SOME RESULTS ON CONSENSUS

Extended Abstract

Paul J. Krasucki
Dept. of Mathematical Sciences
Rutgers University
Camden College of Arts and Sciences
Camden, NJ 08102
email: krasucki@cancer.rutgers.edu

ABSTRACT

We will show how reaching consensus among n individuals communicating in pairs depends on the topology of the communications graph. In particular we show that the consensus on the value of *union consistent* function is guaranteed in any non-cyclic fair protocol. We also anylyze protocols where individuals *exchange* (simultaneously) information in pairs. Finally we prove a surprising result that a certain non-trivial level of common knowledge of some formula which was not initially common knowledge is a *necessary* condition for a *disagreement.*

1 INTRODUCTION

There are many situations in which rational agents, trying to accomplish the same objective, act differently. Typical example are two people who trade stocks. Both want to maximize their profits and one is selling the stocks while the other one is buying the stocks of the same company. If we assume that they are like-minded (given the same information they would have acted in the same way) the only explanation is that they have received different information.

What if they are aware of the others actions so they have reason to believe that the information received by others leads to different conclusions? Can they 'agree to disagree' ? What if they can observe others actions and change their own actions accordingly? Will they start to act the same way?

We will continue the work started by Aumann [Aum76] and subsequently expanded by Geanakoplos and Polemarchakis [GP82] and others. Specifically we will use the model of one-to-one communications among n agents as introduced by Parikh and Krasucki [PK90].

2 DESCRIPTION OF THE MODEL

W is the space of states (or of possible worlds). There is n agents (individuals) numbered from 1 to n. Every individual i has an *accessibility relation* $\approx_i$. If w is the real world, i's information allows him to exclude all the worlds w' which are not accessible from w, but he cannot distinguish between the worlds w'' s.t. $w \approx_i w''$. We will assume that for every i, $\approx_i$ is an equivalence relation so its equivalence classes form a *partition* of W. We will further assume that all the partitions are *finite.* We will use P_i to denote a partition of an individual i given by $\approx_i$.

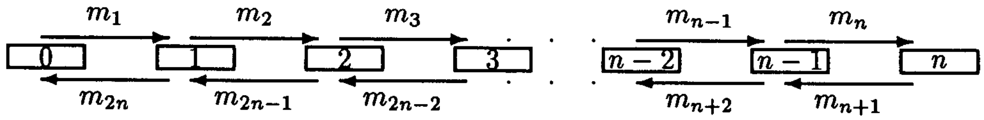

Figure 1: Channel Protocol

$P_i(w)$ will denote a set worlds accessible by $\approx_i$ from w (an element of the partition P_i containing the world w).

We are interested in consensus on the value of a fixed function f from $P(W)$ into some domain D.

It is always assumed that all the relations $\approx_i$ and the function f are common knowledge among all the individuals.

Agents communicate values of f according to a certain protocol which is common knowledge. Upon receiving information an agent refines his information structure; he removes from his set of possible worlds the states incompatible with the information received.

The minimum requirement is that the protocol is *fair*, i.e. that everybody communicates (directly or indirectly) his/her value of f to everybody else infinitely often. This is a necessary condition if we expect consensus, isolated agents don't have to agree with the others. Another necessary condition is *union consistency.*

Function f is *union consistent* iff for X, Y disjoint if $f(X) = f(Y)$, then $f(X) = f(X \cup Y)$ (the same property was also called *the sure thing principle*).

If a function is not union consistent, then if we take two agents such that one's information partition is a refinement of the other's, they may disagree.

There is an example (example 2 in [PK90]) of a function f defined s.t. f has the union consistency property but repeated communication of values of f among 3 individuals according to the round-robin protocol (individual 1 to 2 to 3 to 1 ...), fails to bring about consensus. But there are other natural protocols, in which communications are private. For example a 'channel' (1 to 2, 2 to 3,...,$n-1$ to n, n back to $n-1$, $n-1$ to $n-2$,...,2 to 1, and then again 1 to 2....), or a 'star' (0 to 1, 1 to 0, 0 to 2, 2 to 0,...,0 to n, n to 0, 0 to 1...).

Does the union consistency suffice to force consensus in these protocols? What is the relationship between the topology of a communication graph and a possibility of the disagreement? How can we characterize graphs of protocols which always bring about consensus on the value of a function satisfying the union consistency?

We will try to answer these questions in this paper.

3 BASIC RESULTS

A *protocol* Pr is a pair of functions $s(t), r(t)$ from the set of natural numbers to the set of individuals {1,...,n}. We intend to interprete $s(t)$ as the sender and $r(t)$ as the recipient of the communication which takes place at time t.

A protocol is *fair* iff in this protocol every participant communicates (directly or indirectly) with every other participant infinitely often (see [17]).

Let us assume that individuals communicate values of a function $f : 2^W \rightarrow D$ according to a protocol $Pr = (s(t), r(t))$. We define following [17] $P_i^t(x)$, the set of possible states for i at time t, given that the real state is x.

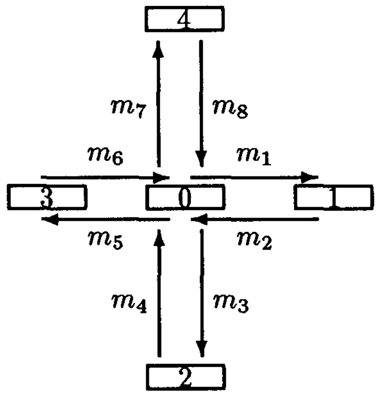

Figure 2: Star Protocol, $n = 4$, m_n same as $m_{(n mod 8)+1}$

$P_i^0(x) = P_i(x)$
If $i = r(t)$, then $P_i^{t+1}(x) = P_i^t(x) \cap \{y | m(y,t) = m(x,t)\}$
where $m(x,t) = f(P_{s(t)}^t(x))$.
Otherwise, $P_i^{t+1}(x) = P_i^t(x)$.

Initially everybody knows only which element of his/her partition contains the real world. At time t if a person doesn't receive any messages, his/her state of knowledge remains unchanged. Upon receiving a message, person excludes from his/her set of possible worlds, worlds incompatible with the received message.

We assume that all the communications are instantaneous, but this assumption is not essential.

We will also assume, that person 1 is the first person sending a message, and that the sender at time t is the person who most recently received new information: the recipient of a message at time $t-1$.

With these additional assumptions, the protocol is characterized by a sequence of agents: $Pr = (a_t)_{t \in N}$, where $a_0 = 1$, a_t is the sender at time $t-1$ as well as the recipient at time $t-2$.

We will need two lemmas which follow from the theorems proved in [PK90]:

Lemma 1: There is a t_0 such that for all x, i, and all $t, t' > t_0$, $P_i^t(x) = P_i^{t'}(x)$. Hence for all x, i, $P_i^t(x)$ has a limiting value $P_i^\infty(x)$. □

Lemma 2: If there is a space W, finite partitions P_1,...,P_n of W and a union consistent function f s.t. for some fair protocol Pr, the sets of possible limiting values are P^1,...,P^n, then there exist finite partitions P_1',..., P_n' of W and protocol Pr' with the same graph as Pr such that executing Pr', we get the *same* set of limiting values, but no one gains any knowledge during the execution of Pr'. I.e. $P_i(x) = P_i^t(x)$ for all i,x,t. □

In particular, if no consensus was reached in the first case, then in the second case also there will be no consensus, and moreover, no learning will ever take place.

Protocols in which no participant gains any knowledge we will call *zero-learning* protocols.

We will use these lemmas to prove a series of results characterizing protocols in which consensus among any n participants on the value of union-consistent f is always reached. Using lemmas 1 and 2 we can without loss of generality consider only situations in which a fair protocol is zero-learning.

Lemma 3: If i communicates values of a union consistent f to j, j sends value of f back to i and neither i nor j has learned anything, then i and j were sending the same value.

Proof: Let $s(t) = i$, $r(t) = j$, $s(t+1) = j$, $r(t+1) = i$. If there is no learning, then for every x, t, $P_i^t(x) = P_i^{t+2}(x)$ and $P_j^t(x) = P_j^{t+2}(x)$. So let $E(x,i,t)$ be a set of sets $P_i^t(y)$ s.t. there is a chain $x_1, x_2, ..., x_k$, s.t. $x_1 = x$, $x_k = y$ and for all $m < k$ $x_m \in P_i^t(x_{m+1})$ or $x_m \in P_j^t(x_{m+1})$. If there is no learning, then all sets in $E(x,i,t)$ are compatible with $P_i^t(x)$, i.e. they give the same value of f as $P_i^t(x)$. Similarly we can define $E(x,j,t)$. $P_j^{t+2}(x) = P_j^t(x) \subseteq \bigcup E(x,i,t)$. Similarly $P_i^{t+2}(x) = P_i^t(x) \subseteq \bigcup E(x,j,t)$. $\bigcup E(x,j,t) = \bigcup E(x,i,t)$. By the union-consistency of f, $f(\bigcup E(x,j,t)) = f(\bigcup E(x,i,t))$, and $f(P_j^t(x)) = f(P_i^t(x))$. □

Corollary 1: If i_1 sends value of f to i_2, i_2 to i_3,..., i_{k-1} to i_k and then i_k back to i_{k-1},...,back to i_1 and nobody has learned anything in the process, then they must have been sending all the same value.

Proof: By the lemma 3, i_k and i_{k-1} are sending the same value. So our protocol is equivalent to a protocol in which there is only $k-1$ agents, and the corollary follows by induction on k. □

Corollary 2: In a 'channel' protocol, consensus on the value of a union consistent function is always reached. □

Corollary 3: In a 'star' protocol, consensus on the value of a union consistent function is always reached.

Proof: A star protocol is equivalent to the channel protocol in which all agents i_{2m} have the same information partitions (all even-numbered agents in a channel protocol are copies of a 'central' agent from the center command protocol). □

Let us call a protocol *cyclic* iff there are agents $i_0, i_1, ..., i_k$, $k \geq 3$, s.t. for $j < k$, i_j communicates directly with i_{j+1}, i_k with i_1 infinitely often i.e. after a certain point T for every j, i_j after sending a message to i_{j+1} will not communicate directly or indirectly with any other agent in a cycle until he receives a message from i_{j-1}.[1]

There is an example of a cyclic protocol in figure 6 and a non-cyclic protocol in figure 5.

Theorem 1: In a fair protocol $Pr = (a_t)_{t \in N}$, consensus on the value of every union-consistent f is guaranteed iff Pr is not cyclic.

Proof outline: Pr is not cyclic so there must be agents (nodes in the graph of Pr) i_j s.t. a message for some k, m_k is sent in Pr by some i_s to i_j and m_{k+1} is sent by i_j to i_s. In a zero-learning protocol by the lemma 3 m_k and m_{k+1} must be the same. Moreover fact that they are the same is common knowledge. So we can delete a node i_j from our communication graph and if there were no consensus between some agents in the original graph, there will be no consensus in the protocol without i_j. Repeating this process of eliminating nodes, we can reduce

[1] Addition and subtraction are mod n

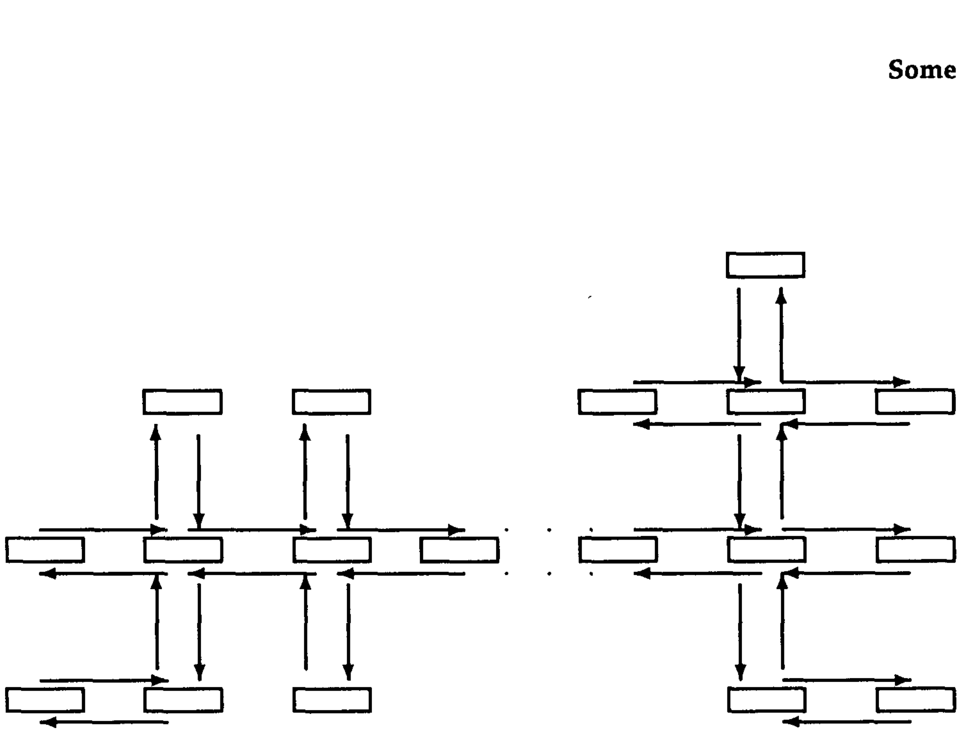

Figure 3: An example of a graph of a non-cyclic protocol

our communication graph to a single node.
If there is a cycle, then we can make our protocol equivalent to a counterexample from [PK90] by giving the agents forming a cycle, partitions from a counterexample and other agents partitions same as partitions forming a cycle. Formally, let $a_1, a_2, ...a_k$ form a cycle. Let the space W equal the set {1,2,3,4,5,6,7,8}. The function f satisfying the union consistency property takes integer values and is defined as follows: let the subsets of W be numbered X_1,X_2,... and let $num(X_i)$ be i.

Now we let:
$f(\{1,2\}) = f(\{3,4\}) = f(\{1,2,3,4\}) = 1$
$f(\{5,6\}) = f(\{7,8\}) = f(\{5,6,7,8\}) = 2$
$f(\{1,3\}) = f(\{5,7\}) = f(\{1,3,5,7\}) = 3$
$f(\{2,4\}) = f(\{6,8\}) = f(\{2,4,6,8\}) = 4$
$f(\{1,5\}) = f(\{2,6\}) = f(\{1,2,5,6\}) = 5$
$f(\{3,7\}) = f(\{4,8\}) = f(\{3,4,7,8\}) = 6$
and for all other subsets X of W, $f(X) = num(X) + 6$.

f has the union consistency property. Partitions are defined by:
$P_{a_1} = \{\{1,2\},\{3,4\},\{5,6\},\{7,8\}\}$
$P_{a_2} = \{\{1,3\},\{2,4\},\{5,7\},\{6,8\}\}$
$P_{a_3} = \{\{1,5\},\{2,6\},\{3,7\},\{4,8\}\}$.

All partitions P_{a_i} for $3 \leq i \leq k$ (remaining agents forming a cycle) are the same as P_{a_3}. For all the agents outside of the cycle we can assign the same partitions as the partitions of appropriate agents in the cycle, i.e. if i_j is not in a cycle $a_1, a_2, ...a_k$ then if there is a path from a_i in a cycle to i_j not going through any of the nodes $a_1, a_2, ..., a_{i-1}, a_{i+1}, ...a_k$, then we assign $P_{i_j} = P_{a_i}$. □

Note that in an example from figure 6 we can take $P_i = P_3$ for all $i \geq 3$, $P_0 = P_1$ and

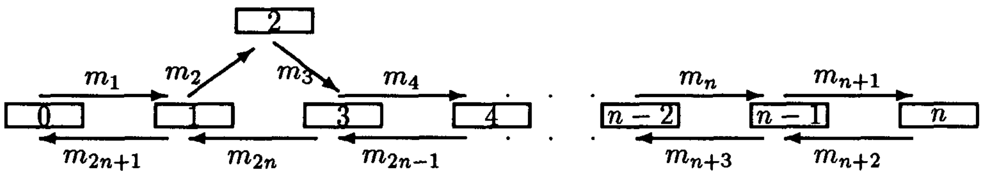

Figure 4: Cyclic Protocol (1,2,3 form a cycle)

P_1, P_2, P_3 and f as in the counterexample from [PK90].

4 EXCHANGE OF INFORMATION

Often when agents communicate, they don't just send messages to the other, but instead they *exchange* information, they hand each other their bids, then after learning from them they reevaluate their positions (an example of such a process could be reaching an agreement between a buyer and a seller of a house). That would correspond in our model to the situation, where sender and receiver in the same instant of time reverse their roles. We will have a single function c (for communicate) which would work as a simultaneous execution of two dual protocols P and Q s.t. $s_P = r_Q$ and $r_P = s_Q$.

Here are the modified definitions:

Let c be a function from the set of natural numbers (time) into the set of pairs of agents. Let $Pr = (c(t))$ be a protocol. There are two messages communicated at time t $m(i,j,x,t)$ (message communicated at t in a real world x by i to j) and $m(j,i,x,t)$, which we define by induction on t.

We will also define $P_i^t(x)$, the set of possible states for i at time t, given that the real state is x.

$P_i^0(x) = P_i(x)$

If $c(t) = (i,j)$, then:

1) $P_i^{t+1}(x) = P_i^t(x) \cap \{y | m(j,i,y,t) = m(j,i,x,t)\}$

2) $P_j^{t+1}(x) = P_j^t(x) \cap \{y | m(i,j,y,t) = m(i,j,x,t)\}$

where $m(j,i,x,t) = f(P_j^t(x))$, $m(i,j,x,t) = f(P_i^t(x))$

Otherwise, $P_i^{t+1}(x) = P_i^t(x)$.

Initially people know elements of their own information partitions containing the real world. In the process of exchanging information they remove worlds incompatible with received information.

Theorem 2: In any fair protocol with information exchange consensus on the value of a union-consistent function must be reached. Moreover it is reached fast: if every agent's partition has k elements, the agreement must be reached in $O(k)$ time.

Proof: The fact that the agreement must be reached follows from the lemma 3.

In order to prove a linear bound on complexity let us notice that in every communication, if one of the agents learns anything, it must be that one equivalence class $P_j^t(x)$ of his partner is incompatible with the message. So he can remove the whole equivalence class. Note that in some cases k may be very large. □

5 SOME COMMON KNOWLEDGE NECESSARY FOR DISAGREEMENT

In his paper Aumann [Aum76] proved that if there is common knowledge among two individuals of their posterior probabilities of a certain fixed event E, then these posterior probabilities must be equal.

Later Cave [Cav83] proved that Aumann's consensus theorem holds true if there is any number n of individuals who communicate their values of any union-consistent function f in a setting where every value communicated immediately becomes common knowledge among *all* the individuals.

These results are not surprising. Having common knowledge means sharing as much knowledge as possible. One can readily expect that like-minded individuals who share a lot of knowledge are likely to agree.

Below we have a result which seems to be unexpected: existence of a certain kind of common knowledge is a necessary condition for *disagreement.*

In a group of n individuals, any formula may be known at many *levels.* Two extreme cases are contradictions which, being always false, cannot be known by anybody, and tautologies which are common knowledge among all individuals. But there are infinitely many levels inbetween, e.g. a true formula may be known to individuls 1,2, and 3 but not known to the others, moreover 1 may know that 2 knows it and 1 may know that 3 knows it, but neither 2 nor 3 knows whether 1 knows it, and for no $i \neq j \neq k$ i knows that j knows that k knows. Similarly there may be formulae which are common knowledge between any group of k individuals but not common knowledge for any group of more than k individuals. For the detailed analysis of levels of knowledge that are attainable, see [Par86, PK90].

Let us say that there is *non-trivial* common knowledge in a group of individuals if there is a formula A which is common knowledge among individuals in a group, and there is a possible world $w \in W$ s.t. formula A is not true in w.

We will show that if there is no non-trivial common knowledge among any two communicating individuals, then consensus on the value of a union-consistent f is guaranteed in any fair protocol.

Our theorem states that in *all* examples in which there is no consensus ever reached on the value of a union-consistent function in a fair protocol, there must be reached a certain (non-empty) level of common knowledge of some formula which is initially not common knowledge among all individuals.

We must know more about each other in order to disagree!

Theorem 3: If there is a space W, finite partitions $P_1,...,P_n$ of W and there is no common knowledge of any non-trivial formula (formula which is false in at least one world in W) between any two communicating agents in a fair protocol Pr, then consensus on a value of a union consistent function f must be reached. □

Proof: By lemmas 1,2 we can assume that Pr is zero-learning. Suppose that i is sending value of f to j. If there are two equivalence classes $P_i^t(x)$ and $P_i^t(y)$ s.t. $f(P_i^t(x)) \neq f(P_i^t(y))$, then since there is no common knowledge between i and j, there must be some chain $x_1, x_2, ..., x_k$, s.t. $x_1 = x$, $x_k = y$ and for all $m < k$ $x_m \in P_i^t(x_{m+1})$ or $x_m \in P_j^t(x_{m+1})$. There must be two sets $P_i^t(x_l), P_i^t(x_{l+2})$ in the chain s.t. $f(P_i^t(x_l)) \neq f(P_i^t(x_{l+2}))$. In such case in a world x_{l+1}, since $P_i^t(x_l) \cap P_j^t(x_{l+1}) \neq \emptyset$ and $P_i^t(x_{l+2}) \cap P_j^t(x_{l+1}) \neq \emptyset$ and values sent by i in $P_i^t(x_l)$ and $P_i^t(x_{l+2})$ are different, in a world x_{l+1}, j would have learned something from the message, contrary to the assumption of zero-learning. So i must be sending the same value in all his equivalence classes. We can repeat the same argument for all participants. By union consistency $f(W) = f(P_i^t(x))$, so for every j, y, $f(P_i^t(x)) = f(P_j^t(y))$. □

In an example from the proof of the theorem 3 (as well as in the example 2 in [PK86]) there is common knowledge between the agents a_1 and a_2 of the fact that 'the real world has a number ≤ 4' (assuming that the real world is e.g. 1). Similarly there is common knowledge between the agents a_3 and a_2 of the fact that 'the real world has an odd number'. and there is common knowledge between the agents a_3 and a_1 of the fact that 'the real world has a number 1,2,5 or 6'. Nothing is common knowledge between all the agents which was not common knowledge between them before the execution of the protocol.

6 SUMMARY

If there is a group of n individuals communicating one to one values of a function satisfying sure-thing principle according to some fair protocol, then if there is at some point common knowledge among individuals of the value of f, then these individuals must agree on the value of f (earlier results, [Aum76, Bac85, Cav83]).

If there is no common knowledge between any two individuals of any non-trivial formula, they also must be in an agreement on the value of f (theorem 3).

If there is common knowledge between some pairs of individuals of some non-trivial formulae (but not the value of f) then there is disagreement possible if a protocol is cyclic. Non-cyclic protocols always lead to agreements.

If we put some additional restrictions on f, we may guarantee agreements even in cyclic protocols (see [Kra90, PK86]).

Acknowledgements

I would like to thank Rohit Parikh for getting me interested in the problems of consensus and for many helpful discussions as well as for reading earlier drafts of this manuscript and valuable comments. I would like to thank Don Beaver for helpful suggestions.

References

[Aum76] R. Aumann. Agreeing to disagree. *Annals of Statistics*, 4:1236–1239, 1976.

[Aum89] R. Aumann. Notes on Interactive Epistemology. 1989.

[Bac85] M. Bacharach. Some extensions of a claim of Aumann in an axiomatic model of knowledge. *Journal of Economic Theory*, 37:167–190, 1985.

[BD87] A. Brandenburger and E. Dekel. Common Knowledge With Probability 1. *J. of Math. Economics*, 16:237–245, 1987.

[Cav83] J. Cave. Learning to agree. *Economics Letters*, 12:147–152, 1983.

[Gen88] J. Geanakoplos. Common Knowledge, Bayesian Learning, and Market Speculation With Bounded Rationality. 1988.

[GP82] J. Geanakoplos and H. Polemarchakis. We can't disagree forever. *Journal of Economic Theory*, 28:192–200, 1982.

[Kra90] P. Krasucki. Reaching Consensus on Decisions. *Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning About Knowledge.* Morgan Kaufmann 1990.

[MP86] R. McKelvey and T. Page. Common Knowledge, Consensus, and Aggregate Information. *Econometrica*, 54:109–127, 1986.

[Mil81] P. Milgrom. An Axiomatic Characterization of Common Knowledge. *Econometrica*, 49:219–222, 1981.

[MS82] P. Milgrom and N. Stokey. Information, Trade and Common Knowledge. *Journal of Economic Theory*, 26:17–27, 1982.

[MN90] Y. Moses and G. Nachum. *Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning About Knowledge.* Morgan Kaufmann 1990.

[NBGMP] L.T. Nielsen, A. Brandenburger, J. Geanakoplos, R. McKelvey and T. Page. Common Knowledge of an Aggregate of Expectations. *Econometrica*, 58:1235–1239, 1990.

[Par86] R. Parikh. Levels of Knowledge in Distributed Computing. *IEEE Symposium on Logic in Computer Science*, Boston 1986, 322–331.

[PK90] R. Parikh and P. Krasucki. Communication, Consensus and Knowledge. *Journal of Economic Theory*, 52:178–189, 1990.

[PK86] R. Parikh and P. Krasucki. Levels of Knowledge in Distributed Computing. Research Report, Brooklyn College, CUNY (1986).

[RW90] A. Rubinstein and A. Wolinsky. On the Logic of "Agreeing to Disagree" Type Results. *Journal of Economic Theory*, 51:184–193, 1990.

[Sam90] D. Samet. Ignoring Ignorance and Agreeing to Disagree. *Journal of Economic Theory*, 52, 1990.

[SG83] J. Sebenius and J. Geanakoplos. Don't Bet On It: Contingent Agreements With Asymmetric Information. *J. of the American Statistical Association*, 78:424–426, 1983.