x}|\hat{s})$ and $m_i(x_i, p|\hat{s}) = m^j$. The second line is due to Lemma 12 because m^j and $n^{j,k}$ are synonymous given h within (T^*, P^*, M^*) if $j \neq k$, or $m^j = n^{j,k}$ if $j = k$. The third line is due to the fact that $\hat{s}$ and s agree for every player after i, and the fourth line is due to $\rho_i(\mathbf{x}|s) = h$ and $m_i(x_i, p|s) = n^{j,k}$. Similarly $P^*(\rho_l(\mathbf{x}|\hat{s})) = P^*(\rho_l(\mathbf{x}|s))$ for any $\mathbf{x}$ such that $\rho_i(\mathbf{x}|s) \in q$. By induction $P^*(\rho_l(\mathbf{x}|\hat{s})) = P^*(\rho_l(\mathbf{x}|s))$ for any $l \in \mathscr{N}$ and $\mathbf{x} \in X$.

Note that $P' = (P^* \backslash \{p, q\}) \cup \{p \cup q\}$. Choose $s' \in S(T', P', M')$ such that

- For any $l \neq i$, $x_l \in X_l$ and $r \in P'_l$:

$$a_l(x_l, r|s') = a_l(x_l, r|\hat{s}).$$

- For any $l < N$, $x_l \in X_l$ and $r \in P'_l$ where $r \neq p \cup q$:

$$m_l(x_l, r|s') = m_l(x_l, r|\hat{s}).$$

- For any $x_i \in X_i$ such that $m_i(x_i, p|\hat{s}) = m^j$ and $m_i(x_i, p|\hat{s}) = m^k$:

$$m_i(x_i, p \cup q|s') = n^{j,k}.$$

We have $\rho_l(\mathbf{x}|s') = \rho_l(\mathbf{x}|\hat{s})$ for any $\mathbf{x}$ such that $\rho_i(\mathbf{x}|s') \notin p \cup q$. Fix $\mathbf{x} = (x_1, ..., x_N)$ such that $\rho_i(\mathbf{x}|s') \in p$. Obviously $\rho_i(\mathbf{x}|s') = \rho_i(\mathbf{x}|\hat{s})$. Denote $h = \rho_i(\mathbf{x}|s')$. By construction of $\hat{s}$ we have $m_i(x_i, p|\hat{s}) = m^j$ and $m_i(x_i, q|\hat{s}) = m^k$ for some $(j,k) \in \{1, ..., K\}^2$. Thus for any $l > i$,

$$\begin{aligned} P'(\rho_l(\mathbf{x}|s')) &= P'(\rho_l(\mathbf{x}|s', i, h, n^{j,k})) \\ &= P'(\rho_l(\mathbf{x}|s', i, h, m^j)) \\ &= P^*(\rho_l(\mathbf{x}|\hat{s}, i, h, m^j)) \\ &= P^*(\rho_l(\mathbf{x}|\hat{s})). \end{aligned}$$

Similarly for any $\mathbf{x}$ such that $\rho_i(\mathbf{x}|s') \in q$ we have $P'(\rho_l(\mathbf{x}|s')) = P^*(\rho_l(\mathbf{x}|\hat{s}))$ for any $l > i$. Thus $P'(\rho_l(\mathbf{x}|s')) = P^*(\rho_l(\mathbf{x}|\hat{s}))$ for any $l \neq i$ and $\mathbf{x} \in X$ by induction. For any $l \in \mathscr{N}$ where $|A_l| > 1$ (and hence $l \neq i$ because $|A_i| = 1$) and $\mathbf{x} = (x_1, ..., x_N) \in X$ we have

$$\begin{aligned} \alpha_l(\mathbf{x}|s') &= a_l(x_l, P'(\rho_l(\mathbf{x}|s'))|s') = a_l(x_l, P^*(\rho_l(\mathbf{x}|\hat{s}))|s') = a_l(x_l, P^*(\rho_l(\mathbf{x}|s))|s') \\ &= a_l(x_l, P^*(\rho_l(\mathbf{x}|s))|\hat{s}) = \alpha_l(x_l, P^*(\rho_l(\mathbf{x}|s))|s) = \alpha_l(\mathbf{x}|s). \end{aligned}$$

It follows that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T^*, P^*, M^*) \subset C(T', P', M')$ because s is arbitrarily chosen from $S(T^*, P^*, M^*)$. Thus (T', P', M') dominates (T^*, P^*, M^*) by Proposition 4, establishing the present lemma. □

A.11 Proof of Theorem 1

Proof. Part 1 is an immediate consequence of Lemmas 2, 3, 4, 6 and 7.

Now we prove Part 2, the "only if" direction. Suppose (T', P', M') dominates (T, P, M). The proofs is broken into several steps.

(Step 1. "Merging".) It is plain to see that an institution $(\hat{T}, \hat{P}, \hat{M})$, where $|\hat{P}_i| = 1$ for every player i whose action set is a singleton, can be obtained from (T, P, M) by a sequence of merging operations. It follows from Lemma 7 that (T, P, M) dominates $(\hat{T}, \hat{P}, \hat{M})$. Thus (T', P', M') dominates $(\hat{T}, \hat{P}, \hat{M})$.

(Step 2. "Refining".) For any $i < N$ define $t_i = \max\{\max_{p \in \hat{P}_i} \hat{M}(p), \frac{\log |\hat{P}_i|}{\log |A_i|}\}$. Suppose $|X_i| \geq t_i$. Since $|X_i| \geq \max_{p \in \hat{P}_i} \hat{M}(p)$, there is a mapping $y_i : \hat{P}_i \times \cup_{p \in \hat{P}_i} \hat{M}(p) \to X_i$ such that $y_i(p, m_i) \neq y_i(p, m_i')$ for any $m_i, m_i' \in \hat{M}(p)$ where $m_i \neq m_i'$. If $|A_i| > 1$ then for each $p \in \hat{P}_i$ there is a mapping $s_p : X_i \to A_i$ such that $s_p \neq s_{p'}$ if $p \neq p'$. To see that, first observe that the total number of mappings from X_i to A_i is $|A_i|^{|X_i|}$. By assumption $|X_i| \geq \frac{\log |\hat{P}_i|}{\log |A_i|}$ or equivalently $|A_i|^{|X_i|} \geq |\hat{P}_i|$. Thus each $p \in \hat{P}_i$ can be assigned with a unique s_p.

Note that for each $p, p' \in \hat{P}_i$ where $p \neq p'$ there is $z_i(p, p') \in X_i$ such that

$s_p(z_i(p,p')) \neq s_{p'}(z_i(p,p'))$. Choose any $s \in S(\hat{T},\hat{P},\hat{M})$ such that

- For any $i \in \mathscr{N}$ where $|A_i| > 1$, $x_i \in X_i$ and $p \in \hat{P}_i$:

$$a_i(x_i, p|s) = s_p(x_i).$$

- For any $i < N$, $p \in \hat{P}_i$ and $m_i \in \hat{M}(p)$:

$$m_i(y_i(p, m_i), p|s) = m_i.$$

By Proposition 4 there is $s' \in S(T',P',M')$ such that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$. For each $h \in \hat{T}$ construct $\mathbf{x}^h = (x_1^h, ..., x_N^h) \in X$ such that $x_i^h = y_i(\hat{P}(h(i-1)), h_i)$ for any $i \leq |h|$. It is straightforward to verify that $h = \rho_{|h|+1}(\mathbf{x}^h|s)$. Define the following objects:

- $\gamma(h) = \rho_{|h|+1}(\mathbf{x}^h|s')$ for any $h \in \hat{T}$.
- $\tau(p) = \{h \in \hat{P}_i : \gamma(h) \in p\}$ for any $p \in P_i'$.
- $\tilde{P}_i = \{\tau(p) : p \in P_i' \text{ and } \tau(p) \neq \emptyset\}$ for any $i \in \mathscr{N}$.

$\tilde{P}_i$ is a partition of $\hat{T}_i$ because: (1) for any $h \in \hat{T}_i$, $\gamma(h) \in p$ for some $p \in P_i'$, and (2) if $h \in \tau(p)$ and $h \in \tau(p')$ then $p = P'(\gamma(h)) = p'$.

Note that for any i where $|A_i| = 1$, we have $|\hat{P}_i| = 1$ by construction and hence $\tilde{P}_i$ is a refinement of $\hat{P}_i$. Fix any i where $|A_i| > 1$. Suppose there exist $h, g \in \hat{T}_i$ such that $\hat{P}(h) \neq \hat{P}(g)$ yet $\tilde{P}(h) = \tilde{P}(g)$. Denote $h' = \gamma(h)$ and $g' = \gamma(g)$. $\tilde{P}(h) = \tilde{P}(g)$ implies $P'(h') = P'(g')$. Choose $\mathbf{x} = (x_1, ..., x_N) \in X$ such that $x_j = x_j^h$ for any $j < i$ and $x_i = z_i(\hat{P}(h), \hat{P}(g))$. Choose $\hat{\mathbf{x}} = (\hat{x}_1, ..., \hat{x}_N) \in X$ such that $\hat{x}_j = x_j^g$ for any $j < i$ and $\hat{x}_i = z_i(\hat{P}(h), \hat{P}(g))$. Obviously $\rho_i(\mathbf{x}|s) = \rho_i(\mathbf{x}^h|s) = h$. Similarly $\rho_i(\hat{\mathbf{x}}|s) = g$. Therefore we have

$$\begin{aligned}\alpha_i(\mathbf{x}|s) &= a_i(z_i(\hat{P}(h), \hat{P}(g)), \hat{P}(h)|s) = s_{\hat{P}(h)}(z_i(\hat{P}(h), \hat{P}(g))) \\ &\neq s_{\hat{P}(g)}(z_i(\hat{P}(h), \hat{P}(g))) = a_i(z_i(\hat{P}(h), \hat{P}(g)), \hat{P}(g)|s) = \alpha_i(\hat{\mathbf{x}}|s).\end{aligned}$$

It follows from $h' = \rho_i(\mathbf{x}^h|s')$ that $h' = \rho_i(\mathbf{x}|s')$ because $\mathbf{x}^h$ and $\mathbf{x}$ agree for the first $i-1$ components. Similarly $g' = \rho_i(\hat{\mathbf{x}}|s')$. Thus

$$\alpha_i(\mathbf{x}|s') = a_i(z_i(\hat{P}(h), \hat{P}(g)), P'(h')|s') = a_i(z_i(\hat{P}(h), \hat{P}(g)), P'(g')|s') = \alpha_i(\hat{\mathbf{x}}|s')$$

because $P'(h') = P'(g')$. It follows that

$$\alpha_i(\mathbf{x}|s') = \alpha_i(\mathbf{x}|s) \neq \alpha_i(\hat{\mathbf{x}}|s) = \alpha_i(\hat{\mathbf{x}}|s') = \alpha_i(\mathbf{x}|s'),$$

a contradiction. Thus $\hat{P}(h) \neq \hat{P}(g)$ implies $\tilde{P}(h) \neq \tilde{P}(g)$, in turn implying $\tilde{P}_i$ is a refinement of $\hat{P}_i$. Let $\tilde{P} = \cup_{i \in \mathscr{N}} \tilde{P}_i$. It then follows that $(\hat{T}, \tilde{P}, \hat{M})$ is obtained from $(\hat{T}, \hat{P}, \hat{M})$ by refining.

(Step 3. "Trimming".) Apply a sequence of trimming operations to $(\hat{T}, \tilde{P}, \hat{M})$ to obtain (T^*, P^*, M^*) such that for any $i < N$ and $p \in P_i^*$ there do not exist $m_i, m_i' \in M^*(p)$ that are synonymous given p within (T^*, P^*, M^*).

(Step 4. "Relabeling".) Clearly (T^*, P^*, M^*) is a sub-institution of $(\hat{T}, \tilde{P}, \hat{M})$. Therefore $P^*(h) = P^*(g)$ if and only if $\tilde{P}(h) = \tilde{P}(g)$ for any $h, g \in T^*$. Let T'' denote the range of γ with domain restricted to T^*. Let P'' denote P' restricted to T'', that is, $P''(h) = P''(g)$ if and only if $P'(h) = P'(g)$ for any $h, g \in T''$. For any $i < N$ and $h \in T_i''$ define $M''(h) = \{m_i : h \times m_i \in T''\}$. We want to show that (T'', P'', M'') is obtained from (T^*, P^*, M^*) by relabeling with relabeling function γ. This will be achieved below by establishing that γ with domain restricted to T^* satisfies R1-R4.[9]

Observe that for any $h, g \in T^*$, if h is the parent of g then the first $|h|$ components of $\mathbf{x}^h$ and $\mathbf{x}^g$ are the same. It follows that $\rho_{|h|+1}(\mathbf{x}^g|s') = \rho_{|h|+1}(\mathbf{x}^h|s')$, implying $\gamma(h)$ is the parent of $\gamma(g)$. Thus γ satisfies R2.

Observe that for any $h, g \in T^*$, $P^*(h) = P^*(g)$ if and only if $\tilde{P}(h) = \tilde{P}(g)$ if and only if $P'(h) = P'(g)$ if and only if $P''(h) = P''(g)$. Thus γ satisfies R3.

To establish that γ is a bijection between T^* and T'' it is sufficient to verify that γ restricted to T^* is one-to-one. Suppose γ is not one-to-one, then there are $g, g' \in T^*$ such that $\gamma(g) = \gamma(g')$. Let i denote the largest index such that $g(i-1) = g'(i-1)$. Denote $f = g(i-1)$, $h = g(i)$, $h' = g'(i)$, $m_i = g_i$ and $m_i' = g_i'$. Thus $h = f \times m_i$ and $h' = f \times m_i'$. Since f is the parent of h and h' it follows that $\gamma(f)$ is the parent of $\gamma(h)$ and $\gamma(h')$. That $\gamma(g) = \gamma(g')$ then implies $\gamma(h) = \gamma(h')$. Denote $p = \hat{P}(f)$ and $p' = P'(\gamma(f))$. By construction of γ, $\gamma(h) = \gamma(h')$ implies $\rho_i(\mathbf{x}^h|s') = \rho_i(\mathbf{x}^{h'}|s')$ and $\mu_i(\mathbf{x}^h|s') = \mu_i(\mathbf{x}^{h'}|s')$. Since $\rho_i(\mathbf{x}^h|s') = \rho_i(\mathbf{x}^{h'}|s') = \gamma(f)$, it follows that

$$m_i(x_i^h, p'|s') = \mu_i(\mathbf{x}^h|s') = \mu_i(\mathbf{x}^{h'}|s') = m_i(x_i^{h'}, p'|s').$$

Thus $m_i(y_i(p, m_i), p'|s') = m_i(y_i(p, m_i'), p'|s')$ because by construction we have $x_i^h = y_i(p, m_i)$ and $x_i^{h'} = y_i(p, m_i')$. Choose any $l \in P^*(f)$ and $l-$cousins $u, u' \in T^*$ where $u_i = m_i$ and $u_i' = m_i'$. $P^*(f) = P^*(l)$ implies $\tilde{P}(f) = \tilde{P}(l) = p$. Thus $x_i^u = y_i(p, m_i)$ and $x_i^{u'} = y_i(p, m_i')$. Also $P^*(f) = P^*(l)$ implies $P'(\gamma(f)) = P'(\gamma(l)) = p'$. Since γ satisfies R2, l being the ancestor of u and u' implies $\gamma(l)$ is the ancestor of $\gamma(u)$ and $\gamma(u')$, then implying $\gamma(l) = \rho_i(\mathbf{x}^u|s') = \rho_i(\mathbf{x}^{u'}|s')$. Moreover we have

$$\begin{aligned}\rho_{i+1}(\mathbf{x}^u|s') &= \rho_i(\mathbf{x}^u|s') \times m_i(x_i^u, P'(\rho_i(\mathbf{x}^u|s'))|s') \\ &= \gamma(l) \times m_i(y_i(p, m_i), p'|s') = \gamma(l) \times m_i(y_i(p, m_i'), p'|s') \\ &= \rho_i(\mathbf{x}^{u'}|s') \times m_i(x_i^{u'}, P'(\rho_i(\mathbf{x}^{u'}|s'))|s') = \rho_{i+1}(\mathbf{x}^{u'}|s')\end{aligned}$$

where $m_i(y_i(p, m_i), p'|s') = m_i(y_i(p, m_i'), p'|s')$ has been established above. Suppose $\rho_j(\mathbf{x}^u|s') = \rho_j(\mathbf{x}^{u'}|s')$ for any $j \leq k$ for some $k \geq i+1$. The inductive

[9] (T'', M'', P'') will be shown to be an improper institution shortly.

hypothesis implies $P'(\rho_k(\mathbf{x}^u|s')) = P'(\rho_k(\mathbf{x}^{u'}|s'))$. Since $\mathbf{x}^{u(k-1)}$ agrees with $\mathbf{x}^u$ for the first $k-1$ components, it follows that $\gamma(u(k-1)) = \rho_k(\mathbf{x}^{u(k-1)}|s') = \rho_k(\mathbf{x}^u|s')$. Similary $\gamma(u'(k-1)) = \rho_k(\mathbf{x}^{u'}|s')$. Thus $P'(\rho_k(\mathbf{x}^u|s')) = P'(\rho_k(\mathbf{x}^{u'}|s'))$ implies $P'(\gamma(u(k-1))) = P'(\gamma(u'(k-1)))$. It follows that $\tilde{P}(u(k-1)) = \tilde{P}(u'(k-1))$ by construction of $\tilde{P}$, which in turn implies $\hat{P}(u(k-1)) = \hat{P}(u'(k-1))$ because $\tilde{P}_k$ is a refinement of $\hat{P}_k$. Since u and u' are cousins, we have $u_k = u'_k$. Thus $y_k(\hat{P}(u(k-1)), u_k) = y_k(\hat{P}(u'(k-1)), u'_k)$. Hence

$$\begin{aligned}\rho_{k+1}(\mathbf{x}^u|s') &= \rho_k(\mathbf{x}^u|s') \times m_k\Big(y_k(\hat{P}(u(k-1), u_k)), P'(\rho_k(\mathbf{x}^u))|s'\Big)\\ &= \rho_k(\mathbf{x}^{u'}|s') \times m_k\Big(y_k(\hat{P}(u'(k-1), u'_k)), P'(\rho_k(\mathbf{x}^{u'}))|s'\Big) = \rho_{k+1}(\mathbf{x}^{u'}|s')\end{aligned}$$

where the first line is due to $x^u_k = y_k(\hat{P}(u(k-1)), u_k)$. Therefore $\rho_{|u|+1}(\mathbf{x}^u|s') = \rho_{|u|+1}(\mathbf{x}^{u'}|s')$ by induction, or equivalently $\gamma(u) = \gamma(u')$. It follows from $P'(\gamma(u)) = P'(\gamma(u'))$ that $\tilde{P}(u) = \tilde{P}(u')$, in turn implying $P^*(u) = P^*(u')$. Hence m_i and m'_i are synonymous given l within (T^*, P^*, M^*). Since l is arbitrarily chosen from $P^*(f)$, it follows that m_i and m'_i are synonymous given $P^*(f)$ within (T^*, P^*, M^*), a contradiction, because by construction (T^*, P^*, M^*) admits no synonymous messages given any perception in P^*. Therefore γ restricted to T^* is one-to-one, establishing R1.

Now show that γ satisfies R4. Choose any $i < N$, $p \in P^*_i$ and $m_i \in M^*(p)$. There exists $\hat{p} \in \hat{P}$ such that $p \subset \hat{p}$. There also exists $p' \in P'$ such that $\gamma(h) \in p'$ for any $h \in p$. For any $h \in p$ we have

$$\begin{aligned}\gamma(h \times m_i) &= \rho_{i+1}(\mathbf{x}^{h\times m_i}|s')\\ &= \rho_i(\mathbf{x}^{h\times m_i}|s') \times m_i\Big(y_i(\hat{p}, m_i), P'(\rho_i(\mathbf{x}^{h\times m_i}|s'))|s'\Big)\\ &= \gamma(h) \times m_i(y_i(\hat{p}, m_i), p'|s')\end{aligned}$$

where the second line is due to $\mathbf{x}^{h\times m_i}_i = y_i(\hat{p}, m_i)$ since $\hat{P}(h) = \hat{p}$. Note that $m_i(y_i(\hat{p}, m_i), p'|s')$ does not depend on the choice of h, thus implying R4.

Now we show that (T'', P'', M'') is indeed an improper institution. The only non-obvious part is that $M''(h) = M''(g)$ if $P''(h) = P''(g)$. Suppose $P''(h) = P''(g)$ yet $M''(h) \neq M''(g)$ for some $h, g \in T''$. Without loss of generality suppose for some $i < N$ there is $m_i \in M''(h)$ such that $m_i \notin M''(g)$. R2 implies $\gamma^{-1}(h \times m_i) = \gamma^{-1}(h) \times \hat{m}_i$ for some $\hat{m}_i \in M^*(\gamma^{-1}(h))$. R3 implies $P^*(\gamma^{-1}(h)) = P^*(\gamma^{-1}(g))$. Thus $\hat{m}_i \in M^*(\gamma^{-1}(g))$. R4 then implies $\gamma(\gamma^{-1}(g) \times \hat{m}_i) = \gamma(\gamma^{-1}(g)) \times m_i = g \times m_i$ since $P^*(\gamma^{-1}(h)) = P^*(\gamma^{-1}(g))$, contradicting the supposition that $m_i \notin M''(g)$. Therefore (T'', P'', M'') is indeed an improper institution and is obtained from (T^*, P^*, M^*) by relabeling with relabeling function γ.

(Step 5. Expanding) Since (T'', P'', M'') is a sub-institution of (T', P', M'), (T', P', M') is obtained from (T'', P'', M'') by expanding. $\square$

A.12 Proof of Proposition 5

The proof is assisted with the following lemma.

Lemma 15. *For institutions (T, P, M) and (T, P', M) respectively representing voting systems (r, d, t) and (r', d, t'), if P'_i is a refinement of P_i for every $i \leq |J|$ and $P'_{|J|+1} = P_{|J|+1}$ then (r', d, t') dominates (r, d, t) for any rule d.*

Proof: Suppose P'_i is a refinement of P_i for every $i \leq |J|$, and $P'_{|J|+1} = P_{|J|+1}$. For every $i \leq |J|$ and $p \in P'_i$ there is $\tau(p) \in P_i$ such that $p \subset \tau(p)$. Choose any $s \in S(T, P, M|d)$. Construct $s' \in S(T, P', M|d)$ such that $m_i(x_i, p|s') = m_i(x_i, \tau(p)|s)$ for every $i \leq |J|$, $x_i \in X_i$ and $p \in P'_i$. Clearly $\rho_1(\mathbf{x}|s') = \rho_1(\mathbf{x}|s)$ for any $\mathbf{x} = (x_1, ..., x_{|J|}, \overline{x}) \in X$. Suppose $\rho_i(\mathbf{x}|s') = \rho_i(\mathbf{x}|s)$ for every $i \leq k$ for some $k \geq 1$. Thus $\tau(P'(\rho_k(\mathbf{x}|s'))) = P(\rho_k(\mathbf{x}|s')) = P(\rho_k(\mathbf{x}|s))$. Therefore

$$\begin{aligned}\rho_{k+1}(\mathbf{x}|s') &= \rho_k(\mathbf{x}|s') \times m_k(x_k, P'(\rho_k(\mathbf{x}|s'))|s') \\ &= \rho_k(\mathbf{x}|s) \times m_k(x_k, P(\rho_k(\mathbf{x}|s))|s) = \rho_{k+1}(\mathbf{x}|s).\end{aligned}$$

Hence $\alpha_{|J|+1}(\mathbf{x}|s') = d(\rho_{|J|+1}(\mathbf{x}|s')) = d(\rho_{|J|+1}(\mathbf{x}|s)) = \alpha_{|J|+1}(\mathbf{x}|s)$. It follows that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T, P, M|d) \subset C(T, P', M|d)$. The lemma follows from Lemma 9. □

Proof of Proposition 5. (Proof of Part 1.) Let (T, P, M) be the institution representing (r, d, t). Let t' be the full disclosure policy. Let (T, P', M) be the institution representing (r, d, t'). Therefore P'_i is a refinement of P_i for every $i \leq |J|$, and $P'_{|J|+1} = P_{|J|+1}$. Thus (r, d, t') dominates (r, d, t) by Lemma 15. Part 1 is established.

(Proof of Part 2.) Suppose t is the full disclosure policy, let r' be the sequential procedure such that $r'(i) > r'(j)$ is $r(i) > r(j)$. Using the same argument as in the previous paragraph we conclude that (r', d, t) dominates (r, d, t), establishing Part 2.

(Proof of Part 3.) Suppose t is the full disclosure policy. Let i denote the player with the largest index among those who vote before the last stage (according to r). It follows that each $p \in P_{|J|}$ can be uniquely identified as $p^{(z_1,...,z_i)}$ such that $P(h) = P(h') = p$ if and only if $h(i) = h'(i) = (z_1, ..., z_i)$.

Let $\hat{Z}$ denote the set of vote profiles $(z_1, ..., z_i)$ from voters voting before the last stage such that $d(z_1, ..., z_i, z_{i+1}, ..., z_{|J|}) = d(z_1, ..., z_i, z'_{i+1}, ..., z'_{|J|})$ for any vote profiles $(z_{i+1}, ..., z_{|J|})$ and $(z'_{i+1}, ..., z'_{|J|})$ from voters voting in the last stage. The collective choice is not determined before voting in the last stage takes place if and only if $\hat{Z}$ is empty. It follows that for any $(z_1, ..., z_i) \in \hat{Z}$, every pair of votes (messages) $z, z' \in Z$ are synonymous given $p^{(z_1,...,z_i)}$ within (T, P, M).

Choose any $s \in S(T, P, M|d)$. Construct $\hat{s} \in S(T, P, M|d)$ such that

1. For any $i < |J|$, $x_i \in X_i$ and $p \in P_i$: $m_i(x_i, p|\hat{s}) = m_i(x_i, p|s)$.

2. For any $x_{|J|} \in X_{|J|}$ and $(z_1, ..., z_i) \notin \hat{Z}$:

$$m_{|J|}(x_{|J|}, p^{(z_1,...,z_i)}|\hat{s}) = m_{|J|}(x_{|J|}, p^{(z_1,...,z_i)}|s).$$

3. For any $x_{|J|} \in X_{|J|}$ and $(z_1, ..., z_i) \in \hat{Z}$:

$$m_{|J|}(x_{|J|}, p^{(z_1,...,z_i)}|\hat{s}) = \overline{z}$$

for some fixed $\overline{z} \in Z$.

Fix any $\mathbf{x} = (x_1, ..., x_{|J|}, \overline{x}) \in X$. Clearly $\rho_{|J|}(\mathbf{x}|\hat{s}) = \rho_{|J|}(\mathbf{x}|s)$. Denote $h = \rho_{|J|}(\mathbf{x}|\hat{s})$. If $P(h) = p^{(z_1,...,z_i)}$ where $(z_1, ..., z_i) \notin \hat{Z}$ then $\mu_{|J|}(\mathbf{x}|\hat{s}) = \mu_{|J|}(\mathbf{x}|s)$ and it follows that $\rho_{|J|+1}(\mathbf{x}|\hat{s}) = \rho_{|J|+1}(\mathbf{x}|s)$. If $P(h) = p^{(z_1,...,z_i)}$ where $(z_1, ..., z_i) \in \hat{Z}$ then $\mu_{|J|}(\mathbf{x}|\hat{s})$ is synonymous to $\mu_{|J|}(\mathbf{x}|s)$ given h within (T, P, M) and hence

$$P(\rho_{|J|+1}(\mathbf{x}|\hat{s})) = P(h \times \mu_{|J|}(\mathbf{x}|\hat{s})) = P(h \times \mu_{|J|}(\mathbf{x}|s)) = P(\rho_{|J|+1}(\mathbf{x}|s)).$$

We have established that $P(\rho_{|J|+1}(\mathbf{x}|\hat{s})) = P(\rho_{|J|+1}(\mathbf{x}|s))$ for any $\mathbf{x} \in X$, implying $\alpha_{|J|+1}(\mathbf{x}|\hat{s}) = \alpha_{|J|+1}(\mathbf{x}|s)$ because $s, s' \in S(T, P, M|d)$.

For any $(z_1, ..., z_i) \in \hat{Z}$ choose $y^{(z_1,...,z_i)} \in Y$ where $y^{(z_1,...,z_i)}$ is different from the candidate that is elected given any vote profile that contains $(z_1, ..., z_i)$. Construct d' such that

$$d'(z_1, ..., z_{|J|}) = \begin{cases} y^{(z_1,...,z_i)} & \text{if } (z_1, ..., z_i) \in \hat{Z} \text{ and } z_{|J|} \neq \overline{z} \\ d(z_1, ..., z_{|J|}) & \text{otherwise.} \end{cases}$$

It is straightforward to verify that under d' the collective choice is not determined before voting in the last stage takes place. Let (T, P', M) be the institution representing (r, d', t). Clearly $P'_i = P_i$ for every $i \leq |J|$. Choose any $s' \in S(T, P', M|d')$ such that $m_i(x_i, p|s') = m_i(x_i, p|\hat{s})$ for any $i \leq |J|$, $x_i \in X_i$ and $p \in P'_i$.

Arbitrarily choose $\mathbf{x} \in X$. Denote $g = \rho_{|J|}(\mathbf{x}|s')$. Clearly $\rho_{|J|+1}(\mathbf{x}|s') = \rho_{|J|+1}(\mathbf{x}|\hat{s})$ for any $\mathbf{x}$. If $g(i) \notin \hat{Z}$ then $\alpha_{|J|+1}(\mathbf{x}|s') = d'(\rho_{|J|+1}(\mathbf{x}|s')) = d(\rho_{|J|+1}(\mathbf{x}|\hat{s})) = \alpha_{|J|+1}(\mathbf{x}|\hat{s}) = \alpha_{|J|+1}(\mathbf{x}|s)$. If $g(i) \in \hat{Z}$ then $\mu(\mathbf{x}|s') = \overline{z}$, and following a similar sequence of equalities we have $\alpha_{|J|+1}(\mathbf{x}|s') = \alpha_{|J|+1}(\mathbf{x}|s)$. Thus $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T, P, M|d) \subset C(T, P', M|d')$. It follows from Lemma 9 that (r, d', t) dominates (r, d, t). □

A.13 Proof of Proposition 6

Proof. Fix institution (T, P, M) where there is some i such that $|P_i| < |T_i|$. Let P' be a partition of T such that $P'_j = P_j$ for every $j \neq i$ and P'_i is a strict refinement of P_i. That $|P_i| < |T_i|$ implies such P'_i exists. Both complexity conditions in the proposition are satisfied by the institution (T, P', M). (T, P', M) dominates (T, P, M) by Lemma 3 because the former is obtained from the latter by refining. If (T, P, M) dominates (T, P', M) then by Theorem 1 (T, P, M) can be obtained from (T, P', M) by a sequence of operations of expanding, refining, trimming or relabeling. Note that merging does not apply because none of the player has a singleton action set. Since none of the operations decrease the number of perceptions of i, [10] it follows that $|P_i| \geq |P'_i|$, a contradiction. Thus (T, P, M) does not dominate (T, P', M). □

A.14 Proof of Proposition 7

Proof. Fix institution (T, P, M) where $|P_{i+1}| \geq 2$ for some $i \geq 2$. Let (T', P', M') be obtained from (T, P, M) by expanding, such that

1. $M'_j = M_j$ for any $j < N$ and $j \notin \{i-1, i\}$.
2. $M_{i-1} \subset M'_{i-1}$.
3. $M'_i = M_i \cup \{\overline{m}_i\}$ for some $\overline{m}_i \notin M_i$.
4. $|P'_j| = |P_j|$ for any $j \in \mathscr{N}$.
5. There is some $p' \in P'_i$ such that $h \in p'$ for any $h \in T'_i \backslash T_i$.
6. For any $m_i, m'_i \in M'_i$ there is some $h \in p'$ such that m_i and m'_i are not synonymous given h within (T', P', M').

An institution satisfying Properties $1-5$ can be easily constructed by making new messages available to players $i-1$ and i, and put message profiles containing the newly introduced messages to existing perceptions as prescribed by Property 5. If the institution (T', P', M') expanded from (T, P, M) satisfying Properties 1-5 does not satisfy Property 6, that is, there are $m_i, m'_i \in M'_i$ that are synonymous given any $h \in p'$ within (T', P', M'), then we can expand (T', P', M') by making an additional message $\overline{m}_{i-1}$ available to player $i-1$, and moreover:

- Put $g \times \overline{m}_{i-1}$ in p' for every $g \in T'_{i-1}$.
- Put $(g \times \overline{m}_{i-1}) \times m_i$ and $(g \times \overline{m}_{i-1}) \times m'_i$ in different perceptions of player $i+1$ for every $g \in T'_{i-1}$.
- Put any message profiles of length $j-1$ containing $\overline{m}_{i-1}$ arbitrarily to existing perceptions of player j for any $j > i+1$.

Let (T'', P'', M'') be the consequent institution. (T'', P'', M'') satisfies Properties 1-5. Let $p'' \in P''_i$ be the consequent perception enlarged from p'. It follows that m_i and m'_i are not synonymous given $g \times \overline{m}_{i-1}$ for any $g \in T''_{i-1}$

[10] By Lemma 13, trimming does not decrease the number of perceptions of any player.

and hence they are not synonymous given p'' within (T'', P'', M''). We can keep applying this particular type of expanding until Property 6 is satisfied, without violating Properties 1-5.

Let (T', P', M') be expanded from (T, P, M) that satisfies Properties 1-6. (T', P', M') dominates (T, P, M) by Lemma 2. If (T, P, M) dominates (T', P', M') then the proof of Theorem 1 implies there are institutions (T^1, P^1, M^1), (T^2, P^2, M^2), (T^3, P^3, M^3) and (T^4, P^4, M^4) such that

1. (T^1, P^1, M^1) is obtained from (T', P', M') by merging.
2. (T^2, P^2, M^2) is obtained from (T^1, P^1, M^1) by refining.
3. (T^3, P^3, M^3) is obtained from (T^2, P^2, M^2) by a sequence of trimming operations.
4. (T^4, P^4, M^4) is obtained from (T^3, P^3, M^3) by relabeling.
5. (T, P, M) is obtained from (T^4, P^4, M^4) by expanding.

$(T^1, P^1, M^1) = (T', P', M')$ because everyone's action set is non-singleton. That $|P'_j| = |P_j|$ for every j implies $(T^2, P^2, M^2) = (T^1, P^1, M^1)$ because strict refining increases the number of perceptions for some player, which will not be decreased by trimming, relabeling or expanding. Thus $(T^2, P^2, M^2) = (T', P', M')$. $p' \cap T^3_i \in P^3_i$ because trimming does not decrease the number of perceptions by Lemma 13. Since there do not exist $m_i, m'_i \in M'_i$ which are synonymous given p' within (T', P', M'), $M^3(p \cap T^3) = M'_i$. Thus for any $h' \in p' \cap T^3$ we have $|M^3(h')| = |M'_i| = |M_i| + 1$, implying there is $h \in T_i$ such that $|M(h)| = |M_i| + 1$ because relabeling and expanding do not decrease the number of children of any message profile. This leads to a contradiction because $|M(h)| = |M_i|$ for any $h \in T_i$. Hence (T, P, M) does not dominate (T', P', M'). $\square$