Generalisation of Samet's (2010) agreement theorem

Bassel Tarbush

23 February 2011

Generalisation of Samet's (2010) agreement theorem

Bassel Tarbush[0]

May 27, 2011

Abstract We develop a framework that allows us to reproduce the generalised agreement theorem of Samet (2010), and extend it to models in which agents can base their decisions on false information, while highlighting the features that distinguish the result from the classic theorems found in the literature. For example, it allows decisions to be based on interactive information, and imposes no requirements on the language in which the states are described. Finally, we produce a result that does not require Samet's assumption of the existence of a completely uninformed agent.

Keywords Agreeing to disagree, knowledge, common knowledge, belief, information, epistemic logic.
JEL classification D80, D83, D89.

1 Introduction

The agreement theorem of Aumann (1976) states that if agents have a common prior, then if their posteriors on some event are common knowledge, these posteriors must be equal, *even if the agents' updates are based on different information.* This was proved for posterior probabilities in the context of a *partitional* information structure.
This result was extended by many authors to generalised decision functions, instead of posterior probabilities (see Cave (1983), Bacharach (1985), Moses and Nachum (1990), Bonanno and Nehring (1998), Aumann and Hart (2006)). However, all these generalisations have relied on the imposition of some version of the Sure-Thing Principle as a condition on the decision functions. Informally, all

[0]Department of Economics, University of Oxford, bassel.tarbush[at]economics.ox.ac.uk
I would like to thank John Quah for invaluable help and Francis Dennig for very useful discussions.

versions of this principle attempt to capture the following intuition: "If I would perform some action when I know that p is the case, and I would perform the same action when I know that p is not the case, then I should also perform that same action when I do not know whether p is the case".

Samet (2010) also derives a generalised agreement theorem in a partitional information structure. However, his approach differs significantly from the classic examples in the literature in that he does not use a standard version of the Sure-Thing Principle. Rather, Samet assumes an "interpersonal" Sure-Thing Principle ($ISTP$) which can informally be stated as: "If I have some information, but I know that whatever information I have about something, you will be better informed about it than me, then if I know your action, I should perform that same action". So, unlike the standard versions of the principle, which are conditions over the decision function of a single agent, the $ISTP$ is a condition imposed on the actions across agents.
To obtain his result, Samet also requires the existence of a "dummy" agent, who is an agent that is less informed than all other agents.
Given this, we can provide an informal statement of Samet's result.

> *If the $ISTP$ holds and there exists a dummy agent, then if the actions of all the agents are common knowledge, then their actions are identical.*

We develop a syntactic framework, using concepts from epistemic logic, which allows us to reproduce Samet's result in a partitional information structure. However, we are also able to keep track of some more subtle features of the result. For example, we show that Samet's result allows for actions to be based on interactive knowledge, whereas standard results require them to be independent of such information. Furthermore, we extend Samet's result to a non-paritional information structure. Partitional information structures imply that agents can only know what is the case; in other words, agents cannot base their actions on false information. But surely, it is perfectly plausible for rational agents to do so. So our extension effectively states that agents cannot agree to disagree even when their actions are based on interactive knowledge (or belief) *and* possibly false information.

In section 2, we introduce the basic concepts that we use from epistemic logic. In section 3, we expand the standard epistemic logic framework to agents performing actions, and we state our main assumptions. We derive our main results in partitional models in section 4, and in non-partitional models in section 5. In the latter section, we also derive an agreement theorem that replaces Samet's assumption about the existence of a dummy agent with an alternative one. All proofs are in the appendix.

2 Epistemic Logic

This section introduces concepts from epistemic logic. All the definitions and results in this section are standard (e.g. see Chellas (1980) and van Benthem (2010) for general reference).

Definition 1 (Basic syntax)**.** Define a finite set of atomic *propositions*, $\mathcal{P}$, which consists of all propositions that cannot be further reduced. Let N denote the set of all agents. We then inductively create all the *formulas* in our language, $\mathcal{L}$, as follows:
(i) Every $p \in \mathcal{P}$ is a formula.
(ii) If ψ is a formula, so is $\neg\psi$.
(iii) If ψ and ϕ are formulas, then so is $\psi \circ \phi$, where $\circ$ is one of the following Boolean operators: $\wedge$, $\vee$, $\rightarrow$, or $\leftrightarrow$.
(iv) If ψ is a formula, then so is $\bullet\psi$, where $\bullet$ is one of the modal operators $\Box_{i\in N}$ or $C_{G\subseteq N}$.
(v) Nothing else is a formula.

Note that $\Box_i$ and C_G are modal operators, while $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ are the standard Boolean operators.

Definition 2 (Modal depth)**.** The modal depth $md(\psi)$ of a formula ψ is the maximal length of a nested sequence of modal operators. This can be defined by the following recursion on our syntax rules: (i) $md(p) = 0$ for any $p \in \mathcal{P}$, (ii) $md(\neg\psi) = md(\psi)$, (iii) $md(\psi \wedge \phi) = md(\psi \vee \phi) = md(\psi \rightarrow \phi) = md(\psi \leftrightarrow \phi) = \max(md(\psi), md(\phi))$, (iv) $md(\Box_i\psi) = 1 + md(\psi)$, (v) $md(C_G\psi) = 1 + md(\psi)$.

So far, we have pure uninterpreted syntax. However, we can now introduce our semantics, to determine the truth or falsity of formulas.

Definition 3 (Kripke semantics)**.** A *frame* is a pair $\langle\Omega, R_{i\in N}\rangle$, where Ω is a finite, non-empty set of *states* (or "possible worlds), and $R_i \subseteq \Omega \times \Omega$ is a binary relation for each agent i, also called the *accessibility relation* for agent i. A *model* on a frame $\langle\Omega, R_{i\in N}\rangle$, is a triple $\mathcal{M} = \langle\Omega, R_{i\in N}, \mathcal{V}\rangle$, where $\mathcal{V} : \mathcal{P} \times \Omega \rightarrow \{0, 1\}$ is a *valuation map.*

Definition 4 (Truth)**.** We say that a proposition $p \in \mathcal{P}$ is *true* at state ω in model $\mathcal{M} = \langle\Omega, R_{i\in N}, \mathcal{V}\rangle$, denoted $\mathcal{M}, \omega \models p$, if and only if $\mathcal{V}(p, \omega) = 1$. Truth is then extended inductively to all other formulas ψ as follows:
(i) $\mathcal{M}, \omega \models \neg\psi$ if and only if it is not the case that $\mathcal{M}, \omega \models \psi$.
(ii) $\mathcal{M}, \omega \models (\psi \wedge \phi)$ if and only if $\mathcal{M}, \omega \models \psi$ and $\mathcal{M}, \omega \models \phi$.[1]
(iii) $\mathcal{M}, \omega \models \Box_i\psi$ if and only if $\forall\omega' \in \Omega$, if $\omega R_i \omega'$ then $\mathcal{M}, \omega' \models \psi$.

[1]The truth of formulas involving the other Boolean operators are similarly defined.

(iv) $\mathcal{M}, \omega \models C_G\psi$ if and only if $\forall\omega' \in \Omega_G(\omega)$, $\mathcal{M}, \omega' \models \psi$.
The *component* of ω, $\Omega_G(\omega)$, is the set of all states that are accessible from ω in a finite sequence of R_i $(i \in G)$ steps.

Note that if $\mathcal{M}, \omega \models C_G\psi$, then one can generate any formula of finite modal depth of the form $\Box_i\Box_j...\Box_r\psi$ with $i, j...r \in G$, and this formula will be true at ω in model $\mathcal{M}$.[2]

Definition 5 (Validity)**.** Formula ψ is valid in a model $\mathcal{M}$, denoted $\mathcal{M} \models \psi$ if and only if $\forall\omega \in \Omega$ in $\mathcal{M}$, $\omega \models \psi$. Formula ψ is valid in a frame $\langle\Omega, R_{i\in N}\rangle$, denoted $\langle\Omega, R_{i\in N}\rangle \models \psi$, if and only if $\forall\mathcal{M}$ over $\langle\Omega, R_{i\in N}\rangle$, $\mathcal{M} \models \psi$. Formula ψ is $\mathcal{T}$-valid (or valid), denoted $\models \psi$, if and only if $\forall\langle\Omega, R_{i\in N}\rangle \in \mathcal{T}$ ($\mathcal{T}$, a collection of frames), $\langle\Omega,\ R_{i\in N}\rangle \models \psi$.

We can identify classes of frames by the restrictions that we impose on the accessibility relations.

Definition 6 (Conditions on frames)**.** We say that a frame $\langle\Omega, R_{i\in N}\rangle$ is,

Reflexive	if $\forall i \in N, \forall\omega \in \Omega, \omega R_i\omega$
Symmetric	if $\forall i \in N, \forall\omega, \omega' \in \Omega,$ if $\omega R_i\omega'$ then $\omega' R_i\omega$
Transitive	if $\forall i \in N, \forall\omega, \omega', \omega'' \in \Omega,$ if $\omega R_i\omega'$ and $\omega' R_i\omega''$ then $\omega R_i\omega''$
Euclidean	if $\forall i \in N, \forall\omega, \omega', \omega'' \in \Omega,$ if $\omega R_i\omega'$ and $\omega R_i\omega''$ then $\omega' R_i\omega''$
Serial	if $\forall i \in N, \forall\omega \in \Omega, \exists\omega' \in \Omega, \omega R_i\omega'$

The system $S5$ consists of all frames that are reflexive, symmetric and transitive; and the system $KD45$ consists of all frames that are serial, transitive and Euclidean. The following formulas are validities in the respective frames, and in fact, the systems can be axiomatised in the sense that if the validities are assumed then they imply the desired restrictions on the accessibility relations:

$S5$ axioms	$KD45$ axioms	Axiom names
$\Box_i(\psi \to \phi) \to (\Box_i\psi \to \Box_i\phi)$	$\Box_i(\psi \to \phi) \to (\Box_i\psi \to \Box_i\phi)$	Distribution
$\Box_i\psi \to \psi$	$\Box_i\psi \to \neg\Box_i\neg\psi$	Veracity; Consistency
$\Box_i\psi \to \Box_i\Box_i\psi$	$\Box_i\psi \to \Box_i\Box_i\psi$	Positive introspection
$\neg\Box_i\psi \to \Box_i\neg\Box_i\psi$	$\neg\Box_i\psi \to \Box_i\neg\Box_i\psi$	Negative introspection

[2]Note that the definition of the operator C_G is drawn from van Benthem (2010), where it is also mentioned that an alternative definition can be given: One can define a new accessibility relation R^*_G for the whole group G as the reflexive transitive closure of the union of all separate relations R_i $(i \in G)$, and then simply let $\mathcal{M}, \omega \models C_G\psi$ if and only if $\forall\omega' \in \Omega,$ if $\omega R^*_G\omega'$ then $\mathcal{M}, \omega' \models \psi$.

It is standard to take the axioms of $S5$ as describing properties of (a rather strong notion of) knowledge. Thus, in $S5$, $\Box_i\psi$ is interpreted as "agent i *knows* that ψ". In $KD45$ however, since veracity is dropped in favour of consistency, we are in a system in which to "know" that something is the case does not imply that it is true. The axioms of $KD45$ are thus rather seen as describing properties of a belief operator, so $\Box_i\psi$ is interpreted as "agent i *believes* that ψ". These two systems mirror the patitional and non-partitional structures mentioned in the introduction.[3]

Similarly, the operator $C_G\psi$ is interpreted as "it is common knowledge to all the agents in G that ψ" in $S5$, and as "it is common belief to all the agents in G that ψ" in $KD45$.

3 Models with information and decisions

Let P be a finite set of atomic propositions. Since P is finite, its closure under the standard Boolean operators, denoted P^*, is tautologically finite.[4] So P^* is just the set of all possible inequivalent formulas that can be created out of the propositions in P and the Boolean operators. Let Ψ_0^r be the set of all possible modal formulas that can be generated from P^* with modal depth 0 up to r for an arbitrary $r \in \mathbb{N}_0$. Again, since P^* is finite, so is Ψ_0^r, so $|\Psi_0^r| = m$, for some $m \in \mathbb{N}$; and note that $\Psi_0^0 = P^*$.[5]

Definition 7 (New operators)**.** For each agent $i \in N$ create a set of modal operators, $O_i = \{\Box_i, \hat{\Box}_i, \dot{\Box}_i\}$, where for every formula ψ, $\hat{\Box}_i\psi := \Box_i\neg\psi$ and $\dot{\Box}_i\psi := \neg(\Box_i\psi \vee \hat{\Box}_i\psi)$.

The interpretation, for example in $S5$, is that $\hat{\Box}_i\psi$ stands for "agent i knows that it is not the case that ψ", and $\dot{\Box}_i\psi$ stands for "agent i does not know whether it is the case that ψ". There are similar counterpart interpretations in $KD45$.

Definition 8 (Kens)**.** Order the set Ψ_0^r into a vector of length m: $(\psi_1, \psi_2, ..., \psi_m)$, and for each agent $i \in N$, create the sets

$$U_i = \{(\nu_i^1\psi_1 \wedge \nu_i^2\psi_2 \wedge ... \wedge \nu_i^m\psi_m) | \forall n \in \{1, ..., m\}, \nu_i^n \in O_i\}$$

$$V_i = \{\nu_i \in U_i | \models \neg(\nu_i \leftrightarrow (p \wedge \neg p))\}$$

[3]The philosophical grounds for these systems originated in Hintikka (1962), and for an extensive formal treatment, see Chellas (1980).

[4]In the sense that there is only a finite number of inequivalent formulas (so p and $p \wedge p$ count as one).

[5]If $P = \{p, q\}$, then one can generate 20 inequivalent formulas: 2 from p alone, 2 from q alone and 16 out of p and q together, so $|P^*| = 20$.

A *ken* ($\nu_i \in V_i$) for agent i, describes i's information concerning every formula in Ψ_0^r. So, calling $\nu_i^n \psi_n$ the n^{th} *entry* of i's ken, the formula $\nu_i^n \psi_n$ states - in $S5$ - whether i knows that the formula ψ_n is the case, or knows that it is not the case, or does not know whether it is the case.
Note that V_i is a restriction of U_i to the set of kens that are not logically equivalent to a contradiction; so only the logically consistent kens are considered.[6]

The following lemma shows that at each state, there exists a ken for each agent which holds at that state, and moreover, that any two different kens must be contradictory at any given state.

Lemma 1. *(i)* $\forall \omega \in \Omega, \exists \nu_i \in V_i, \omega \models \nu_i$, *(ii)* $\forall \omega \in \Omega, \forall \nu_i, \mu_i \in V_i$, *if* $\nu_i \neq \mu_i$ *then* $\omega \models \neg(\nu_i \wedge \mu_i)$.

By the above lemma, there is a unique ken in V_i that holds at a given state.

Definition 9 (Informativeness)**.** Create an order $\succsim \subseteq V_i \times V_j$ for all $i, j \in N$. We say that the ken ν_i is *more informative* than the ken μ_j, denoted $\nu_i \succsim \mu_j$, if and only if whenever i knows that ψ then j either also knows that ψ or does not know whether ψ, and whenever i does not know whether ψ, then so does j.[7]
Note that $\succsim$ is not a complete order on kens. For example, consider any two kens ν_i and μ_i for agent i, in which the n^{th} entry is $\nu_j^n \psi_n = \Box_i \psi_n$ and $\mu_j^n \psi_n = \hat{\Box}_i \psi_n$. These two kens would not be comparable with $\succsim$.
Finally, note that $\nu_i \sim \mu_j$ denotes $\nu_i \succsim \mu_j$ and $\mu_j \succsim \nu_i$; which is interpreted as ν_i and μ_j carrying the *same* information, but seen from the perspectives of agents i and j respectively.

Definition 10 (Actions)**.** We will add formulas of the form d_i^x to our syntax, which are read as "Agent i performs action x".

We assume that there can only be one action for any given ken; and since there is a unique ken that is true at any given state, there is a unique action per state. Furthermore, we assume - uncontroversially - that if the same ken is true at different states, then the same action must be taken at those states.

3.1 Main assumption

We will assume that the Interpersonal Sure-Thing Principle is a formula, $ISTP$, that is valid in every model that we will consider.

[6] An example of a logically inconsistent ken would be one containing $\Box_i p$, $\Box_i q$ and $\Box_i(p \to \neg q)$.

[7] Formally, (i) if $\nu_i^n \psi_n = \Box_i \psi_n$ then ($\mu_j^n \psi_n = \Box_j \psi_n$ or $\mu_j^n \psi_n = \dot{\Box}_j \psi_n$), (ii) if $\nu_i^n \psi_n = \hat{\Box}_i \psi_n$ then ($\mu_j^n \psi_n = \hat{\Box}_j \psi_n$ or $\mu_j^n \psi_n = \dot{\Box}_j \psi_n$), and (iii) if $\nu_i^n \psi_n = \dot{\Box}_i \psi_n$ then ($\mu_j^n \psi_n = \dot{\Box}_j \psi_n$).

Assumption 1 (Interpersonal Sure-Thing Principle - *ISTP*)**.** *For all* $\omega \in \Omega$, $\omega \models (\nu_i \wedge \nu_j) \wedge \Box_i(\nu_j \rightarrow \nu_j \succsim \nu_i) \rightarrow (\Box_i(d_j^x) \rightarrow d_i^x)$

The above states that for any agents i and j, if i knows (believes) that j's ken is more informative than hers ($\nu_j \succsim \nu_i$), then if i knows (believes) that j performs action x, then i performs action x. Note that this does *not* require i to know (believe) j's ken! Rather, it simply requires i to think that *if* j's ken *were* ν_j - whatever it may be - then this ken would be more informative than hers.

4 Results in $S5$

In $S5$, the accessibility relation R_i is an equivalence relation for each $i \in N$. Let $I_i(\omega) = \{\omega' \in \Omega | \omega R_i \omega'\}$ be the *information cell* of i at ω. One can verify that the set $\mathcal{I}_i = \{I_i(\omega) | \omega \in \Omega\}$ is a partition of the state space Ω.

The following lemma states that at any state in which the information cell of agent i is a subset of agent j's cell at that state, then j's ken is more informative than i's ken at that state.

Lemma 2. *For any* $\omega \in \Omega$ *such that* $I_i(\omega) \subseteq I_j(\omega)$, *if* $\omega \models \nu_i \wedge \nu_j$ *then* $\omega \models \nu_i \succsim \nu_j$.

We will require two further lemmas.

Lemma 3. $\forall i \in G$, $\bigcup_{\omega' \in \Omega_G(\omega)} I_i(\omega') = \Omega_G(\omega)$.

Lemma 4. *If for some* $\omega' \in I_i(\omega)$, $\omega' \models \nu_i$, *then for all* $\omega'' \in I_i(\omega)$, $\omega'' \models \nu_i$.

Samet (2010) assumes that there always exists an "epistemic dummy": An agent whose information cell is equal to the entire component $\Omega_G(\omega)$.

Assumption 2 (Epistemic dummy)**.** $\exists h \in G, I_h(\omega) = \Omega_G(\omega)$.

Theorem 1. *Suppose that there exists an epistemic dummy, ISTP holds, and that the system is* $S5$. *Let* $G = \{i, j, h\}$ *with* h *the epistemic dummy. Then,* $\models C_G(d_i^x \wedge d_j^y \wedge d_h^z) \rightarrow (x = y = z)$.

Note that there is a slight abuse of notation in the statement of the theorem above. Technically, "=" is not part of our syntax, so $x = y$ should not appear anywhere. However, we simply use it as shorthand. Our results should really be read as: $\models C_G(d_i^x \wedge d_j^y \wedge d_h^z) \rightarrow (d_i^w \wedge d_j^w \wedge d_h^w)$, and $w = z = x = y$.

4.1 Discussion

The intuition driving the result is that by assuming that there exists an epistemic dummy, one is assuming that there is an agent h whose performed action is based on a ken that is less informative than every other agents'. However, h knows the performed actions of the other agents, and knows that those actions are based on information that is more informative than her ken. She therefore models her choice on the performed actions of each of the other agents. But if those more informed agents were taking different actions then she would have to simultaneously copy two different actions, which is impossible, thus the actions of the more informed agents must be the same.

Note that the result would also hold if we had used this alternative version of the dummy assumption:

Assumption 3 (Dummy*). *There is an $h \in G$ such that for all $\omega' \in \Omega_G(\omega)$, if $\omega' \models \nu_h \wedge \nu_i$, then $\nu_i \succsim \nu_h$ for all $i \neq h$.*

In fact, this assumption is *weaker* than the original epistemic dummy assumption because in principle, it allows the dummy agent to have different information across different states within the same component, whereas the original assumption forces the dummy to have the same information across all states within a component.

In Tarbush (2011) it is shown that previous agreement theorems require the assumption that actions only be based on kens where Ψ_0^r is such that $r = 0$. That is, actions cannot be based on *interactive* information.[8] So, in previous results, agents *can* agree to disagree if say i bases her decision on what she knows about what j knows. However, one of the main distinguishing features of Samet's result is that this restriction does not need to be imposed.
Furthermore, when the "Disjoint Sure-Thing Principle" is imposed on decision functions in previous results (which emulates Bacharach's (1985) original condition), the language must be assumed to be "rich" enough to guarantee that information (or kens) are, in a sense, "disjoint".[9] The implication is that whether or not the agreement results hold depends on the way in which the states are described! However, again, Samet's result requires no such condition.

[8] As explained in the paper, this is in response to the criticism (Moses and Nachum (1990)) of Bacharach (1985).

[9] The language in a component $\Omega_G(\omega)$ is *rich* if and only if for all $i \in G$, and any $\omega', \omega'' \in \Omega_G(\omega)$ such that $\omega' \models \nu_i$, $\omega'' \models \mu_i$ and $\nu_i \neq \mu_i$, there is $n \in \{1, ..., m\}$ such that $\nu_i^n = \Box_i$ and $\mu_i^n = \hat{\Box}_i$.

5 Results in $KD45$

We can now analyse the consequences of using a model for *belief* rather than knowledge. So we impose a $KD45$ frame rather than an $S5$ frame.
Essentially, the only difference between knowledge and belief that we will consider is that belief is not *infallible*. In $S5$, agents cannot know something that is false, because reflexivity implies that if one knows that p at some state, then p must be true at that state (Veracity). On the other hand, $KD45$ allows agents to believe what is false, and thus to base decision on *false* information, by dropping reflexivity. In fact, $S5 = KD45 +$ reflexivity.

We can provide a description of the links between states in a $KD45$ frame: Some sets of states within Ω are "completely connected", in the sense that the accessibility relation over states within such sets in an equivalence relation, so these sets have the same properties as information cells in $S5$; and, for each one of these completely connected sets there exists a (possibly empty) set of "associated" states that have arrows pointing from them to every state in the completely connected set, but with no arrow (by the same agent) pointing towards them. The set of all completely connected sets and their set of associated states exhaust the state space.
Formally, let $S_i(\omega) = \{\omega' \in \Omega | \omega E_i \omega'\}$, where E_i is an equivalence relation. We call this set of completely connected states the *information sink* of state ω for player i. The set S_i do not necessarily partition the state space, hence we have a *non-partitional* model. Note, that this way of defining the sink guarantees that if $S_i(\omega) \neq \emptyset$ then $\omega \in S_i(\omega)$. Furthermore, we define ω's *set of associated states* as $A_i(\omega) = \{\omega'' \in \Omega | \forall \omega''' \in S_i(\omega), \omega'' F_i \omega'''\}$, where F_i is now a simple arrow. So, note that now, for any agent i, we have that $R_i = E_i \cup F_i$. Finally, we can define $J_i(\omega) = S_i(\omega) \cup A_i(\omega)$, and note that $\mathcal{J}_i = \{J_i(\omega) | \omega \in \Omega\}$ exhausts the entire state space.

Proposition 1. *The above is a complete characterisation of the $KD45$ state space.*

We now use lemmas that are analogous to the ones used in $S5$.

Lemma 5. *For any $\omega \in \Omega$ such that $S_i(\omega) \subseteq S_j(\omega)$, if $\omega \models \nu_i \wedge \nu_j$ then $\omega \models \nu_i \succsim \nu_j$.*

Lemma 6. *$\forall i \in G$, $\bigcup_{\omega' \in \Omega_G(\omega)} S_i(\omega') \subseteq \Omega_G(\omega) \subseteq \bigcup_{\omega' \in \Omega_G(\omega)} J_i(\omega')$.*

Lemma 7. *If for some $\omega' \in J_i(\omega)$, $\omega' \models \nu_i$, then for all $\omega'' \in J_i(\omega)$, $\omega'' \models \nu_i$.*

We now require an assumption that is analogous to the epistemic dummy assumption.

Assumption 4 (Doxastic dummy)**.** *$\exists h \in G$, $\exists \omega' \in \Omega_G(\omega)$, such that for $k \in G \backslash \{h\}$, $\bigcup_{\omega'' \in \Omega_G(\omega)} S_k(\omega'') \subseteq S_h(\omega')$ and $J_h(\omega') = \Omega_G(\omega) \cup \{\omega\}$.*

This assumption requires that some agent's (the dummy's) unique information sink be a superset of the union of the information sinks of every other agent in the component.

Theorem 2. *Suppose that there exists a doxastic dummy, $ISTP$ holds, and that the system is $KD45$. Let $G = \{i, j, h\}$ with h the doxastic dummy. Then, $\models C_G(d_i^x \wedge d_j^y \wedge d_h^z) \rightarrow (x = y = z)$.*

5.1 Discussion

The only substantial difference between theorem in $S5$ and the one is $KD45$ is the assumption made about the dummy. Note that an alternative assumption could have been: $\exists h \in G$, $S_h(\omega) = \Omega_G(\omega) \cup \{\omega\}$. One can verify that this *implies* the doxastic dummy assumption. However, we see it as being unreasonably strong: It implies that if the "actual" state ω is not in the sink of any of the agents other than the dummy's, then it must at least be in the dummy's sink. In such a case, the dummy would be somewhat of a "wise fool" in the sense that all other agents would be deeming ω impossible, whereas the dummy does not rule out *any* possibility, including ω itself. This implication does not necessarily hold when the doxastic dummy assumption is taken as it is originally stated.

As before, the weaker assumption, Dummy*, would have sufficed for the above theorem to hold.

One rather worrying feature of Theorem 2, however, can be illustrated by the following example. Consider model $\mathcal{M}$ in Figure 1 with $\omega \models p$ and $\omega' \models \neg p$. At every state of this model, i believes that $\neg p$ and at every state, j believes that p. In this model, the condition of "heterogeneity" fails, so all the agreement theorems mentioned in the introduction would concede that i and j can agree to disagree (see Tarbush (2011)).[10] Now, consider adding an epistemic dummy h to this model, to obtain model $\mathcal{M}'$. Heterogeneity would again fail, so the agents can again agree to disagree according to all the agreement theorems other than Samet's. However, according to Theorem 2, the agents cannot agree to disagree. But what drives the result in this case?

Agent i must surely perform his action as though he were certain that $\neg p$ is the case, since $\neg p$ is the only proposition that i believes, regardless of the state.

[10] For all $i \in G$, if for all $\omega' \in \Omega_G(\omega)$, we have $\omega \models \nu_i$, $\omega' \models \mu_i$ and $\nu_i = \mu_i$, then for all $\omega' \in \Omega_G(\omega)$, with $\omega' \models \nu_i' \wedge \nu_j'$ we have $\nu_i' \sim \nu_j'$.

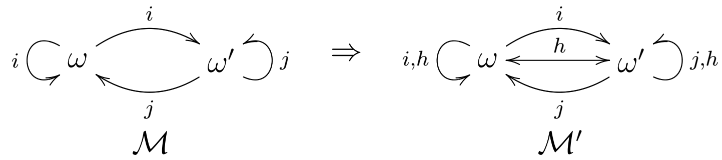

Figure 1: $\mathcal{M}' = \mathcal{M}$ plus dummy h

Similarly, agent j must surely perform her action as though she were certain that p is the case. However, by the presence of h, the agents i and j must perform the same action. So the existence of the dummy must collapse the action that one would perform when p and when $\neg p$ to the *same* action.
One can interpret this in one of two ways: (i) The existence of the dummy can be seen as a constraint on the decision functions, requiring them to be independent of one's information regarding p. But this then makes agreement trivial. Or, (ii) the decision functions do depend on p, but the existence of the dummy implies that the more informed agents must nevertheless perform the same action. However, this must be the action that the agents would perform when they do not "know" whether p is true, *even though*, in this example, the more informed agents are effectively *certain* of their information regarding p.

We can now provide a further theorem without a doxastic dummy.

Definition 11. Condition (1.a): For all $\omega \in \Omega$ and $i, j \in G$, there exists an $\omega' \in \Omega_G(\omega)$ such that $S_i(\omega') = S_j(\omega')$. Condition (1.b): For all $\omega \in \Omega$ and $i, j \in G$, there exists an $\omega' \in \Omega_G(\omega)$ such that $\omega' \models \nu_i \wedge \nu_j$ for some ν_i, ν_j, such that $\nu_i \sim \nu_j$.[11]

Condition (1.a) states that in any component, there must exist a state in which both agents have the same sink. Syntactically, condition (1.b) states that it must not be commonly believed, among the agents, that they do not have the same information. This condition can be seen as a requirement that there be a "grain of agreement" among the agents - in the sense that there must exist some state within each component in which the agents have the same information.
By Lemma 5, one can verify that condition (1.a) implies (1.b), however the converse does not hold.

Theorem 3. *Suppose that ISTP and condition (1.a) hold, and that the system is KD45. Let $G = \{i, j\}$. Then,* $\models C_G(d_i^x \wedge d_j^y) \to (x = y)$.

[11]These conditions are also discussed in Tarbush (2011).

As a result of the assumption used, this theorem does *not* apply in the model $\mathcal{M}'$ represented in Figure 1.

Furthermore, the result could be extended to more than two agents. We would simply have to require the existence of a "grain of agreement" between pairs of agents, enough to cover all agents. For example, if there are three agents i, j and k, then we would need condition (1.a) to hold between say i and j and between j and k - but we would not require for it to hold between i and k for example.

In terms of limitations , it should be noted that condition (1.b) would not be sufficient for the result. Also, the result does also apply in $S5$ - if we replace sinks by cells in the statement of condition (1.a) - however, in such frames, it only generates agreement in a trivial sense. Indeed, one can verify that condition (1.a), with sinks replaced by cells, implies that all agents have the same information at every state within a component, which renders agreement trivial.

Appendix

Proof of Lemma 1 (i) Consider an arbitrary $i \in N$ and $\omega \in \Omega$, and suppose that $\omega \models \psi$, for some formula $\psi \in \Psi_0^r$. It must be the case that either (i.a) $\forall \omega' \in \Omega$, if $\omega R_i \omega'$ then $\omega' \models \psi$, or (i.b) $\forall \omega' \in \Omega$, if $\omega R_i \omega'$ then $\omega' \models \neg\psi$, or (i.c) $\exists \omega', \omega'' \in \Omega$, such that $\omega R_i \omega'$ and $\omega R_i \omega''$, and $\omega' \models \psi$ and $\omega'' \models \neg\psi$ (i.e. neither (i.a) nor (i.b)). If (i.a) is the case, then $\omega \models \Box_i \psi$. If (i.b) is the case, then $\omega \models \hat{\Box}_i \psi$, and finally, if (i.c) is the case, then $\omega \models \dot{\Box}_i \psi$. Therefore, in all cases, the operator over ψ belongs to the set O_i, and since this holds for any $\psi \in \Psi_0^r$, it holds for each entry of a ken. Furthermore, $\models$ can only generate consistent lists of formulas, so kens cannot be inconsistent. This implies that a ken must exist that belongs to V_i.
(ii) Consider an arbitrary $i \in N$ and $\omega \in \Omega$. Let $\nu_i, \mu_i \in V_i$, and consider the n^{th} entry of each ken such that $\nu_i^n \psi_n \neq \mu_i^n \psi_n$. Case (ii.a): Suppose $\omega \models \nu_i^n \psi_n = \Box_i \psi_n$. So, $\forall \omega' \in \Omega$, if $\omega R_i \omega'$, then $\omega' \models \psi_n$. By definition, this rules out the possibility that also, $\omega \models \hat{\Box}_i \psi_n$, or $\omega \models \dot{\Box}_i \psi_n$. For cases (ii.b), $\omega \models \nu_i^n \psi_n = \hat{\Box}_i \psi_n$, and (ii.c), $\omega \models \nu_i^n \psi_n = \dot{\Box}_i \psi_n$, proceed analogously to (ii.a).

Proof of Lemma 2 Consider some arbitrary state $\omega \in \Omega$. Suppose $I_i(\omega) \subseteq I_j(\omega)$ and $\omega \models \nu_i \wedge \nu_j$. Consider the nth entry of these kens.
(a) Suppose $\omega \models \nu_i^n \psi_n = \Box_i \psi_n$, and suppose that $\omega \models \nu_j^n \psi_n = \hat{\Box}_j \psi_n$. Then, $\forall \omega' \in I_j(\omega)$, $\omega' \models \neg\psi_n$. But if $I_i(\omega) \subseteq I_j(\omega)$, then $\forall \omega' \in I_i(\omega)$, $\omega' \models \neg\psi_n$, which contradicts the statement that $\omega \models \Box_i \psi_n$. Therefore, $\omega \models (\nu_j^n \psi_n = \Box_j \psi_n \vee \nu_j^n \psi_n = \dot{\Box}_j \psi_n)$.
Cases (b), $\omega \models \nu_i^n \psi_n = \hat{\Box}_i \psi_n$ and (c) $\omega \models \nu_i^n \psi_n = \dot{\Box}_i \psi_n$ can dealt with analogously to case (a).

Proof of Lemma 3 Suppose $\omega'' \in \bigcup_{\omega' \in \Omega_G(\omega)} I_i(\omega')$. So, $\omega'' \in I_i(\omega')$ for some $\omega' \in \Omega_G(\omega)$. But, $\omega' R_i \omega''$, and there exists a sequence of R_i ($i \in G$) steps such that ω' is reachable from ω. Therefore, there exists a sequence, one step longer, such that ω'' is reachable from ω. So, $\omega'' \in \Omega_G(\omega)$. (And, note that $I_i(\omega'') \subseteq \Omega_G(\omega)$).
Suppose $\omega'' \in \Omega_G(\omega)$. Reflexivity guarantees that $\omega'' \in I_i(\omega'')$. So, for some $\omega^* \in \Omega_G(\omega)$, $\omega'' \in I_i(\omega^*)$, so $\omega'' \in \bigcup_{\omega' \in \Omega_G(\omega)} I_i(\omega')$.

Proof of Lemma 4 Suppose $\omega' \models \nu_i$ for some $\omega' \in I_i(\omega)$. Consider the n^{th} entry of the ken, namely, $\nu_i^n \psi_n$.
(a) Suppose $\omega' \models \nu_i^n \psi_n = \Box_i \psi_n$. Then, for all $\omega'' \in \Omega$, $\omega' R_i \omega''$ implies $\omega'' \models \psi_n$. So, for all $\omega'' \in I_i(\omega')$, $\omega'' \models \psi_n$. But since R_i is an equivalence relation, and $\omega' \in I_i(\omega)$, it follows that $I_i(\omega') = I_i(\omega)$. So, for all $\omega'' \in I_i(\omega)$, $\omega'' \models \psi_n$, from which it follows that for all $\omega'' \in I_i(\omega)$, $\omega'' \models \Box_i \psi_n$.
Case (b), $\omega' \models \nu_i^n \psi_n = \hat{\Box}_i \psi_n$ and (c), $\omega' \models \nu_i^n \psi_n = \dot{\Box}_i \psi_n$ are analogous to case (a).

Proof of Theorem 1 Suppose that there exists an epistemic dummy, $ISTP$ holds, and that the system is $S5$. Let $\omega \in \Omega$, and consider the set $\Omega_G(\omega)$. It must be the case that at ω, $\omega \models \nu_h$ for some ν_h. So by Lemma 4 and the existence of an epistemic dummy, for all $\omega' \in \Omega_G(\omega)$, $\omega' \models \nu_h$. By Lemma 3, we know that $\bigcup_{\omega' \in \Omega_G(\omega)} I_i(\omega') = \Omega_G(\omega)$. So for each state in each of i's information cells, and therefore for each $\omega'' \in \Omega_G(\omega)$ with $\omega'' \models \nu_i$ for some ν_i, it must be the case that $\nu_i \succsim \nu_h$ by Lemma 2. This is true at every state in the component, so in particular, if $\omega \models \nu_i'$ then $\omega \models \Box_h(\nu_i' \rightarrow \nu_i' \succsim \nu_h)$. (Agent h knows $\nu_i' \rightarrow \nu_i' \succsim \nu_h$ because it must be true at every state of the information cell $I_h(\omega)$).
Finally, by the assumption that $\omega \models C_G(d_i^x)$, it follows that $\omega \models \Box_h(d_i^x)$. By $ISTP$, it follows that $\omega \models d_h^x$.
Reasoning similarly, between the dummy h and agent j, we find that $\omega \models d_h^y$. Therefore $\omega \models d_h^x \wedge d_h^y$, which is not possible unless $x = y$.

Proof of Proposition 1 Let "i-arrow" refer to an arrow of i's accessibility relation. Firstly, we can show that $R_i = E_i \cup F_i$. An arbitrary $\omega \in \Omega$ either has an i-arrow pointing to it or it does not. If it does not, by seriality, it points to another state. If it does, then there exists a state ω' that points to ω which itself points to some state ω'' by seriality. Transitivity implies that ω' points to ω'' and Euclideaness implies that ω'' points to ω. From here it is easy to prove that ω, ω' and ω'' are in an equivalence class.
Secondly, we show that if $J_i(\omega') \neq J_i(\omega'')$ then $J_i(\omega') \cap J_i(\omega'') = \emptyset$. Suppose $\omega \in J_i(\omega') \cap J_i(\omega'')$. If $\omega \in S_i(\omega') \cap S_i(\omega'')$ then $S_i(\omega')$ and $S_i(\omega'')$ are indistinguishable, and one can verify that $J_i(\omega') = J_i(\omega'')$. If $\omega \in S_i(\omega') \cap A_i(\omega'')$ then ω both does have and does not have an i-arrow pointing to it. Finally, if $\omega \in A_i(\omega') \cap A_i(\omega'')$ then by Euclideaness, ω' and ω'' are indistinguishable, and $J_i(\omega') = J_i(\omega'')$.
Thirdly, we can show that $\cup_{\omega \in \Omega} J_i(\omega) = \Omega$. Suppose $\omega' \in \cup_{\omega \in \Omega} J_i(\omega)$, then by the definitions of S_i and A_i, $\omega' \in \Omega$. On the other hand, suppose $\omega \in \Omega$. Then if there is an i-arrow pointing to ω, $\omega \in S_i(\omega) \subseteq J_i(\omega)$. If there is no i-arrow pointing to it, then by seriality, there is an ω' that ω points to, so $\omega \in A_i(\omega') \subseteq J_i(\omega')$. So, $\omega \in \cup_{\omega \in \Omega} J_i(\omega)$.

Proof of Lemma 5 Entirely analogous to the proof of Lemma 2.

Proof of Lemma 6 Suppose $\omega'' \in \bigcup_{\omega' \in \Omega_G(\omega)} S_i(\omega')$. So, $\omega'' \in S_i(\omega')$ for some $\omega' \in \Omega_G(\omega)$. But, $\omega' E_i \omega''$, and there exists a sequence of R_i ($i \in G$) steps such that ω' is reachable from ω. Therefore, there exists a sequence, one step longer, such that ω'' is reachable from ω. So, $\omega'' \in \Omega_G(\omega)$.

Suppose $\omega'' \in \Omega_G(\omega)$. Either ω'' has an i-arrow pointing towards it, in which case $\omega'' \in S_i(\omega'')$. So, $\omega'' \in S_i(\omega'') \cup A_i(\omega'') = J_i(\omega'')$, or, ω'' has no i-arrow pointing towards it, in which case, by seriality, there exists some ω''' such that $\omega'' \in A_i(\omega''')$. Note that ω''' must be in $\Omega_G(\omega)$ since it is reachable from ω''. So, $\omega'' \in S_i(\omega''') \cup A_i(\omega''') = J_i(\omega''')$. In either case, for some $\omega^* \in \Omega_G(\omega)$, $\omega'' \in J_i(\omega^*)$, so $\omega'' \in \bigcup_{\omega' \in \Omega_G(\omega)} J_i(\omega')$.

Proof of Lemma 7 Suppose $\omega' \models \nu_i$ for some $\omega' \in J_i(\omega)$. Firstly, suppose $\omega' \in S_i(\omega)$, and consider the n^{th} entry of the ken, namely, $\nu_i^n \psi_n$.
(a) Suppose $\omega' \models \nu_i^n \psi_n = \Box_i \psi_n$. Then, for all $\omega'' \in \Omega$, $\omega' E_i \omega''$ implies $\omega'' \models \psi_n$. So, for all $\omega'' \in S_i(\omega')$, $\omega'' \models \psi_n$. But since E_i is an equivalence relation, and $\omega' \in S_i(\omega)$, it follows that $S_i(\omega') = S_i(\omega)$. So, for all $\omega'' \in S_i(\omega)$, $\omega'' \models \psi_n$, from which it follows that for all $\omega'' \in S_i(\omega)$, $\omega'' \models \Box_i \psi_n$. Also, each $\omega''' \in A_i(\omega)$ has an arrow pointing to each state in $S_i(\omega)$, so for all $\omega^* \in S_i(\omega)$, if $\omega''' F_i \omega^*$, $\omega^* \models \psi_n$. So, for all $\omega''' \in A_i(\omega)$, $\omega''' \models \Box_i \psi_n$. It follows that for all $\omega'' \in J_i(\omega)$, $\omega'' \models \Box_i \psi_n$.
Case (b), $\omega' \models \nu_i^n \psi_n = \hat{\Box}_i \psi_n$ and (c), $\omega' \models \nu_i^n \psi_n = \dot{\Box}_i \psi_n$ are analogous to case (a).
Now, suppose $\omega' \in A_i(\omega)$, and consider the n^{th} entry of the ken, namely, $\nu_i^n \psi_n$.
(d) Suppose $\omega' \models \nu_i^n \psi_n = \Box_i \psi_n$. Then, for all $\omega'' \in \Omega$, $\omega' F_i \omega''$ implies $\omega'' \models \psi_n$. So, for all $\omega'' \in S_i(\omega')$, $\omega'' \models \psi_n$. This implies that $\omega'' \models \Box_i \psi_n$ for all $\omega'' \in S_i(\omega)$, and $\omega''' \models \Box_i \psi_n$ for all other states $\omega''' \in A_i(\omega)$. It follows that for all $\omega'' \in J_i(\omega)$, $\omega'' \models \Box_i \psi_n$.
Case (e), $\omega' \models \nu_i^n \psi_n = \hat{\Box}_i \psi_n$ and (f), $\omega' \models \nu_i^n \psi_n = \dot{\Box}_i \psi_n$ are analogous to case (d).

Proof of Theorem 2 Suppose that there exists a doxastic dummy, $ISTP$ holds, and that the system is $KD45$. Let $\omega \in \Omega$, and consider the set $\Omega_G(\omega)$. It must be the case that at ω, $\omega \models \nu_h$ for some ν_h. So by Lemma 7 and the existence of a doxastic dummy, for all $\omega' \in \Omega_G(\omega) \cup \{\omega\}$, $\omega' \models \nu_h$. By Lemma 6, we know that $\bigcup_{\omega' \in \Omega_G(\omega)} S_i(\omega') \subseteq \Omega_G(\omega)$. So for each state ω'' in each of i's information sinks, it must be the case that $\omega'' \models \nu_i''$ for some ν_i'', and that $\nu_i'' \succsim \nu_h$ by Lemma 5. However, this must also be true at every state ω''' that is in the component but not in any of i's sinks (by Lemma 7); that is, at every state in $\Omega_G(\omega) \cup \{\omega\}$. So, in particular, if $\omega \models \nu_i'$ then $\omega \models \Box_h(\nu_i' \rightarrow \nu_i' \succsim \nu_h)$. (Agent h knows $\nu_i' \rightarrow \nu_i' \succsim \nu_h$ because it must be true at every state of the information sink and associated state $J_h(\omega)$).
Finally, by the assumption that $\omega \models C_G(d_i^x)$, it follows that $\omega \models \Box_h(d_i^x)$. By $ISTP$, it follows that $\omega \models d_h^x$.
Reasoning similarly, between the dummy h and agent j, we find that $\omega \models d_h^y$. Therefore $\omega \models d_h^x \wedge d_h^y$, which is not possible unless $x = y$.

Proof of Theorem 3 If condition (1.a) holds, then without loss of generality, there is some state $\omega' \in \Omega_G(\omega)$ such that $\omega' \models \nu'_i \wedge \mu'_j$ for some ν'_i, μ'_j and $\mu'_j \succsim \nu'_i$ by Lemma 5, and $\omega' \models \Box_i \mu'_j$. The latter is true because μ'_j is invariant across the sink $S_j(\omega')$, but this is equal to $S_i(\omega')$, so is invariant across this sink as well. Now, suppose that $\omega \models C_G(d_i^x \wedge d_j^y)$. Firstly, $d_i^x \wedge d_j^y$ must be true at any state ω'' such that $\omega R_i \omega''$, and since agent i's kens are the same at ω and ω'' (Lemma 7), the actions must also be the same at those states, so (reasoning similarly for j) we have $\omega \models d_i^x \wedge d_j^y$. Secondly, at ω' itself, we have $\omega' \models \Box_i(d_j^y)$. By $ISTP$, we have that $\omega' \models d_i^y$. But by common belief, $\omega' \models d_i^x$ since it is also reachable from ω. So $\omega' \models d_i^x \wedge d_i^y$, which is not possible unless $x = y$; which also holds at ω itself.

References

Aumann, R. J.: 1976, 'Agreeing to disagree'. *The annals of statistics* **4**(6), 1236–1239.

Aumann, R. J. and S. Hart: 2006, 'Agreeing on decisions, The Hebrew University of Jerusalem'. *Center for Rationality, unpublished DP.*

Bacharach, M.: 1985, 'Some extensions of a claim of Aumann in an axiomatic model of knowledge'. *Journal of Economic Theory* **37**(1), 167–190.

Bonanno, G. and K. Nehring: 1998, 'Assessing the Truth Axiom under incomplete information'. *Mathematical social sciences* **36**(1), 3–29.

Cave, J. A.: 1983, 'Learning to agree'. *Economics Letters* **12**(2), 147–152.

Chellas, B. F.: 1980, *Modal logic.* Cambridge University Press Cambridge.

Hintikka, J.: 1962, *Knowledge and belief: an introduction to the logic of the two notions.* Cornell University Press.

Moses, Y. and G. Nachum: 1990, 'Agreeing to disagree after all'. In: *Proceedings of the 3rd conference on Theoretical aspects of reasoning about knowledge.* pp. 151–168.

Samet, D.: 2010, 'Agreeing to disagree: The non-probabilistic case'. *Games and Economic Behavior* **69**(1), 169–174.

Tarbush, B.: 2011, 'Agreeing to disagree: a syntactic approach'. *University of Oxford, Department of Economics, Working Paper.*

van Benthem, J.: 2010, 'Modal Logic for Open Minds'. *Center for the Study of Language and Information.*