Agreeing to Disagree

The previous analysis demonstrated a formal connection between agreement and common knowledge.

To coordinate their attack, the generals have to agree to attack together at a particular time. We have shown that common knowledge is necessary and sufficient condition for such agreement to hold.

Common knowledge has surprizing consequences in applications in which the players are attempting to agree to také *different actions.*

Unlike to the coordinated-attack where the agents are attempting to agree to take essentially *the same action.*

Example. (Trading in stock market)

Here, the transaction occurs when one side *buys* and the other side sells.

Why people trade?

Some trades are certainly due to the fact that people may have different utilities for having money at a given moment:

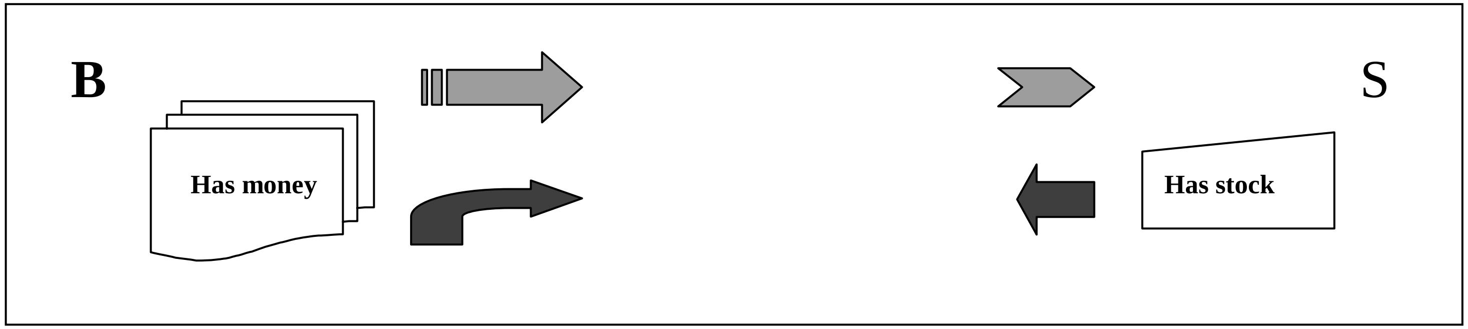
Case 1. The simple and naive one. One person may need to make a big payment and will therefore want to sell stock, while the other may have just received a large sum of money and wish to invest some of it in the stock market.

The seller thinks the price of a given stock is likely to go down, while the buyer believes it will go up .

Perhaps the buyer has some information leading him to belive that the company that issued that stock go well in the next year, while the seller has information indicating that the company may fail.

Some trades are certainly due to the fact that people may have different utilities for having money at a given moment:

The seller sells and the buyer buys.



For a trade to take place, both sides have to agree to the transaction.

Each agent uses information of his/her.

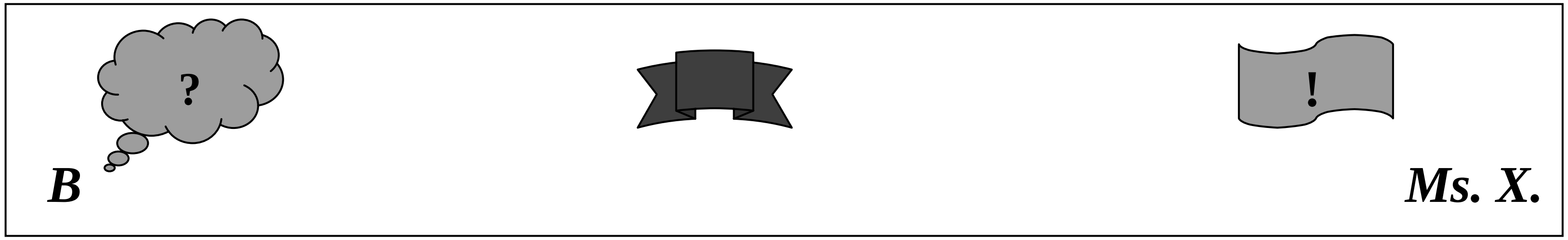
No *Common K nowledge* is needed.

Case 2. Speculation - a somewhat extreme case.

The seller is Ms. X, a top executive in the company whose stock is being traded.

“Clearly,” the intended buyer should reason, “if Ms. X is selling, then the stock price is likely to drop. Thus, if she is willing to sell the stock for $\$ k$, then I should not buy it for that amount.”

Since the participants in the trade have reached the common knowledge when the trade takes place, they should make use of this knowledge when making their decisions.



Surprisingly, if they use this knowledge, the the trade cannot take place.

More precisely, we show that if both sides act according to the same rules, then the common knowledge that would arise prevents the trade from taking place.

Comment. Roughly speaking, the result says that players cannot „agree to disagree“, i.e. , they cannot have common k nowledge that they are taking different actions, such as buying and selling.

Notice that the word „agree“plays two different roles in the phrase „agree to disagree“:

„Agree“ refers to common knowledge, while „disagree“refers to reaching different decisions.

To describe correctly this result, we need some definitions.

Recall that the actions taken by the players are prescribed by their protocols, where a protocol for player *i* is a function of player *i*‘s local state.

In many applications, it is more appropriate to view the players actions as depending not on her local state, but on the set of points she consideres possible.

For example, suppose that a player wants maximize some payoff that depends on the point. Since the player does not know what the actual point is, her decision actually depends on the set of points that she considers possible.

In two different systems, she may have the same local state, but consider a different set of points possible, and thus takes different actions.

This is a phenomenon that is a topic of knowledge-base programs.

We generalize our notion of protocol in an attempt to capture this intuition.

Definition. (Information sets)

Given a local state ℓ for player i in a system R , let $\mathbf{IS}_i(\ell , R)$ denotes the set of points (r , m) in R such that $r_i (m) = \ell$. If $I = (\ell , R)$ is an interpreted system, we identify $\mathbf{IS}_i(\ell , I)$ with $\mathbf{IS}_i(\ell , R)$.

Comment. In terminology of event-based approach (Aumann), $\mathbf{IS}_i(\ell , R)$ is the *information set* of player i when in local state ℓ ,in the interpreted system R .

We want to to view the player i‘s action as a function of her information set rather than as a function of her local state. Essentially, this amounts to making a player’saction function of her *knowledge.*

Definition. (Decision functions)

Given a set G of global states, let S be the set of points over G i.e. S is the set of points (r, m) in a run r over G. If ACT_i is the set of actions for player i, we define a *decision function* for player i (over G) to be function with the domain consisting of some subsets of S with the range ACT_i.

Comment. Thus, an decision function prescribes an action for the subsets S in its domain.

We have another method for prescribing actions for player i. Namely , by means of a protocol.

How are decision functions and protocols related ?

In a precise sense, we can view decision functions as more general than protocols.

To know what action a decision function prescribes for player i in a given local state, we need to know what player i's information is.

But this depends not just player's local state, but on the whole system.

We may be able to associate a protocol with a given decision function once we have a system in hand to determine what the information sets are.

Definition. (Protocols compatible with Decision functions)

Given a decision function D for player i and a system R, we say that a protocol P_i for player i is *compatible with* D *in* R if

$$P_i(\ell , R) = D(\mathbf{IS}_i(\ell , R)) \quad \text{for all } \ell \text{ ,in } L_i$$

Comment. Note that the definition requires that the domain of D includes all the information sets of player i in R .

It is not hard to see that every deterministic protocol is compatible with some decision function D : i.e. , to show that if P_i is compatible with D in all systems R .

As follows from our discussion, we are mainly interested to apply decision function to information sets. But we have, however, allowed decision functions to be defined on *arbitrary sets of points.*

There are many many sets, that cannot be information sets. In particular, any set that includes points (r, m) and (r', m') such that

$$\neg (r_i(m) = r_i'(m'))$$

Why are we allowing decision functions to be defined on such sets ?

We have two reasons:

(i) this makes it possible for us to talk about all player‘s using *the same* decision function, as we do in examples to be shown later on.

(ii) as these exaples will show, it is often the case that the decision fuction we have in mind is most naturally thought of as a function on arbitrary sets of points.

As we have already said, we are interested in situations where all players use the same decision function.

Consequently, we assume that all the player‘s actions are all from the same set ACT .

Definition. (Protocols implementing decision functions)

We say that the joint protocol $P = (P_1, \dots, P_n)$ *implements* the decision function D in context γ if P_i is compatible with D in $\mathbf{R}^{rep}(P, \gamma)$, for $i = 1, \dots, n$.

Comment. Thus, if P implements D, then the actions prescribed by both P and D agree in the system representing P.

We are now almost ready to state the *Agreement Theorem* a fundamental result in Game Theory.

What we want to show is that if two players use the same decision function, the they cannot agree to perform different actions.

To capture this, we restrict our attention to *interpreted contexts* (γ, π^{ag}) *for agreement.*

How to simplify the contexts ?

As in the case of coordinated attack, we want to put as few restrictions on our contexts as possible. We assume that

- the player‘s actions are all taken from the same set ACT ,
- γ is a recording context ,
- for each action **a** , there is a primitive proposition $perf_i\,(\mathbf{a})$,

• (π^{ag})(s)($perf_i\,(\mathbf{a})$) is true in a state s if the action **a** was performed by player i , as recorded in the environment‘s state,

- as we did for the case of the coordinated attack, we take $act_i\,(\mathbf{a})$ to be an abbreviation for

$$\neg\, perf_i\,(\mathbf{a}) \;\;\&\;\; \mathrm{O}\, perf_i\,(\mathbf{a})$$

- thus, $act_i\,(\mathbf{a})$ is true if player i is about to perform action **a** .

Definition. (Union-consistent decision functions)

A decision function D is said to be *union-consistent* if it satisfies the following condition:

for every action **a** and for every collection $T_1, T_2, \dots, T_k, \dots$ of pairwise disjoint subsets of S the following holds :

$D(T_j) =$ **a** for all j -> ($\mathrm{U}_j(\mathrm{T}_j)$ is in the domain of D & $D(\mathrm{U}_j(\mathrm{T}_j))$ = **a**)

Comment. Intuitively, the function D is union-consistent if, whenever it prescribes the same action **a** for pairwise disjoint sets of points, it prescribes **a** for the union of all these sets as well.

Union-consistency seems fairly reasonable : it says that if a player performs the action **a** whenever she consideres T_j to be the set of possible worlds , then she should also perform **a** if she consideres $\mathrm{U}_j(T_j)$ possible.

Recall that we observed that any deterministic protocol can be obtained from some decision function.

In fact, it can be shown that any deterministic protocol can be obtained from some *union-consistent* decision function. (Exercise).

We give some examples of union-consistent functions after stating the theorem.

We can now formally state the Agreement Theorem. We shall write simply C instead of $C_{\{1,2\}}$ to represent common knowledge among the two players.

Agreement Theorem.

Suppose that $I = \mathbf{I}^{rep}(P, \gamma, \pi^{ag})$*, where* P *is a joint protocol and* (γ, π^{ag}) *is an interpreted context for agreement. If* P *implements some union-consistent decision function in context* γ*, then for all actions* **a , b** *in ACT,*

$$\textit{if} \quad I \models C(act_1(\mathbf{a}) \,\&\, act_2(\mathbf{b})) \quad \text{then} \quad \mathbf{a} = \mathbf{b} \,.$$

Comment. Thus, if two agents use the same union-consistent function , i.e. they act according to the same rules, they cannot have common knowledge that they are taking different actions. In other words, they cannot agree to disagree.

We observed earlier that every protocol for player i is compatible with *some* union-consistent decision function.

The crux of the Agreement Theorem is the requirement that both players use *the same union-consistent decision function.*

How reasonable is this requirement ?

We now describe three examples of situations where this arises.

Example 1. (A roulette game)

Suppose that two players each perform an action and receive a payoff as a result of that action.

Moreover, suppose that the payoff to player i depends solely on the action that player i performs and the global state at which the action is performed.

This means that, in particular, the payoff is independent of the action that the other player performs and both players receive the same payoff if they perform the same action.

For example, if the two players are betting on numbers in a roulette game, and we assume that the winning number is completely determined by the global state, then each player‘s payoff depends only on the global state (and not on the bet made by the other player), and the players receive the same payoff if they make the same bet.

Of course, the problem is that, in such scenarios, the players do not know what the global state is, so they do not know what their payoff will be.

Definition. (Risk averse players)

We say that a player is *risk averse* , if she choose the action for which the worst-case payoff is maximal.

(i) Formally, if s is a global state and $\mathbf{a}$ is an action, let *payoff* (s, $\mathbf{a}$) be the payoff for performing action $\mathbf{a}$ at the global state s.

(ii) Let $S`$ be a set of points, **a** be an action. We define

$$\mathit{payoff}\,(S`, \mathbf{a}) = \min\{\, \mathit{payoff}\,(r(m), \mathbf{a}) \mid (r, m) \text{ in } S` \}$$

To be precise, we should use the infimum rather then the minimum, since a minimum over an infinite set of payoffs may not exist.)

Comment. Clearly, $\mathit{payoff}\,(S`, \mathbf{a})$ is the worst-case payoff if the action **a** is performed in the set $S`$.

Definition. (Risk averse decision functions)

Let ACT be a finite set of actions and let for all subsets $S`$ of points and for all distinct pairs of actions **a** and **b ,** we have

$$\neg\, (payoff\,(S`, \mathbf{a}\,) \;=\; payoff\,(S`, \mathbf{b}\,))\;.$$

Under these assumptions, we define $D^{ra}(\,S`\,)$ to be the unique action **a** such that

$$payoff\,(S`, \mathbf{a}\,) \;>\; payoff\,(S`, \mathbf{b}\,)$$

for all actions **b** different from **a .**

Comment. The *´ra´* in D^{ra} stands for *rick averse.*

Thus , according to the decision function D^{ra} , the action chosen in $S`$ is the one that maximizes the worst-case payoff in $S`$.

It is easy to check that D^{ra} is a union-consistent function (Excersise).

It follows that if the players are both risk averse in this sense, they cannot agree to disagree. Thus, if they discuss their actions until they have common knowledge of the actions they are about to perform, then these actions must be the same.

In our remaining two examples, the player‘s decisions are defined in terms of probability.

Example 2. (Probability-based)

Definition. Suppose, we have a probability distribution $\Pr$ defined on certain subsets of S, the set of points. Suppose that e is a fixed subset of points (i.e. e is an event) , ACT the set of actions, just consists of all numbers in the closed interval [0 , 1].

Let the decision function D^{pr} be defined on subsets $S‘$ of S for which Pr(e / S‘) is defined. On these subsets, we define

$$D^{pr}(S‘) = \mathrm{Pr}(e / S‘)$$

Thus, $D^{pr}(S‘)$ is the conditional probability of e given $S‘$.

Excercise . It is easy to show that D^{pr} is union-consistent.

Proposition.

Under the above assumptions, if player i is in local state ℓ , then his estimate of the probability of e is given by the conditional probability $\mathrm{Pr}(e / \mathbf{IS}_i (\ell\ , R))$.

Thus, according to the Agreement Theorem, if the players have common knowledge of their estimate of the probability of e , then these estimate must be the same.

Comment. To bring our the perhaps surprising nature of this example, suppose that the players start with the same probability distribution on the set of points and then receive some information that causes them to revise their estimate of the probability of e, using conditioning. They can then exchange their estimates of e.

Doing so can cause them to revise further their estimates of e, since their information changes.

Suppose that after a while they reach steady state, and no further exchanges of their information can cause them to revise these estimates.

It is possible to show that in a large class of contexts, the players are in fact guaranteed to reach steady state and to attain common knowledge of the fact that they are in steady state.

Once this happens, their estimates are common knowledge, so according to the Agreement Theorem, they must be the same.

Thus, although the players might have different information, they cannot agree to disagree on their estimates of the probability of e .

Example 3. (A probabilistic version of Example 1)

We return to the setting of the first example but, as in the second example, we assume that we have a probablistic distribution on the set of points.

Rather then being risk averse, as in the first example, suppose that the players perform the action that has the highest expected rate of return.

That is, we now define the function $\mathit{payoff}\,`(S`, \mathbf{a})$ to be the expected payoff of the action $\mathbf{a}$ over the set $S`$.

Definition. (Utility maximizer)

Assume that *ACT* is a finite set of actions and for all distinct actions **a** and **b** , we have

$$\neg\,(payoff\,(S`, \mathbf{a}\,) \;=\; payoff\,(S`, \mathbf{b}\,))\;.$$

Under these conditions, we define $D^{um}(S`)$ to be the unique action **a** such that

$$payoff\,(S`, \mathbf{a}\,) \;>\; payoff\,(S`, \mathbf{b}\,)$$

for every pair of distinct actions **a , b .**

Comment. The „*um*“ in D^{um} stands for *utility maximizer*.

It is easy to see that D^{um} is union consistent. (Exercise)

Again, the Agreement Theorem tells us that if the player‘s protocols are consistent with D^{um} , then they cannot agree to disagree.

Is the Agreement Theorem counterintuitive ?

If we have look at actions in Examples 1 and 3 , and take them to be *buy* and *sell* , then we are back to the scenario with which we began this section.

Thus, in this setting, the Agreement Theorem tells us that speculative trading between players who follow the same rules (e.g. have the same payoff function and are both risk averse in Example 1, or have the same payoff function and probability distribution and are both utility maximisers in Example 3) is impossible.

This certainly seems counterintuitive.

Visions.

There has been a lot of work in Game Theory on trying to understand the implications of this result and to avoid the apparent paradox.

Some of the approaches involve so called *limited reasoners* a topic we will discuss later on.