Reasoning about Knowledge and Beliefs

Lecture 12

Eric Pacuit

University of Maryland, College Park

pacuit.org
epacuit@umd.edu

October 21, 2013

Robert Aumann. *Agreeing to Disagree*. Annals of Statistics **4** (1976).

Theorem. Suppose that n agents share a common prior and have different private information. If there is common knowledge in the group of the posterior probabilities, then the posteriors must be equal.

Robert Aumann. *Agreeing to Disagree*. Annals of Statistics **4** (1976).

Theorem. Suppose that n agents share a common prior and have different private information. If there is common knowledge in the group of the posterior probabilities, then the posteriors must be equal.

S. Morris. *The common prior assumption in economic theory*. Economics and Philosophy, 11, pgs. 227 - 254, 1995.

Generalized Aumann's Theorem

Qualitative versions: *like-minded individuals cannot agree to make different decisions*.

M. Bacharach. *Some Extensions of a Claim of Aumann in an Axiomatic Model of Knowledge*. Journal of Economic Theory (1985).

J.A.K. Cave. *Learning to Agree*. Economic Letters (1983).

D. Samet. *Agreeing to disagree: The non-probabilistic case*. Games and Economic Behavior, Vol. 69, 2010, 169-174.

The Framework

Knowledge Structure: $\langle W, \{\Pi_i\}_{i \in \mathcal{A}} \rangle$ where each Π_i is a partition on W ($\Pi_i(w)$ is the cell in Π_i containing w).

Decision Function: Let D be a nonempty set of **decisions**. A decision function for $i \in \mathcal{A}$ is a function $\mathbf{d}_i : W \to D$. A vector $\mathbf{d} = (d_1, \ldots, d_n)$ is a decision function profile. Let $[\mathbf{d}_i = d] = \{w \mid \mathbf{d}_i(w) = d\}$.

The Framework

Knowledge Structure: $\langle W, \{\Pi_i\}_{i \in \mathcal{A}} \rangle$ where each Π_i is a partition on W ($\Pi_i(w)$ is the cell in Π_i containing w).

Decision Function: Let D be a nonempty set of **decisions**. A decision function for $i \in \mathcal{A}$ is a function $\mathbf{d}_i : W \to D$. A vector $\mathbf{d} = (d_1, \ldots, d_n)$ is a decision function profile. Let $[\mathbf{d}_i = d] = \{w \mid \mathbf{d}_i(w) = d\}$.

(A1) Each agent knows her own decision:

$$[\mathbf{d}_i = d] \subseteq K_i([\mathbf{d}_i = d])$$

Comparing Knowledge

$[j \succeq i]$: agent j is at least as knowledgeable as agent i.

$$[j \succeq i] := \bigcap_{E \in \wp(W)} (K_i(E) \Rightarrow K_j(E)) = \bigcap_{E \in \wp(W)} (\neg K_i(E) \cup K_j(E))$$

Comparing Knowledge

$[j \succeq i]$: agent j is at least as knowledgeable as agent i.

$$[j \succeq i] := \bigcap_{E \in \wp(W)} (K_i(E) \Rightarrow K_j(E)) = \bigcap_{E \in \wp(W)} (\neg K_i(E) \cup K_j(E))$$

$w \in [j \succeq i]$ then j knows at w every event that i knows there.

Comparing Knowledge

$[j \succeq i]$: agent j is at least as knowledgeable as agent i.

$$[j \succeq i] := \bigcap_{E \in \wp(W)} (K_i(E) \Rightarrow K_j(E)) = \bigcap_{E \in \wp(W)} (\neg K_i(E) \cup K_j(E))$$

$w \in [j \succeq i]$ then j knows at w every event that i knows there.

$[j \sim i] = [j \succeq i] \cap [i \succeq j]$

The Sure-Thing Principle

A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant.

The Sure-Thing Principle

A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant. So, to clarify the matter to himself, he asks whether he would buy if he knew *that the Democratic candidate were going to win, and decides that he would.*

The Sure-Thing Principle

A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant. So, to clarify the matter to himself, he asks whether he would buy if he knew *that the Democratic candidate were going to win, and decides that he would. Similarly, he considers whether he would buy if he* knew *a Republican candidate were going to win, and again he finds that he would.*

The Sure-Thing Principle

A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant. So, to clarify the matter to himself, he asks whether he would buy if he knew *that the Democratic candidate were going to win, and decides that he would. Similarly, he considers whether he would buy if he* knew *a Republican candidate were going to win, and again he finds that he would. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains, or will obtain, as we would ordinarily say.* *(Savage, 1954)*

Sure-Thing Principle

Should I study or have a beer?

Sure-Thing Principle

Should I study or have a beer? Either I pass or I won't pass the exam.

Sure-Thing Principle

Should I study or have a beer? Either I pass or I won't pass the exam. If I pass, it is better to drink and pass, so I should drink. If I fail, it is better to drink and fail, so I should drink.

Sure-Thing Principle

Should I study or have a beer? Either I pass or I won't pass the exam. If I pass, it is better to drink and pass, so I should drink. If I fail, it is better to drink and fail, so I should drink. I should drink in either case, so I should have a drink.

Sure-Thing Principle

It is not the logical principle $\varphi \to \chi$ and $\psi \to \chi$ then $\varphi \vee \psi \to \chi$.

Sure-Thing Principle

It is not the logical principle $\varphi \to \chi$ and $\psi \to \chi$ then $\varphi \vee \psi \to \chi$.
There is a book I want to read which was written by one of two authors.

Sure-Thing Principle

It is not the logical principle $\varphi \to \chi$ and $\psi \to \chi$ then $\varphi \vee \psi \to \chi$. There is a book I want to read which was written by one of two authors. If I know it is written by author A then I will read it. If I know it is written by author B then I will read it.

Sure-Thing Principle

It is not the logical principle $\varphi \to \chi$ and $\psi \to \chi$ then $\varphi \vee \psi \to \chi$. There is a book I want to read which was written by one of two authors. If I know it is written by author A then I will read it. If I know it is written by author B then I will read it. If I know it is written by either author A or author B then I may not choose to read the book.

Sure-Thing Principle

There are three candidates, republican, independent and democrat.

Sure-Thing Principle

There are three candidates, republican, independent and democrat. I will buy stock if the democrat looses and I will buy stock if the republican looses.

Sure-Thing Principle

There are three candidates, republican, independent and democrat. I will buy stock if the democrat looses and I will buy stock if the republican looses. Either the republican or democrat will loose. So, I should buy the stock.

R. Aumann, S. Hart and M. Perry. *Conditioning and the Sure-Thing Principle*. manuscript, 2005.

The Nixon Diamond

You're told (from a reliable source) that Nixon is a republican, which suggests that he is a Hawk. You're also told (from a reliable source) that Nixon is a Quaker, which suggests that he is a Dove.

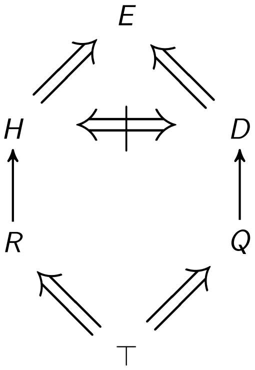
The Nixon Diamond

You're told (from a reliable source) that Nixon is a republican, which suggests that he is a Hawk. You're also told (from a reliable source) that Nixon is a Quaker, which suggests that he is a Dove. Either being a Hawk or a Dove implies having extreme political views.

The Nixon Diamond

You're told (from a reliable source) that Nixon is a republican, which suggests that he is a Hawk. You're also told (from a reliable source) that Nixon is a Quaker, which suggests that he is a Dove. Either being a Hawk or a Dove implies having extreme political views. Should you conclude that Nixon has extreme political views?

Floating Conclusions



J. Horty. *Skepticism and floating conclusions*. Artificial Intelligence, 135, pp. 55 - 72, 2002.

Your parents have 1M inheritance which will is split between you mother and father (each may give you 0.5M).

Your parents have 1M inheritance which will is split between you mother and father (each may give you 0.5M). Your brother (a reliable source) says that you will receive the money from your Mother (but not your Father).

Your parents have 1M inheritance which will is split between you mother and father (each may give you 0.5M). Your brother (a reliable source) says that you will receive the money from your Mother (but not your Father). Your sister (a reliable source) says that you will receive the money from your Father (but not your Mother).

Your parents have 1M inheritance which will is split between you mother and father (each may give you 0.5M). Your brother (a reliable source) says that you will receive the money from your Mother (but not your Father). Your sister (a reliable source) says that you will receive the money from your Father (but not your Mother). You want to buy a yacht which requires a large deposit and you can only afford it provided you inherit the money.

Your parents have 1M inheritance which will is split between you mother and father (each may give you 0.5M). Your brother (a reliable source) says that you will receive the money from your Mother (but not your Father). Your sister (a reliable source) says that you will receive the money from your Father (but not your Mother). You want to buy a yacht which requires a large deposit and you can only afford it provided you inherit the money. Should you make a deposit on the yacht?

Interpersonal Sure-Thing Principle (ISTP)

For any pair of agents i and j and decision d,

$$K_i([j \succeq i] \cap [\mathbf{d}_j = d]) \subseteq [\mathbf{d}_i = d]$$

Interpersonal Sure-Thing Principle (ISTP): Illustration

Suppose that Alice and Bob, two detectives who graduated the same police academy, are assigned to investigate a murder case.

Interpersonal Sure-Thing Principle (ISTP): Illustration

Suppose that Alice and Bob, two detectives who graduated the same police academy, are assigned to investigate a murder case. If they are exposed to different evidence, they may reach different decisions.

Interpersonal Sure-Thing Principle (ISTP): Illustration

Suppose that Alice and Bob, two detectives who graduated the same police academy, are assigned to investigate a murder case. If they are exposed to different evidence, they may reach different decisions. Yet, being the students of the same academy, the method by which they arrive at their conclusions is the same.

Interpersonal Sure-Thing Principle (ISTP): Illustration

Suppose that Alice and Bob, two detectives who graduated the same police academy, are assigned to investigate a murder case. If they are exposed to different evidence, they may reach different decisions. Yet, being the students of the same academy, the method by which they arrive at their conclusions is the same. Suppose now that detective Bob, a father of four who returns home every day at five oclock, collects all the information about the case at hand together with detective Alice.

Interpersonal Sure-Thing Principle (ISTP): Illustration

However, Alice, single and a workaholic, continues to collect more information every day until the wee hours of the morning — information which she does not necessarily share with Bob.

Interpersonal Sure-Thing Principle (ISTP): Illustration

However, Alice, single and a workaholic, continues to collect more information every day until the wee hours of the morning — information which she does not necessarily share with Bob. Obviously, Bob knows that Alice is at least as knowledgeable as he is.

Interpersonal Sure-Thing Principle (ISTP): Illustration

However, Alice, single and a workaholic, continues to collect more information every day until the wee hours of the morning — information which she does not necessarily share with Bob. Obviously, Bob knows that Alice is at least as knowledgeable as he is. Suppose that he also knows what Alices decision is.

Interpersonal Sure-Thing Principle (ISTP): Illustration

However, Alice, single and a workaholic, continues to collect more information every day until the wee hours of the morning — information which she does not necessarily share with Bob. Obviously, Bob knows that Alice is at least as knowledgeable as he is. Suppose that he also knows what Alices decision is. Since Alice uses the same investigation method as Bob, he knows that had he been in possession of the more extensive knowledge that Alice has collected, he would have made the same decision as she did. Thus, this is indeed his decision.

Implications of ISTP

Proposition. If the decision function profile $\mathbf{d}$ satisfies ISTP, then

$$[i \sim j] \subseteq \bigcup_{d \in D} ([\mathbf{d}_i = d] \cap [\mathbf{d}_j = d])$$

ISTP Expandability

Agent i is an **epistemic dummy** if it is always the case that all the agents are at least as knowledgeable as i. That is, for each agent j,

$$[j \succeq i] = W$$

A decision function profile $\mathbf{d}$ on $\langle W, \Pi_1, \ldots, \Pi_n \rangle$ is **ISTP expandable** if for any expanded structure $\langle W, \Pi_1, \ldots, \Pi_{n+1} \rangle$ where $n+1$ is an epistemic dummy, there exists a decision function $\mathbf{d}_{n+1}$ such that $(\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_{n+1})$ satisfies ISTP.

ISTP Expandability: Illustration

Suppose that after making their decisions, Alice and Bob are told that another detective, one E.P. Dummy, who graduated the very same police academy, had also been assigned to investigate the same case.

ISTP Expandability: Illustration

Suppose that after making their decisions, Alice and Bob are told that another detective, one E.P. Dummy, who graduated the very same police academy, had also been assigned to investigate the same case. In principle, they would need to review their decisions in light of the third detectives knowledge: knowing what they know about the third detective, his usual sources of information, for example, may impinge upon their decision.

ISTP Expandability: Illustration

But this is not so in the case of detective Dummy. It is commonly known that the only information source of this detective, known among his colleagues as the couch detective, is the TV set.

ISTP Expandability: Illustration

But this is not so in the case of detective Dummy. It is commonly known that the only information source of this detective, known among his colleagues as the couch detective, is the TV set. Thus, it is commonly known that every detective is at least as knowledgeable as Dummy.

ISTP Expandability: Illustration

But this is not so in the case of detective Dummy. It is commonly known that the only information source of this detective, known among his colleagues as the couch detective, is the TV set. Thus, it is commonly known that every detective is at least as knowledgeable as Dummy. The news that he had been assigned to the same case is completely irrelevant to the conclusions that Alice and Bob have reached. Obviously, based on the information he gets from the media, Dummy also makes a decision.

ISTP Expandability: Illustration

But this is not so in the case of detective Dummy. It is commonly known that the only information source of this detective, known among his colleagues as the couch detective, is the TV set. Thus, it is commonly known that every detective is at least as knowledgeable as Dummy. The news that he had been assigned to the same case is completely irrelevant to the conclusions that Alice and Bob have reached. Obviously, based on the information he gets from the media, Dummy also makes a decision. We may assume that the decisions made by the three detectives satisfy the ISTP, for exactly the same reason we assumed it for the two detectives decisions

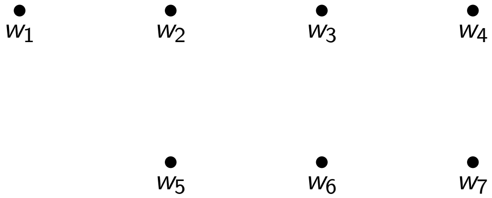
Generalized Agreement Theorem

If $\mathbf{d}$ is an ISTP expandable decision function profile on a partition structure $\langle W, \Pi_1, \ldots, \Pi_n \rangle$, then for any decisions $d_1, \ldots, d_n$ which are not identical, $C(\bigcap_i [\mathbf{d}_i = d_i]) = \emptyset$.

Robert Aumann. *Agreeing to Disagree*. Annals of Statistics **4** (1976).

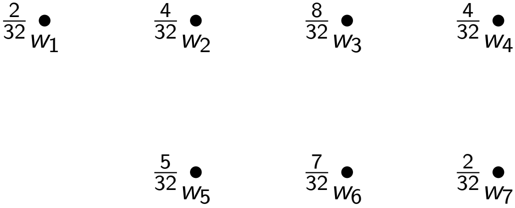
Theorem. Suppose that n agents share a common prior and have different private information. If there is common knowledge in the group of the posterior probabilities, then the posteriors must be equal.

2 Scientists Perform an Experiment



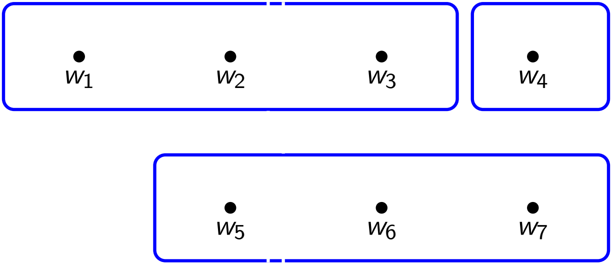
They agree the true state is one of seven different states.

2 Scientists Perform an Experiment



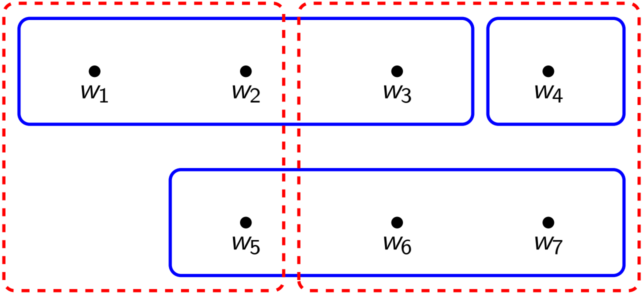
They agree on a common prior.

2 Scientists Perform an Experiment



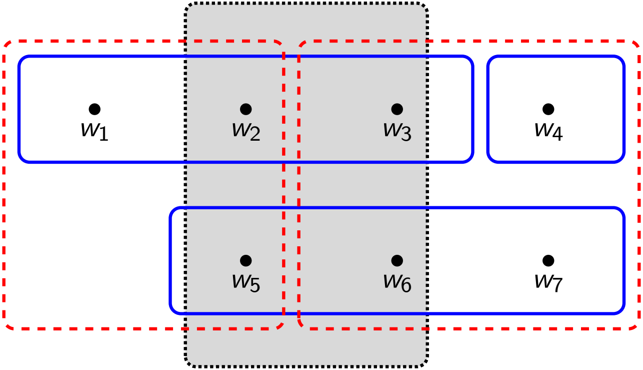
They agree that Experiment 1 would produce the blue partition.

2 Scientists Perform an Experiment



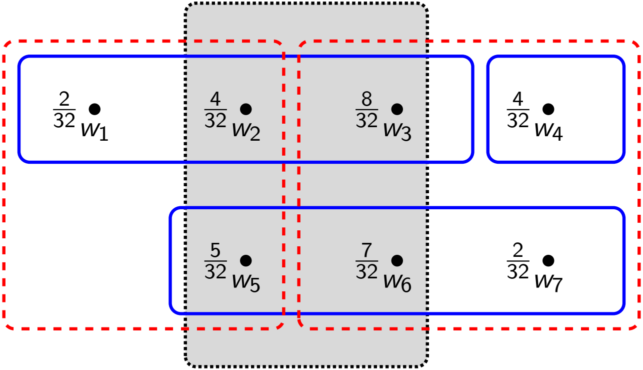
They agree that Experiment 1 would produce the blue partition and Experiment 2 the red partition.

2 Scientists Perform an Experiment



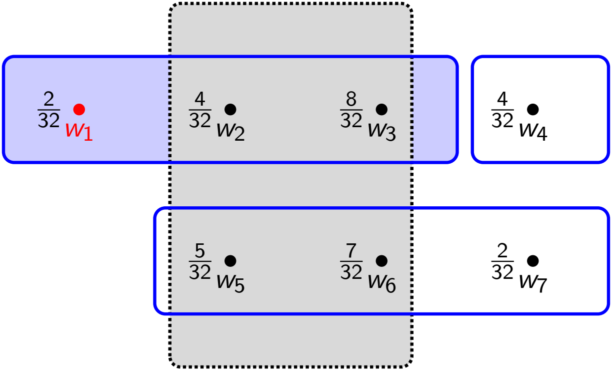
They are interested in the truth of $E = \{w_2, w_3, w_5, w_6\}$.

2 Scientists Perform an Experiment



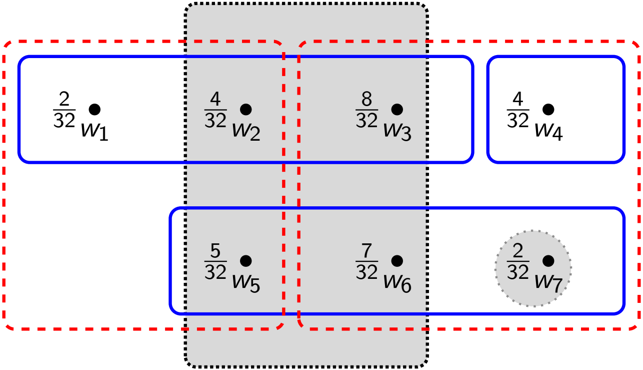
So, they agree that $P(E) = \frac{24}{32}$.

2 Scientists Perform an Experiment



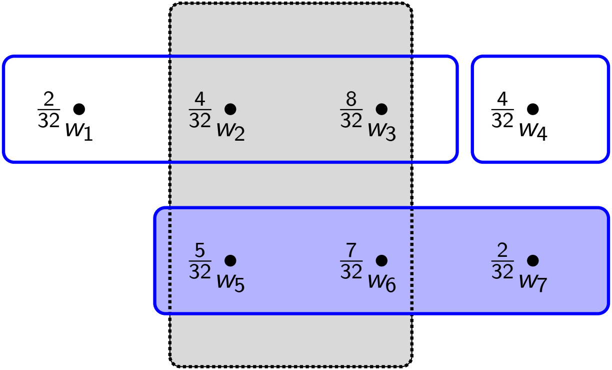
Also, that *if the true state is* w_1, then Experiment 1 will yield $P(E|I) = \frac{P(E \cap I)}{P(I)} = \frac{12}{14}$

2 Scientists Perform an Experiment



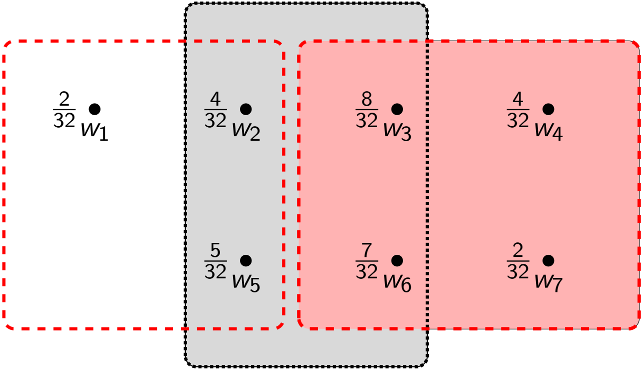
Suppose the true state is w_7 and the agents preform the experiments.

2 Scientists Perform an Experiment



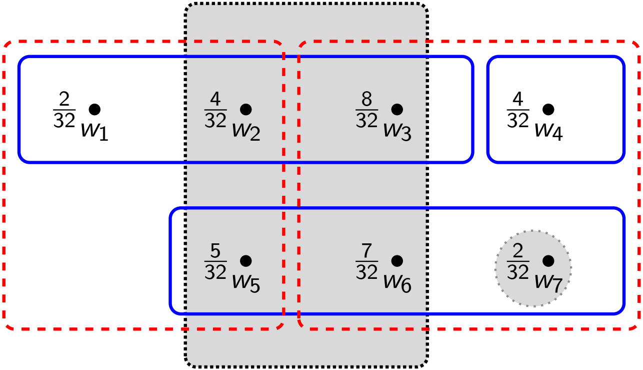
Suppose the true state is w_7, then $Pr_1(E) = \frac{12}{14}$

2 Scientists Perform an Experiment



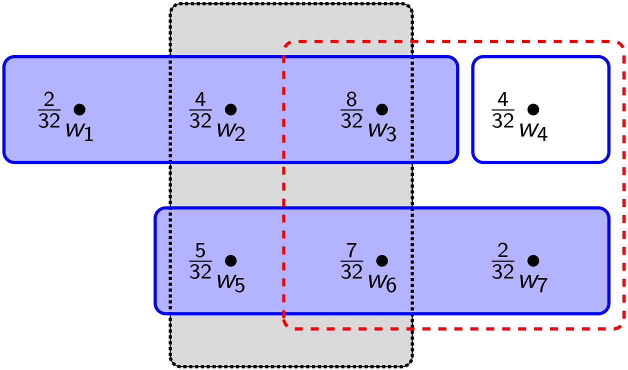
Then $Pr_1(E) = \frac{12}{14}$ and $Pr_2(E) = \frac{15}{21}$

2 Scientists Perform an Experiment



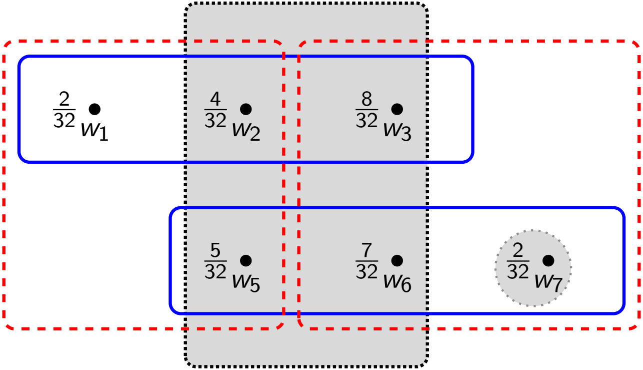
Suppose they exchange emails with the new subjective probabilities: $Pr_1(E) = \frac{12}{14}$ and $Pr_2(E) = \frac{15}{21}$

2 Scientists Perform an Experiment



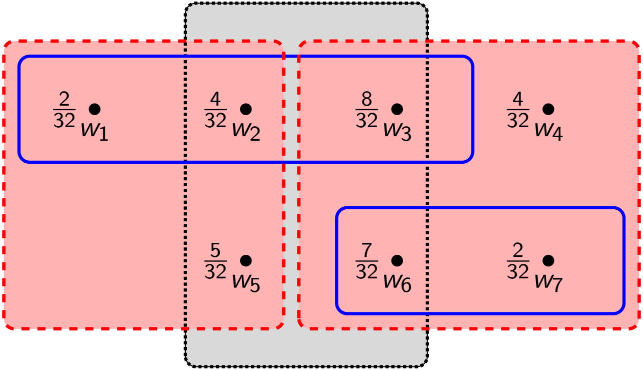
Agent 2 learns that w_4 is **NOT** the true state (same for Agent 1).

2 Scientists Perform an Experiment



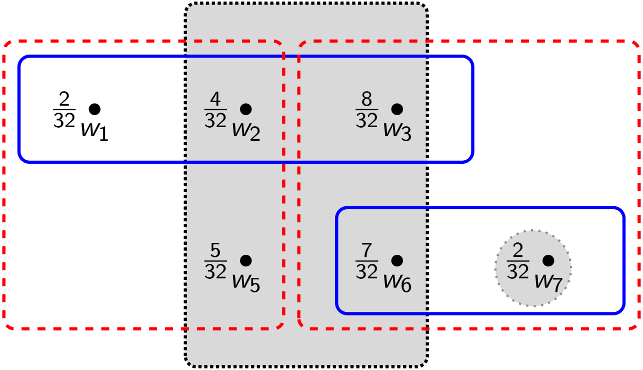
Agent 2 learns that w_4 is **NOT** the true state (same for Agent 1).

2 Scientists Perform an Experiment



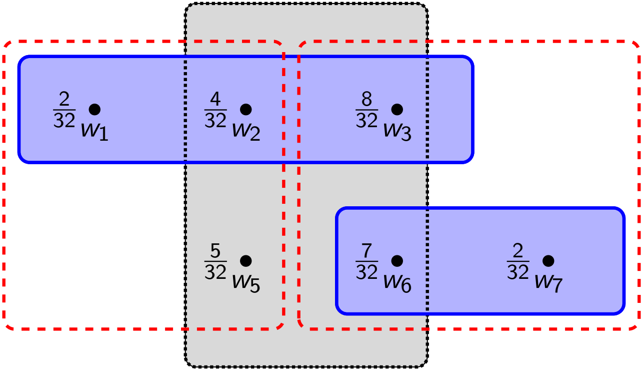
Agent 1 learns that w_5 is **NOT** the true state (same for Agent 1).

2 Scientists Perform an Experiment



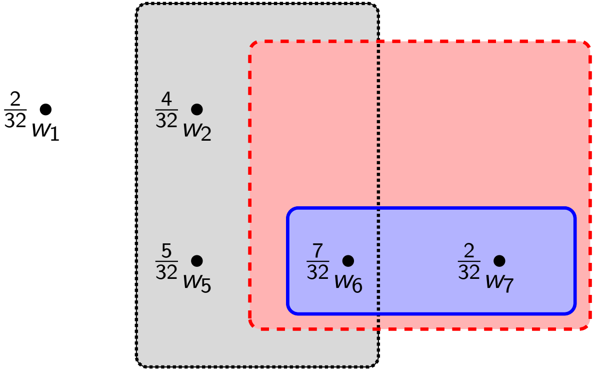
The new probabilities are $Pr_1(E|I') = \frac{7}{9}$ and $Pr_2(E|I') = \frac{15}{17}$

2 Scientists Perform an Experiment



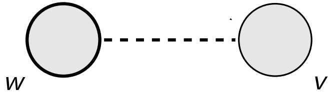
After exchanging this information ($Pr_1(E|I') = \frac{7}{9}$ and $Pr_2(E|I') = \frac{15}{17}$), Agent 2 learns that w_3 is **NOT** the true state.

2 Scientists Perform an Experiment



No more revisions are possible and the agents agree on the posterior probabilities.

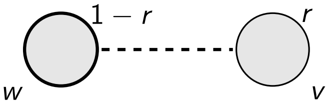
Models of Hard and Soft Information



$\mathcal{M} = \langle W, \{\Pi_i\}_{i \in \mathcal{A}} \rangle$
Π_i is agent i's partition with $\Pi_i(w)$ the partition cell containing w.

$K_i(E) = \{w \mid \Pi_i(w) \subseteq E\}$

Models of Hard and Soft Information



$\mathcal{M} = \langle W, \{\Pi_i\}_{i\in\mathcal{A}}, \{p_i\}_{i\in\mathcal{A}}\rangle$
for each i, $p_i : W \to [0,1]$ is a probability measure

$B^p(E) = \{w \mid p_i(E \mid \Pi_i(w)) = \frac{\pi_i(E \cap \Pi_i(w))}{p_i(\Pi_i(w))} \geq p\}$

1. $B_i^p(B_i^p(E)) = B_i^p(E)$

2. If $E \subseteq F$ then $B_i^p(E) \subseteq B_i^p(F)$

3. $\pi(E \mid B_i^p(E)) \geq p$

Common p-belief

The typical example of an event that creates common knowledge is a **public announcement**.

Common p-belief

The typical example of an event that creates common knowledge is a **public announcement**.

Shouldn't one always allow for some small probability that a participant was absentminded, not listening, sending a text, checking facebook, proving a theorem, asleep, ...

Common *p*-belief

The typical example of an event that creates common knowledge is a **public announcement**.

Shouldn't one always allow for some small probability that a participant was absentminded, not listening, sending a text, checking facebook, proving a theorem, asleep, ...

"We show that the weaker concept of "common belief" can function successfully as a substitute for common knowledge in the theory of equilibrium of Bayesian games."

D. Monderer and D. Samet. *Approximating Common Knowledge with Common Beliefs*. Games and Economic Behavior (1989).

Common p-belief: definition

$$B_i^p(E) = \{w \mid p(E \mid R_i(w)) \geq p\}$$

Common p-belief: definition

$$B_i^p(E) = \{w \mid p(E \mid R_i(w)) \geq p\}$$

An event E is **evident p-belief** if for each $i \in \mathcal{A}$, $E \subseteq B_i^p(E)$

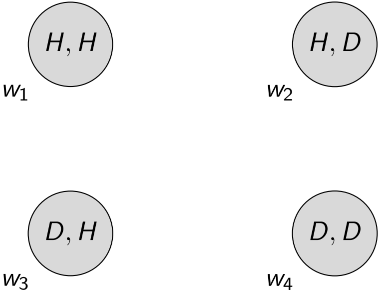
Common p-belief: definition

$$B_i^p(E) = \{w \mid p(E \mid R_i(w)) \geq p\}$$

An event E is **evident p-belief** if for each $i \in \mathcal{A}$, $E \subseteq B_i^p(E)$

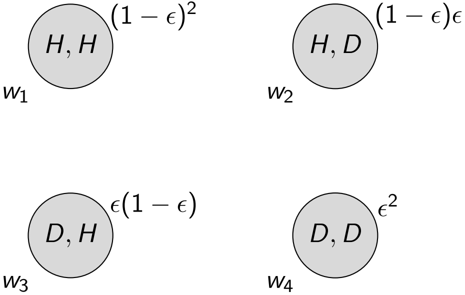
An event F is **common p-belief** at w if there exists and evident p-belief event E such that $w \in E$ and for all $i \in \mathcal{A}$, $E \subseteq B_i^p(F)$

Common p-belief: example



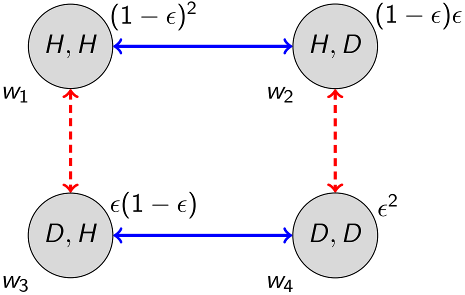
Two agents either hear (H) or don't hear (D) the announcement.

Common p-belief: example



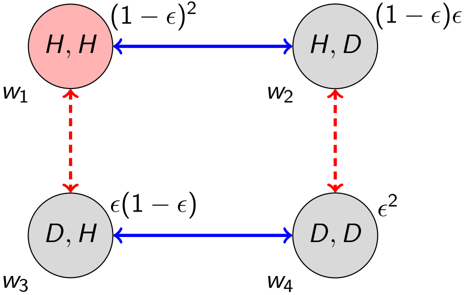
The probability that an agent hears is $1 - \epsilon$.

Common p-belief: example



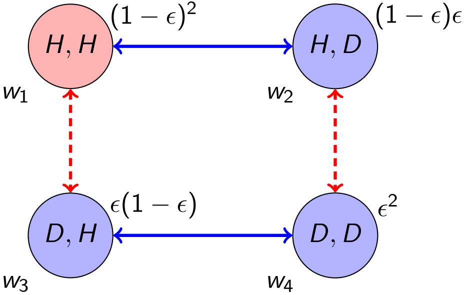
The agents *know* their "type".

Common p-belief: example



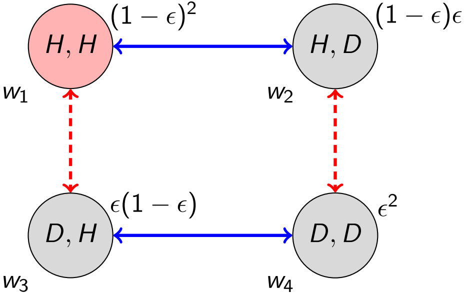
The event "everyone hears" ($E = \{w_1\}$)

Common p-belief: example



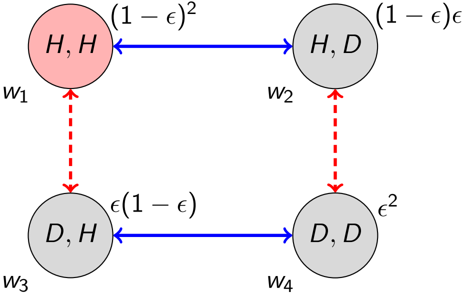
The event "everyone hears" ($E = \{w_1\}$) is **not** common knowledge

Common p-belief: example



The event "everyone hears" ($E = \{w_1\}$) is **not** common knowledge, but it is common $(1-\epsilon)$-belief

Common p-belief: example



The event "everyone hears" ($E = \{w_1\}$) is **not** common knowledge, but it is common $(1-\epsilon)$-belief:
$B_i^{(1-\epsilon)}(E) = \{w \mid p(E \mid \Pi_i(w)) \geq 1-\epsilon\} = \{w_1\} = E$, for $i = 1, 2$

Common p-belief

Theorem. If the posteriors of an event X are common p-belief at some state w, then any two posteriors can differ by at most $1 - p$.

D. Samet and D. Monderer. *Approximating Common Knowledge with Common Beliefs*. Games and Economic Behavior, Vol. 1, No. 2, 1989.