



Munich Personal RePEc Archive

**Agreement theorems with interactive  
information: possibilities and  
impossibilities**

Tarbush, Bassel

17 August 2011

Online at <https://mpra.ub.uni-muenchen.de/32850/>  
MPRA Paper No. 32850, posted 16 Aug 2011 23:52 UTC

# Agreement theorems with interactive information: possibilities and impossibilities

Bassel Tarbush<sup>0</sup>

August 17, 2011

**Abstract** Following from Tarbush (2011a), we explore the implications of using two different definitions of informativeness over kens; one that ranks objective, and the other subjective information. With the first, we create a new semantic operation that allows us to derive agreement theorems even when decision functions are based on interactive information (for any  $r \geq 0$ ). Effectively, this operation, unlike information cell union captures the notion of an agent becoming “more ignorant” *for all modal depths*. Using the definition that ranks subjective information however, we show an impossibility result: In generic models, agreement theorems using the standard Sure-Thing Principle do not hold when decision functions depend on interactive information (when  $r > 0$ ).

**Keywords** Agreeing to disagree, knowledge, common knowledge, belief, information, epistemic logic.

**JEL classification** D80, D83, D89.

## 1 Introduction

Bacharach (1985) proved an agreement theorem showing that if the agents of like-minded agents are common knowledge, then their actions must be the same. This result is proven with general decision functions,  $D_i : \mathcal{F} \rightarrow \mathcal{A}$ , that map from a field  $\mathcal{F}$  of subsets of  $\Omega$  into an arbitrary set  $\mathcal{A}$  of *actions*. To derive the result, it is assumed that agents have the same decision function (termed “like-mindedness”), and that the decision functions satisfy what we call the Disjoint Sure-Thing Principle (*DSTP*):  $\forall E \in \mathcal{E}$ , if  $D_i(E) = x$  then  $D_i(\cup_{E \in \mathcal{E}} E) = x$ , where  $\mathcal{E}$  is a set of disjoint events. The idea is that if one decides to perform action  $x$  in each case

---

<sup>0</sup>Department of Economics, University of Oxford, [bassel.tarbush\[at\]economics.ox.ac.uk](mailto:bassel.tarbush[at]economics.ox.ac.uk)

when one is more informed (when  $E$  and when  $F$ ), then one must also decide to perform action  $x$  when one is more ignorant (when  $E \cup F$ ). Moses and Nachum (1990) show that taking the union of information cells is conceptually problematic because it does not adequately capture of being “more ignorant”.

Tarbush (2011a) shows that the union of information cells does capture the notion of being “more ignorant”, but only for non-interactive information (see Lemma 6), and thus proposes a solution to the conceptual problem by providing explicit accounts of the information that each agent has at each state, and then truncating the information so that decisions only depend on non-interactive information. Effectively, the paper preserves a similar semantic operation to cell union and bypasses the problem by focussing on the non-problematic information that remains after the operation.

It is not shown that truncating the information is *necessary* to obtain the agreement result, only that it is sufficient - given that the semantic operation is cell union. However, this leaves open the possibility that agreement results hold for non-truncated, that is, interactive information, if an operation other than cell union is used - which is what we explore here.

So, in this paper, we define a new operation that allows us to talk about agents becoming “more ignorant” *even* when considering interactive information. The upshot is that Bacharach’s result does hold, but requires a different operation from cell union; and indeed, any paper that resorts to cell union to capture more ignorance in an interactive agent setting should really be using the operation proposed here.

Note that all the notation and definitions referred to in the rest of this paper are drawn from Tarbush (2011a).

To carry out the analysis regarding interactive information, we need to be careful how we rank information at the interactive level: In Tarbush (2011a) a notion of informativeness was defined in order to rank kens, both for a single agent, and across agents. Here, we show that we can define at least two different versions of this ranking; and which one we decide to use has far-reaching implications for the results we eventually obtain.

The following is the definition of informativeness given in Tarbush (2011a):

**Definition 1** (Informativeness). Create an order  $\succ \subseteq V_i \times V_j$  for all  $i, j \in N$ . We say that the ken  $\nu_i$  is *more informative* than the ken  $\mu_j$ , denoted  $\nu_i \succ \mu_j$ , if and only if whenever  $i$  knows that  $\psi$  then  $j$  either also knows that  $\psi$  or does not know whether  $\psi$ , and whenever  $i$  does not know whether  $\psi$ , then  $j$  also does not know whether  $\psi$ . Formally,

- If  $\nu_i^n \psi_n = \square_i \psi_n$  then  $(\mu_j^n \psi_n = \square_j \psi_n \text{ or } \mu_j^n \psi_n = \dot{\square}_j \psi_n)$ ,
- If  $\nu_i^n \psi_n = \hat{\square}_i \psi_n$  then  $(\mu_j^n \psi_n = \hat{\square}_j \psi_n \text{ or } \mu_j^n \psi_n = \dot{\square}_j \psi_n)$ , and
- If  $\nu_i^n \psi_n = \dot{\square}_i \psi_n$  then  $(\mu_j^n \psi_n = \dot{\square}_j \psi_n)$ .

Finally, note that  $\nu_i \sim \mu_j$  denotes  $\nu_i \succsim \mu_j$  and  $\mu_j \succsim \nu_i$ ; which is interpreted as  $\nu_i$  and  $\mu_j$  carrying the *same* information, but seen from the perspectives of agents  $i$  and  $j$  respectively.

This means that if the  $n^{\text{th}}$  entry of  $\nu_i$  is  $\square_i \square_j p$ , and  $\nu_i \sim \mu_j$  then the  $n^{\text{th}}$  entry of  $\mu_j$  is  $\square_j \square_i p$ . This notion, in a sense, captures a comparison about what information the agents *objectively* have about the world: Here, we are asking what each of the agents know about  $j$ 's knowledge about  $p$ .

Contrast this with the alternative definition of informativeness given below.

**Definition 2** (Informativeness\*). Create an order  $\succsim^* \subseteq V_i \times V_j$  for all  $i, j \in N$ , defined by:

- If  $\nu_i^n \psi_n = \square_i \psi_n$  then  $(\mu_j^{\sigma(n)} \psi_{\sigma(n)} = \square_j \psi_{\sigma(n)} \text{ or } \mu_j^{\sigma(n)} \psi_{\sigma(n)} = \dot{\square}_j \psi_{\sigma(n)})$ ,
- If  $\nu_i^n \psi_n = \hat{\square}_i \psi_n$  then  $(\mu_j^{\sigma(n)} \psi_{\sigma(n)} = \hat{\square}_j \psi_{\sigma(n)} \text{ or } \mu_j^{\sigma(n)} \psi_{\sigma(n)} = \dot{\square}_j \psi_{\sigma(n)})$ , and
- If  $\nu_i^n \psi_n = \dot{\square}_i \psi_n$  then  $(\mu_j^{\sigma(n)} \psi_{\sigma(n)} = \dot{\square}_j \psi_{\sigma(n)})$ .

This definition is applicable if there are only two agents in the model.  $\sigma$  is a bijection. If  $\psi_n$  is an atomic formula,  $\sigma$  maps  $n$  into itself. On the other hand, if  $\psi_n$  has modal depth one or greater, then  $\psi_{\sigma(n)}$  is identical to  $\psi_n$  except that all the  $i$  and  $j$  indices are switched.

Now, this means that if the  $n^{\text{th}}$  entry of  $\nu_i$  is  $\square_i \square_j p$ , and  $\nu_i \sim^* \mu_j$  then the  $\sigma(n)^{\text{th}}$  entry of  $\mu_j$  is  $\square_j \square_i p$ . This notion captures a comparison about the *subjective* states of minds of the agents: Here, for example, we are saying that  $i$  knows that  $j$  knows that  $p$ , and we say that  $j$  is in a “similar” state of mind if we can simply switch their indices around - that is, if they can somehow stand in each other’s shoes, in which case we would say that  $j$  knows that  $i$  knows that  $p$ .

Note that both orderings are identical over atomic propositions, so all results in Tarbush (2011a) that imposed  $r = 0$  are unchanged by our choice of definition for informativeness. Also, regardless of which definition of informativeness we use, our definition of the infimum of kens is unchanged, since it is only applicable to the kens of a single agent. The infimum of  $\nu_i$  and  $\mu_i$ , denoted  $\inf\{\nu_i, \mu_i\}$ , is the most informative ken that is less informative than  $\nu_i$  and  $\mu_i$ .

**Lemma 1.** For any  $\nu_i, \mu_i \in V_i$ ,  $\inf\{\nu_i, \mu_i\}$  exists in  $V_i$  and is characterised by:

$$\inf\{\nu_i, \mu_i\}^n \psi_n = \square_i \psi_n \text{ iff } (\nu_i^n \psi_n = \mu_i^n \psi_n = \square_i \psi_n)$$

$$\inf\{\nu_i, \mu_i\}^n \psi_n = \hat{\square}_i \psi_n \text{ iff } (\nu_i^n \psi_n = \mu_i^n \psi_n = \hat{\square}_i \psi_n)$$

$$\inf\{\nu_i, \mu_i\}^n \psi_n = \dot{\square}_i \psi_n \text{ iff } (\nu_i^n \psi_n \neq \mu_i^n \psi_n \text{ or } \nu_i^n \psi_n = \mu_i^n \psi_n = \dot{\square}_i \psi_n)$$

In this paper, we show that with the ordering  $\succsim$ , all the previous results found in Tarbush (2011a) and Tarbush (2011b) still hold. In fact, they can be extended: Unlike the *cell merge* and *sink merge* operations defined in Tarbush (2011a), in section 2, we show that there is an operation that recovers the infimum of kens in a component for all agents and for all  $r \geq 0$ , and this allows us to derive agreement theorems that hold for all  $r \geq 0$  - in contrast with our previous results that only held for  $r = 0$ .

However, if we had used the ordering  $\succsim^*$  then many of the results in Tarbush (2011a) and Tarbush (2011b) would only hold for  $r = 0$ . In particular, the results would all be unchanged in Tarbush (2011a), but Proposition 3 would only hold for  $r = 0$ . Furthermore, in Tarbush (2011b), Lemmas 2 and 5, and therefore all the theorems in that paper, would only hold for  $r = 0$ .

In fact, with  $\succsim^*$ , we demonstrate a rather stark result in section 3: In generic models, agreement theorems (using the standard versions of the Sure-Thing Principle) do not hold when decisions are based on interactive information (that is, when  $r > 0$ ).

## 2 New operation

Consider an *S5* model  $\mathcal{M} = \langle \Omega, R_i, \mathcal{V} \rangle$ . Now create a new model  $\mathcal{M}^+ = \langle Q, Z_i, \mathcal{T} \rangle$ , where:

$$Q = \Omega \cup \{s\} \tag{1}$$

$$Z_i = R_i \cup \{(s, \omega) | \omega \in \Omega\} \tag{2}$$

$$\mathcal{T} = \mathcal{V} \tag{3}$$

This means that  $\mathcal{M}^+$  is the same model as  $\mathcal{M}$  except that it contains one extra state,  $s$ , and for each agent, there are arrows pointing from  $s$  to *every* other state. The valuation maps for the atomic propositions are identical in both models, and the atomic propositions that are true at  $s$  itself is irrelevant since  $s$  does not point to itself.<sup>1</sup>

---

<sup>1</sup>An analogous operation exists in *KD45* where  $s$  has  $i$ -arrows only pointing to states that are in information sinks for agent  $i$ .

We will show below that the ken of every agent at the state  $s$  is the infimum of all the kens at all other states in the model, for all  $r > 0$ .<sup>2</sup>

**Proposition 1.** *For all  $i \in N$ ,  $\mathcal{M}^+, s \models \inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\}$  (for any  $r \geq 0$ ).*

*Proof.* Suppose we are working in  $S5$ . Consider agent  $i$ , and any formula  $\psi$  of some modal depth  $r > 0$ . If  $\psi$  is true in every one of  $i$ 's information cells in the model  $\mathcal{M}$ , then clearly,  $s \models \Box_i \psi$ . If  $\psi$  is true at some states and false in others then  $s \models \dot{\Box}_i \psi$ . That is, whenever  $i$  knows that  $\psi$  in model  $\mathcal{M}$ , this information is preserved in  $s$ ; and in all other cases - when  $i$  does not know whether  $\psi$  or has contradictory information about  $\psi$  in different information cells of the model  $\mathcal{M}$ , then  $i$  does not know whether  $\psi$  at the state  $s$ . This is precisely the characterisation of the infimum, and holds for any  $r$ . (The proof is identical for  $KD45$ , but being careful to talk about information sinks rather than information cells).  $\square$

The above suggests that in order to derive his agreement theorem, Bacharach (1985) should have used this operation rather than the union of information cells.

Consider the following characterisation of a class of models, which we denote by  $\mathbf{M}$ :

- If the modal operator of the  $n^{\text{th}}$  entry of *every* ken of some agent  $i$  is  $\Box_i$ , then the modal operator of the  $n^{\text{th}}$  entry of *every* ken of all other agents  $j$  is also  $\Box_j$ .
- If the modal operator of the  $n^{\text{th}}$  entry of *every* ken of some agent  $i$  is  $\hat{\Box}_i$ , then the modal operator of the  $n^{\text{th}}$  entry of *every* ken of all other agents  $j$  is also  $\hat{\Box}_j$ .
- If the modal operators of the  $n^{\text{th}}$  entry of different kens of some agent  $i$  are different, or if the modal operator of the  $n^{\text{th}}$  entry of *every* ken of the agent is  $\dot{\Box}_i$ , then *either* the modal operators of the  $n^{\text{th}}$  entry of different kens are also different for all other agents  $j$ , *or*, the modal operator of the  $n^{\text{th}}$  entry of *every* ken of all other agents  $j$  is  $\dot{\Box}_j$ .

**Proposition 2.**  $\mathcal{M}^+, s \models \inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\} \wedge \inf\{\nu_j | \omega \in \Omega \ \& \ \omega \models \nu_j\}$  and  $\inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\} \sim \inf\{\nu_j | \omega \in \Omega \ \& \ \omega \models \nu_j\}$  if and only if  $\mathcal{M} \in \mathbf{M}$ .

---

<sup>2</sup>Naturally, for the purposes of using this operation for the proof of an agreement theorem, we would have to redefine the operation to work not at the level of the entire model, but rather component by component. This is trivial, but the complications that this would introduce in the notation here might blur the main point.

*Proof.* Suppose  $\mathcal{M}^+, s \models \inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\} \wedge \inf\{\nu_j | \omega \in \Omega \ \& \ \omega \models \nu_j\}$  and  $\inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\} \sim \inf\{\nu_j | \omega \in \Omega \ \& \ \omega \models \nu_j\}$ .

Now, suppose the modal operator of the  $n^{\text{th}}$  entry of *every* ken of some agent  $i$  is  $\square_i$ . Then, the modal operator of the  $n^{\text{th}}$  entry of  $i$ 's ken at state  $s$  is also  $\square_i$ . If it is not the case that the modal operator of the  $n^{\text{th}}$  entry of *every* ken of all other agents  $j$  is also  $\square_j$ , then the modal operator of the  $n^{\text{th}}$  entry of some agent  $j$ 's ken at state  $s$  is not  $\square_j$ ; thus contradicting our initial assumption that all agents have the “same” ken at state  $s$ . Proceed similarly for the other cases.

The other direction is trivial. □

One can easily verify that all *S5* models are in fact models in the class **M**. Therefore, we can use the above proposition instead of Lemma 6 in Tarbush (2011a) to derive Theorems 1 and 2; but now they would hold for all  $r \geq 0$ , as stated below.<sup>3</sup>

**Theorem 1.** *Consider  $\Psi_0^r$  for any  $r \geq 0$ , the agents are like-minded, and the system is *S5*. Let  $G = \{i, j\} \subseteq N$ .*

*If either, *NDSTP* holds; or, *DSTP* holds and the language is rich in every component, then  $\models C_G(d_i^x \wedge d_j^y) \rightarrow (x = y)$ .*

*Proof.* See proof of Theorems 1 and 2 in Tarbush (2011a) but replace Lemma 6 with Proposition 2 here. □

### 3 Impossibility in generic models

Consider the following characterisation of a class of two agent models, which we denote by **M\***:

- If the modal operator of the  $n^{\text{th}}$  entry of *every* ken of some agent  $i$  is  $\square_i$ , then the modal operator of the  $\sigma(n)^{\text{th}}$  entry of *every* ken of agent  $j$  is also  $\square_j$ .
- If the modal operator of the  $n^{\text{th}}$  entry of *every* ken of some agent  $i$  is  $\hat{\square}_i$ , then the modal operator of the  $\sigma(n)^{\text{th}}$  entry of *every* ken of agent  $j$  is also  $\hat{\square}_j$ .
- If the modal operators of the  $n^{\text{th}}$  entry of different kens of some agent  $i$  are different, or if the modal operator of the  $n^{\text{th}}$  entry of *every* ken of the agent is  $\square_i$ , then *either* the modal operators of the  $\sigma(n)^{\text{th}}$  entry of different kens are also different for agent  $j$ , *or*, the modal operator of the  $\sigma(n)^{\text{th}}$  entry of *every* ken of agent  $j$  is  $\square_j$ .

---

<sup>3</sup>With appropriate modifications, analogous results should hold for *KD45* as well.

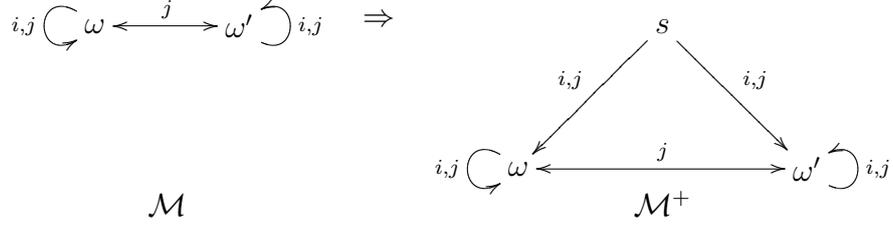


Figure 1: Impossibility with  $r > 0$

**Proposition 3.**  $\mathcal{M}^+, s \models \inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\} \wedge \inf\{\nu_j | \omega \in \Omega \ \& \ \omega \models \nu_j\}$  and  $\inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\} \sim^* \inf\{\nu_j | \omega \in \Omega \ \& \ \omega \models \nu_j\}$  if and only if  $\mathcal{M} \in \mathbf{M}^*$ .

*Proof.* Suppose  $\mathcal{M}^+, s \models \inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\} \wedge \inf\{\nu_j | \omega \in \Omega \ \& \ \omega \models \nu_j\}$  and  $\inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\} \sim^* \inf\{\nu_j | \omega \in \Omega \ \& \ \omega \models \nu_j\}$ .

Now, suppose the modal operator of the  $n^{\text{th}}$  entry of *every* ken of some agent  $i$  is  $\square_i$ . Then, the modal operator of the  $n^{\text{th}}$  entry of  $i$ 's ken at state  $s$  is also  $\square_i$ . If it is not the case that the modal operator of the  $\sigma(n)^{\text{th}}$  entry of *every* ken of all other agents  $j$  is also  $\square_j$ , then the modal operator of the  $\sigma(n)^{\text{th}}$  entry of some agent  $j$ 's ken at state  $s$  is not  $\square_j$ ; thus contradicting our initial assumption that all agents have the “same” ken at state  $s$ . Proceed similarly for the other cases.

The other direction is trivial.  $\square$

The above proposition shows that in generic cases - that is, for any model  $\mathcal{M} \notin \mathbf{M}^*$  -, the standard agreement theorems (using the assumptions of like-mindedness and the Sure-Thing Principle) will generally not hold if decision functions depend on interactive information (that is, when  $r > 0$ ) - since Lemma 6 in Tarbush (2011a) would no longer hold (with the ordering  $\succ^*$ ).

### 3.1 Example

We consider a very simple example in which the agreement theorem cannot hold for  $r > 0$ , which suggests that it does not hold in generic cases when decision functions are allowed to depend on interactive information.

Model  $\mathcal{M}$  in Figure 1 is again the coin in a box example, where  $\omega \models h$  and  $\omega' \models t$ . Agent  $j$  does not know whether the coin is facing heads up or tails up, but  $i$  does, and  $j$  knows that  $i$  knows. That is, at every state of  $\mathcal{M}$  it is the case that  $\square_j(\square_i h \vee \square_i t)$  and  $\hat{\square}_i(\square_j h \vee \square_j t)$ .

It follows that at the state at which the infimum holds for each agent, we have  $s \models \square_j(\square_i h \vee \square_i t) \wedge \hat{\square}_i(\square_j h \vee \square_j t)$ . So clearly, it is *not* the case that  $\inf\{\nu_i | \omega \in \Omega \ \& \ \omega \models \nu_i\} \sim^* \inf\{\nu_j | \omega \in \Omega \ \& \ \omega \models \nu_j\}$ . (Note, this shows that model  $\mathcal{M}$  in Figure 1 does not belong to the class  $\mathbf{M}^*$ ).

The simplicity of this case illustrates how general this phenomenon is, and therefore that in general, agreement theorems do not hold when decision functions are allowed to depend on interactive information (when  $r > 0$ ).

In fact, it should be easy to see that unlike the class  $\mathbf{M}$ , the class  $\mathbf{M}^*$  is actually very special: In order to construct a model in  $\mathbf{M}^*$ , one would effectively need to create a “symmetric” model. That is, one in which, if  $(\omega, \omega') \in R_i$ , then  $(\omega'', \omega''') \in R_j$  where for all  $p \in P^*$ ,  $\mathcal{V}(p, \omega) = \mathcal{V}(p, \omega'')$  and  $\mathcal{V}(p, \omega') = \mathcal{V}(p, \omega''')$ .

### 3.2 Other implications of $\succsim^*$

Note that Lemma 2 in Tarbush (2011b) does hold for  $r = 0$  even with the ordering  $\succsim^*$ . However, it does not hold for all  $r \geq 0$ . Indeed, it is possible that  $\omega \square_i \square_i p \wedge \square_j \square_j p$  and  $I_i(\omega) \subseteq I_j(\omega)$ . Therefore, it is possible that  $I_i(\omega) \subseteq I_j(\omega)$ ,  $\omega \models \nu_i \wedge \mu_j$  without it being the case that  $\nu_i \succsim^* \mu_j$ .

For the same reason, Lemma 5 in Tarbush (2011b) also fails for  $r > 0$ ; and since the theorems in the rest of the paper depend on these lemmas, those results also do not hold for  $r > 0$ .

Finally, the results in Tarbush (2011a) are largely unaffected since most already impose the restriction  $r = 0$ . However, the proof of Proposition 3 in that paper follows a similar logic to the proofs of Lemmas 2 and 5 in Tarbush (2011b); so although Proposition 3 still does hold for  $r = 0$ , it fails for  $r > 0$  if the ordering  $\succsim^*$  were used.

## References

- Bacharach, M.: 1985, ‘Some extensions of a claim of Aumann in an axiomatic model of knowledge’. *Journal of Economic Theory* **37**(1), 167–190.
- 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.
- Tarbush, B.: 2011a, ‘Agreeing to disagree: a syntactic approach’. *University of Oxford, Department of Economics, Working Paper*.
- Tarbush, B.: 2011b, ‘Generalisation of Samet’s (2010) agreement theorem’. *University of Oxford, Department of Economics, Working Paper*.