Carnegie Mellon University

Topological Semantics for Asynchronous Common Knowledge

Siddharth Namachivayam

Department of Philosophy (advised by Adam Bjorndahl and Kevin Kelly)

Motivation: Byzantine Generals

Two generals Alice and Bob are deciding whether to attack an enemy.

- The conditions may or may not be favorable to attack.
- If the conditions are favorable to attack, Alice and Bob want to simultaneously attack in order to win the battle.
- If the conditions are not favorable to attack, or only one of the generals attacks, then anyone who attacks has their army obliterated.

The generals want to find joint strategies that are guaranteed to protect both of them from losing their respective armies.

Motivation: Learning Game

Two agents Alice and Bob are deciding whether to report a proposition $P \subseteq \Omega$ is true or defer judgement.

- If P is true, Alice and Bob want to both eventually converge on reporting P is true.
- If P is false, or only one of Alice or Bob converges on reporting P is true, then anyone who converges on reporting P is true gets perpetually shamed.

The agents want to find joint strategies that are guaranteed to protect both of them from being perpetually shamed.

Motivation: Setup

- Denote Alice and Bob by a and b.
- Ω is a countable set of possible worlds.
- $\mathscr{B}_a$ and $\mathscr{B}_b$ represent the possible *information states* of Alice and Bob.
- Both agents have perfect recall.
- So, $\mathscr{B}_a$ and $\mathscr{B}_b$ are topological bases over Ω.
- Alice and Bob have non-zero prior μ over Ω.
- Alice and Bob also have *retraction budget* $n \in \mathbb{N}$.
- That is, they are willing to switch their answer from Yes to deferring judgment at most n times.

Carnegie Mellon University

Motivation: Strategies

- A *strategy* for $i \in N = \{a, b\}$ is a map $s_i : \mathscr{B}_i \to \{\text{Yes}, ?\}$.
- An m*-retraction sequence* for s_i is a finite downward sequence $B_1 \supseteq B_2 \ldots \supseteq B_{2m}$ of information states in $\mathscr{B}_i$ such that $s_i(B_{2k-1}) = \text{Yes}$ and $s_i(B_{2k}) = ?$.
- An n-retraction *strategy* for i has no m-retraction sequences for any $m > n$.
- So, Alice and Bob pick n-retraction strategies s_a and s_b.

Carnegie Mellon University

Motivation: Payoffs

- $\sigma(s_i, w)$ returns the output strategy s_i converges to in world w.
- Payoffs are given by $\sum_{w \in \Omega} u_{i|w}(s_a, s_b) \cdot \mu(w)$ where:

$$u_{i|w}((s_a, s_b)) = \begin{cases} 1 & w \in P \land \forall j \in N, \sigma(s_j, w) = \text{Yes}; \\ 0 & \sigma(s_i, w) = \text{?}; \\ -\infty & \sigma(s_i, w) = \text{Yes} \land (w \notin P \lor \exists j \in N, \sigma(s_j, w) = \text{?}) \,. \end{cases}$$

- Nash equilibria?

Carnegie Mellon University

Coordination

- *Coordination* corresponds to strategy profiles where Alice converges to Yes iff Bob does.
- All equilibria must be coordinated since otherwise at least one player would save money by never reporting P is true.
- Further, Alice and Bob must be correct in the limit when they converge to Yes otherwise they would again save money by never reporting P is true.
- So, in equilibrium we have *protected coordination*:

$$\sigma(s_a, \cdot\,)^{-1}(\text{Yes}) = \sigma(s_b, \cdot\,)^{-1}(\text{Yes}) \subseteq P.$$

- That is also sufficient.

Carnegie Mellon University

Protected Coordination Theorem

Theorem: (s_a, s_b) is a Nash equilibrium iff $\sigma(s_a, \cdot\,)^{-1}(\text{Yes}) = \sigma^{-1}(s_b, \cdot\,)(\text{Yes}) \subseteq P$.

- **Proof:** Necessity is already proved.
- For sufficiency, let s_a and s_b be strategies such that $\sigma(s_a, \cdot\,)^{-1}(\text{Yes}) = \sigma^{-1}(s_b, \cdot\,)(\text{Yes}) \subseteq P$.
- By hypothesis s_i has non-negative expectation for i.
- Now suppose for contradiction that Alice has a profitable deviation s'_a.
- Note s'_a has non-negative expectation for a.
- So, it must be $\sigma(s'_a, \cdot\,)^{-1}(\text{Yes}) \subseteq \sigma^{-1}(s_b, \cdot\,)(\text{Yes})$, else Alice would expect to receive $-\infty$.
- But that means $\sigma(s'_a, \cdot\,)^{-1}(\text{Yes}) \subseteq \sigma^{-1}(s_a, \cdot\,)(\text{Yes})$, so s'_a cannot be profitable. $\perp$. $\square$.

So Nash equilibria are exactly those joint strategies which are guaranteed to protect both agents from perpetual shame.

Carnegie Mellon University

0-Retraction Equilibria as Fixed Points

- There exists a 0-retraction strategy s_i so that $\sigma(s_i, \cdot\,)^{-1}(\text{Yes}) = X$ iff X is open in the topology $\mathcal{T}_i$ generated by $\mathcal{B}_i$.
- So, we can *characterize* 0-retraction equilibria by finding those subsets of P which are open in $\mathcal{T}_a \cap \mathcal{T}_b$.
- Let int_i denote the interior operator with respect to $\mathcal{T}_i$.
- Consider the map from $\mathcal{P}(\Omega)$ to itself which sends $X \mapsto \bigcap_{i \in N} \text{int}_i(P \cap X)$.

0-Retraction Equilibria as Fixed Points

- Fixed points of $X \mapsto \cap_{i \in N} \text{int}_i(P \cap X)$ are exactly those subsets of P which are open in $\mathscr{T}_a \cap \mathscr{T}_b$.
- Unions of open sets are open.
- So, the map has a greatest fixed point: the union of all open sets in $\mathscr{T}_a \cap \mathscr{T}_b$ which entail P.
- Take int_N to be the interior operator with respect to $\mathscr{T}_a \cap \mathscr{T}_b$.
- Thus, the greatest fixed point of $X \mapsto \cap_{i \in N} \text{int}_i(P \cap X)$ is $\text{int}_N P$.
- So, $\text{int}_N P$ characterizes *welfare maximizing* 0-retraction equilibria.

Carnegie Mellon University

Common Verifiability

- $\text{int}_N P$ is also epistemically interesting.
- Tarski's fixed point theorem implies that $\text{int}_N P$ corresponds to the transfinite conjunction of the statements:

$1)$ Everyone can verify P

$2)$ Everyone can verify $1.$

…

$k)$ Everyone can verify $k - 1.$

…

$\omega)$ $\forall k < \omega$, Everyone can verify k.

$\omega + 1)$ Everyone can verify ω.

Etc.

- Thus $\text{int}_N P$ corresponds to the worlds where P is "commonly verifiable."
- **Question:** Is P commonly verifiable iff P can become asynchronous common knowledge?

Carnegie Mellon University

Gonczarowski and Moses 2024

- Paradox: $\Omega = \{w_1, w_2, w_3\}$ and $P = \{w_1, w_2\}$ is decidable for Alice.
- Alice sends a message telling Bob P is true if she has verified it at time 0.
- No message is sent if she verifies $\neg P$ at time 0.
- The message is guaranteed to arrive in either w_1) 1 minute or w_2) 2 minutes.
- There is no global clock.
- Alice has a local clock which starts upon deciding P and Bob has a local clock which starts upon receiving a message from Alice.
- Clearly, P is not common knowledge at time 0 in both w_1 and w_2.
- More surprisingly, P is **never** common knowledge in either world!

Gonczarowski and Moses 2024

- Suppose for a contradiction P is common knowledge among Alice and Bob at some time in w_1.
- Common knowledge arises synchronously.
- So there must exist a first time t where P becomes common knowledge in w_1 that is the same for Alice and Bob.
- But since w_1 and w_2 are indistinguishable for both agents, this means the first time P becomes common knowledge in w_2 is $t+1$ for Bob and t for Alice. $\perp$.
- Thus P is never common knowledge in w_1.
- Symmetrically, we can argue P is never common knowledge in w_2.

Carnegie Mellon University

Gonczarowski and Moses 2024

- Why is this a paradox?
- Common knowledge is supposed to guide coordination.
- We typically think *successful coordination* is possible in some equilibrium iff P can become common knowledge.
- Since P cannot become common knowledge, successful coordination should not be possible.
- However, it is obvious how to successfully coordinate!
- Alice should report P is true upon verifying it and Bob should report P is true upon receiving the message.

Asynchronous Common Knowledge

- Solution: Allow common knowledge to arise *asynchronously*.
- Intuition: Alice knows "Bob knows P upon receiving the message" upon sending the message.
- G&M refine possible worlds $w \in \Omega$ into *states of affairs* $(w, t) \in \Omega \times \mathbb{N}$ at a particular time.
- Each agent $i \in N$ has an *information partition* $\mathscr{P}_i$ over $\Omega \times \mathbb{N}$.
- Fix $X \subseteq \Omega \times \mathbb{N}$.
- $\Diamond X$ and $\Box X$ denote the states of affairs where where X holds at some time and at all times.
- $\mathscr{P}_i(w, t)$ denotes the states of affairs that are indistinguishable at (w, t) for agent i.
- Write $(w, t), \mathscr{P}_i \vDash K_i^* X$ iff $\mathscr{P}_i(w, t) \subseteq X$.
- Say a *witness* $W_i \subseteq \Omega \times \mathbb{N}$ is i-*local* iff $K_i^* W_i = W_i$.
- Say a *witness profile* $\overline{W} \subseteq (\Omega \times \mathbb{N})^N$ is N-*local* iff $\forall i \in N,\ W_i$ is i-local.

Carnegie Mellon University

Asynchronous Common Knowledge

- Given an N-local witness profile $\overline{W}$, $K^*_{i@W_i}P$ denotes "i knows P upon certifying W_i", i.e.:

$$(w,t),\mathscr{P}_i \vDash K^*_{i@W_i}P \text{ iff } (w,t),\mathscr{P}_i \vDash \Diamond W_i \wedge \Box(W_i \rightarrow K^*_i P)$$

- Similarly, $E^*_{N@\overline{W}}P$ denotes "everyone knows P upon asynchronous certification of $\overline{W}$", i.e.:

$$(w,t),(\mathscr{P}_i)_{i\in N} \vDash E^*_{N@\overline{W}}P \text{ iff } (w,t),\mathscr{P}_i \vDash K^*_{i@W_i}P \text{ for all } i \in N$$

- Finally, $C^*_{N@\overline{W}}P$ denotes "P is common knowledge upon asynchronous certification of $\overline{W}$", i.e.:

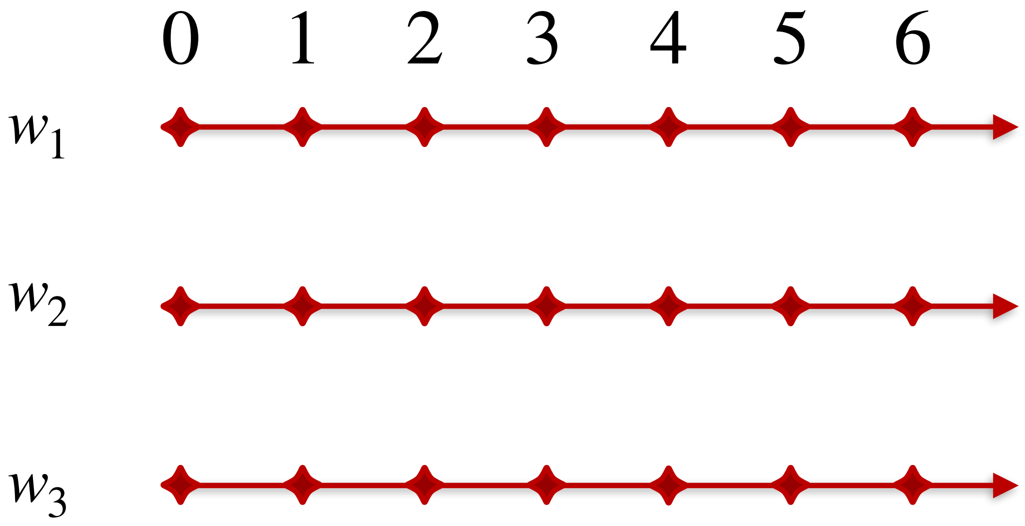
$$(w,t),(\mathscr{P}_i)_{i\in N} \vDash M^{*1}_{N@\overline{W}}P \text{ iff } (w,t),(\mathscr{P}_i)_{i\in N} \vDash E^*_{N@\overline{W}}P$$

$$(w,t),(\mathscr{P}_i)_{i\in N} \vDash M^{*k}_{N@\overline{W}}P \text{ iff } (w,t),(\mathscr{P}_i)_{i\in N} \vDash E^*_{N@\overline{W}}M^{*k-1}_{N@\overline{W}}P$$

$$(w,t),(\mathscr{P}_i)_{i\in N} \vDash C^*_{N@\overline{W}}P \text{ iff } (w,t),(\mathscr{P}_i)_{i\in N} \vDash M^{*k}_{N@\overline{W}}P \text{ for all } k \in \mathbb{N}\backslash\{0\}$$

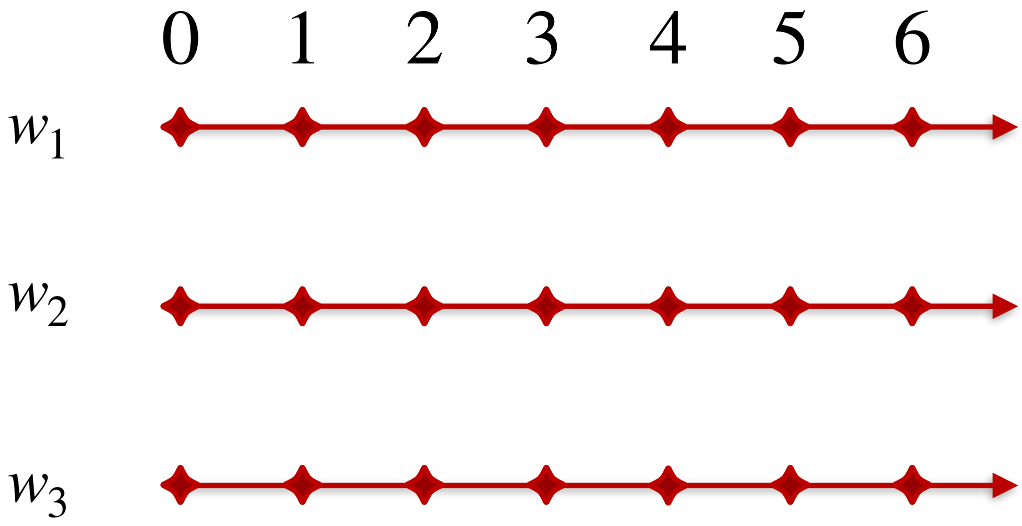
Carnegie Mellon University

Example



Example

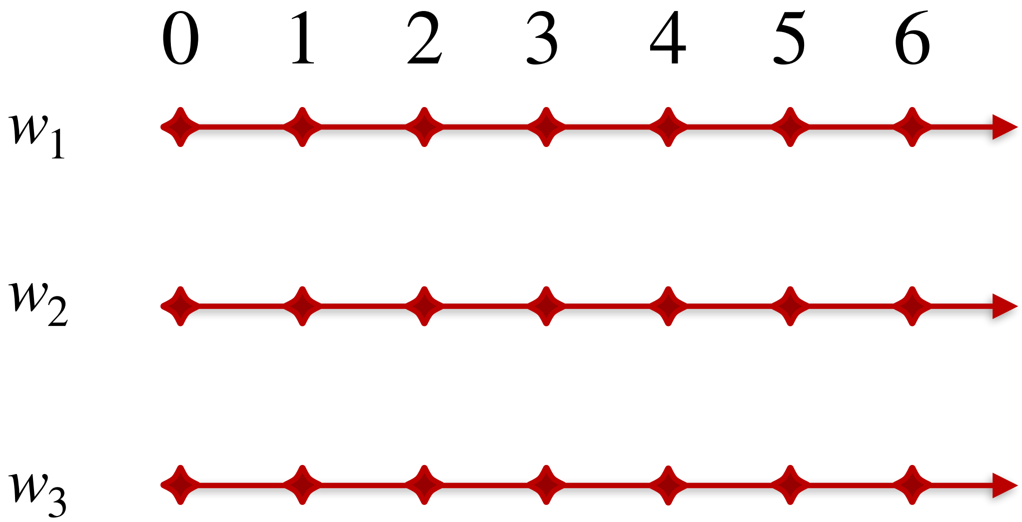
Alice verifies P and message arrives in 1 min



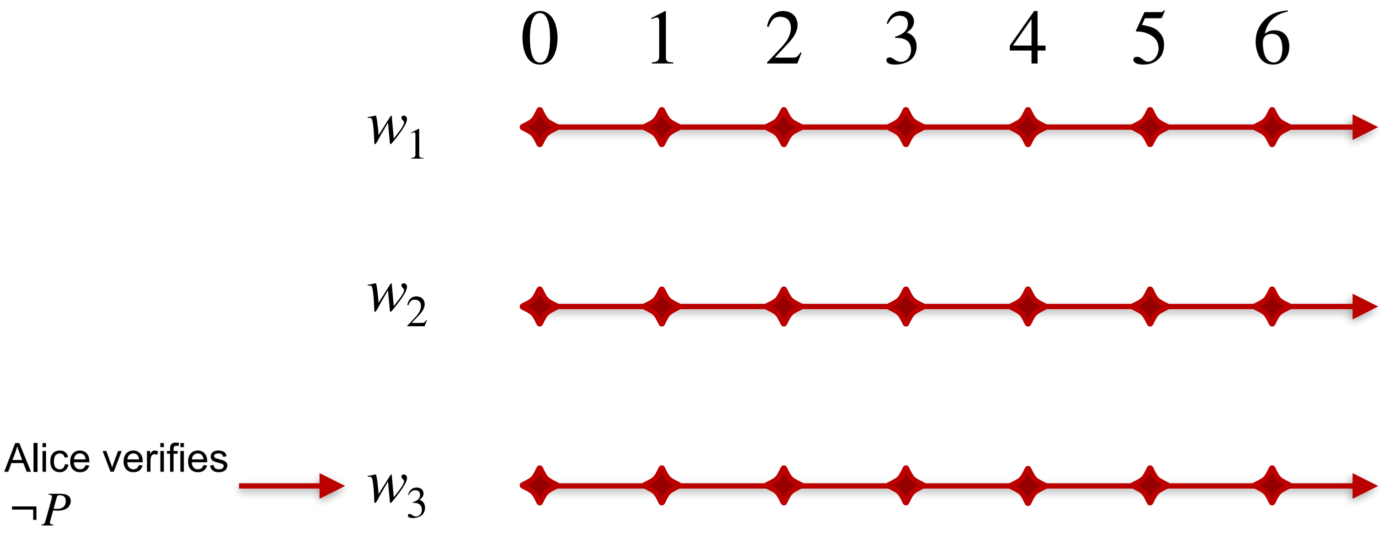
Carnegie Mellon University

Example

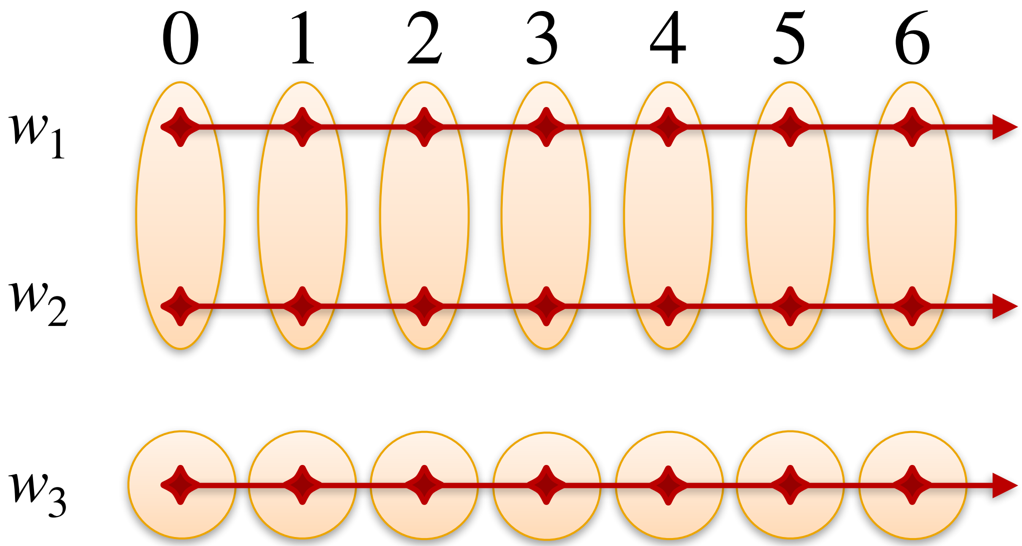
Alice verifies P
and message
arrives in 2 min



Example

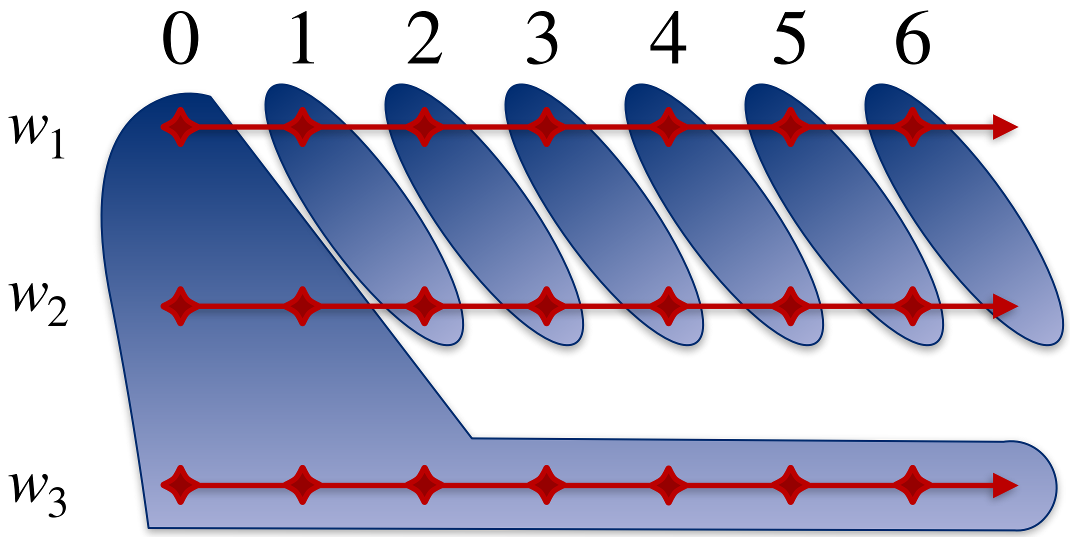


Example- Alice's Partition



Clock starts upon deciding P

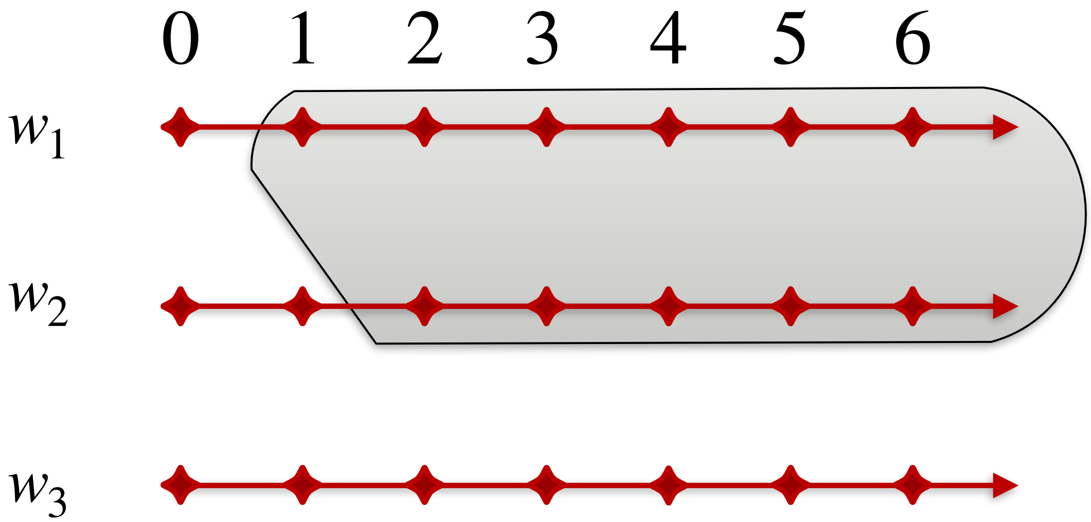
Example- Bob's Partition



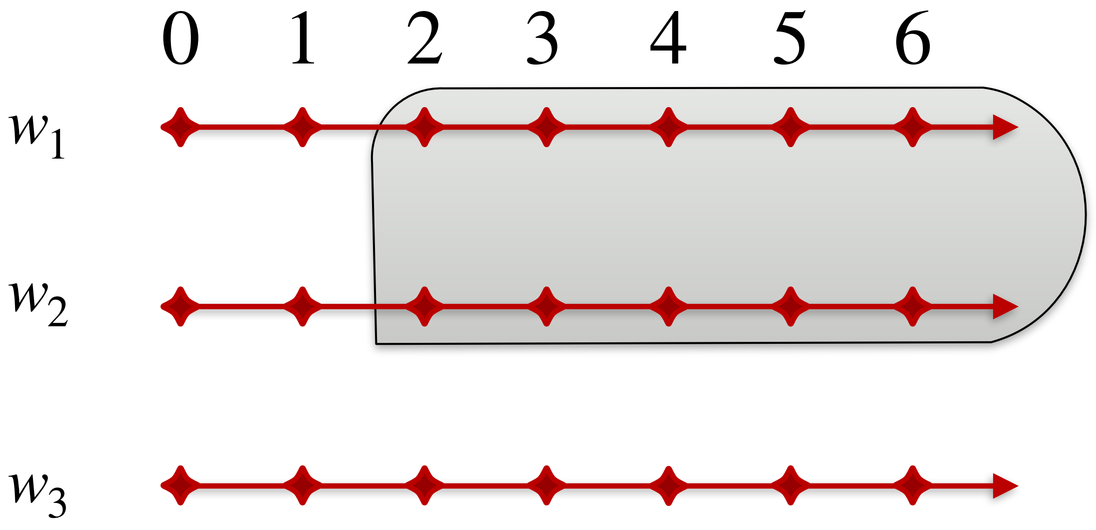
Clock starts upon receiving message

Carnegie Mellon University

Example- Everyone Knows P

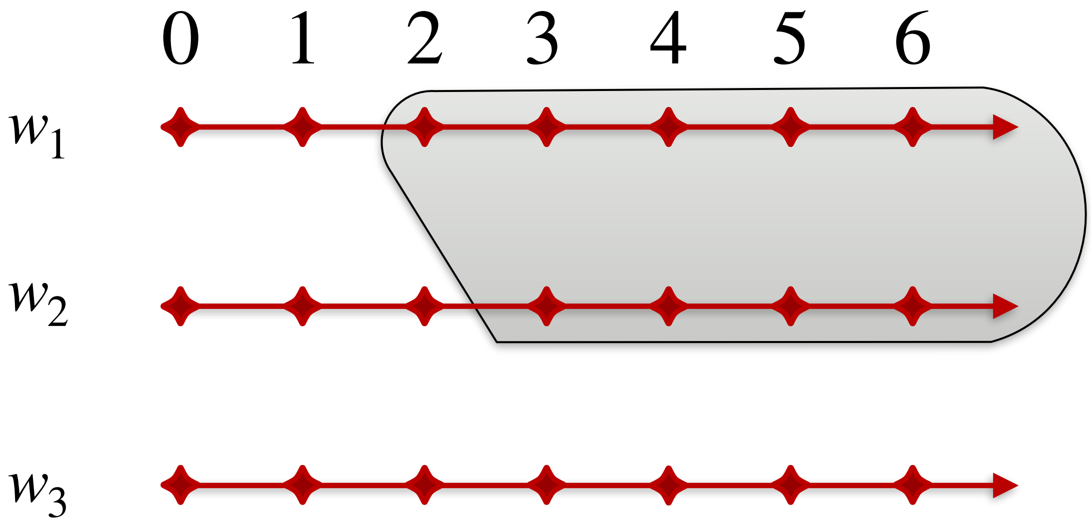


Example- (Everyone Knows)^2 P



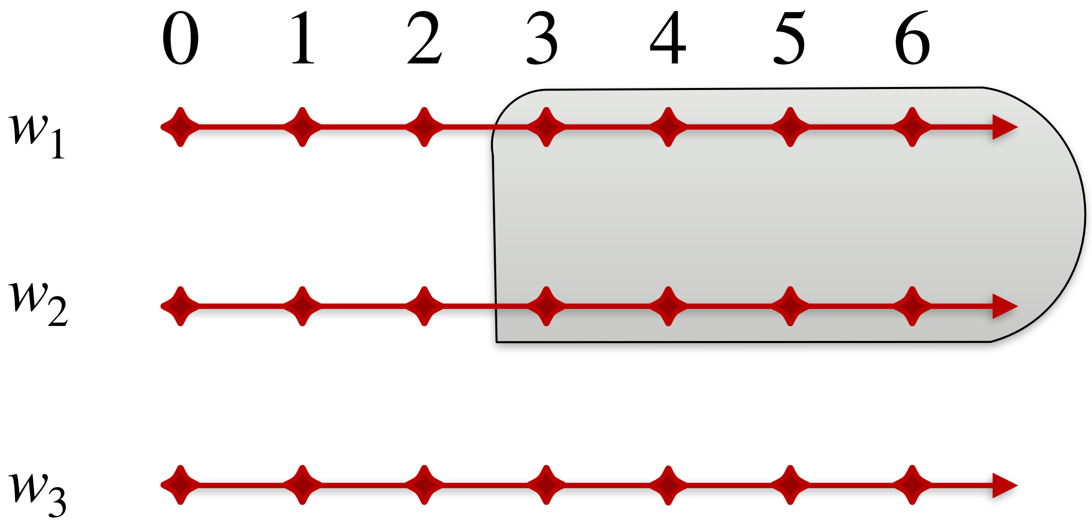
Carnegie Mellon University

Example- (Everyone Knows)³ P

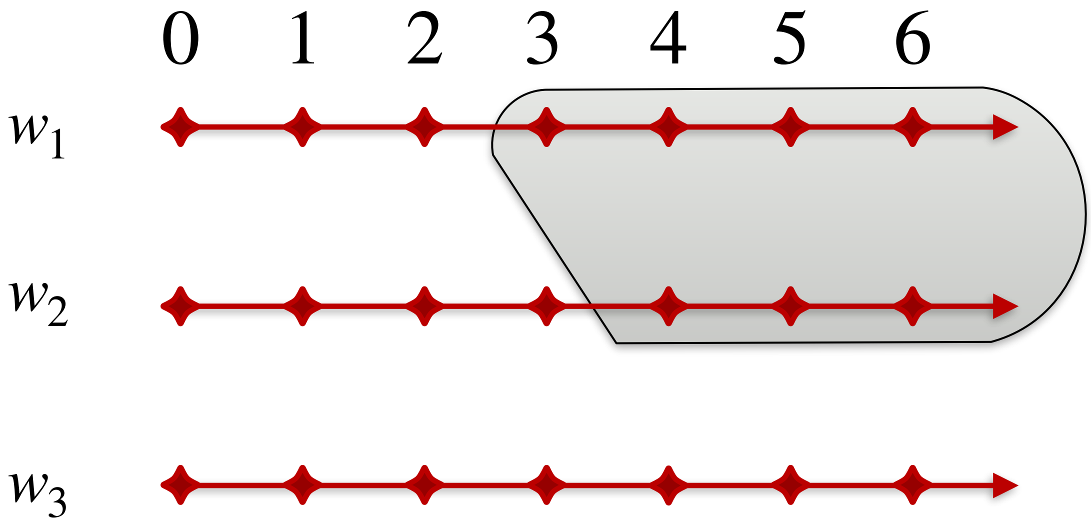


Carnegie Mellon University

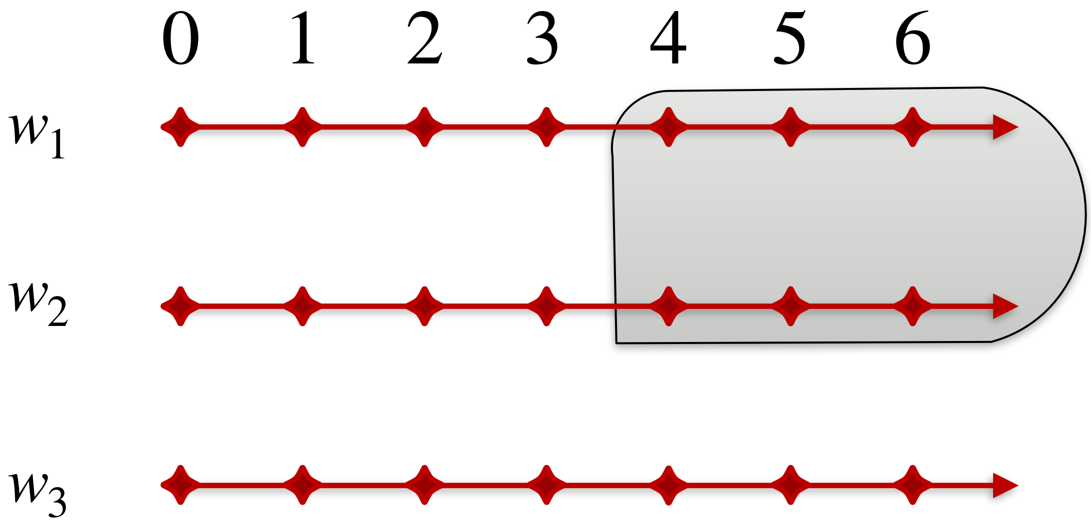
Example- (Everyone Knows)^4 P



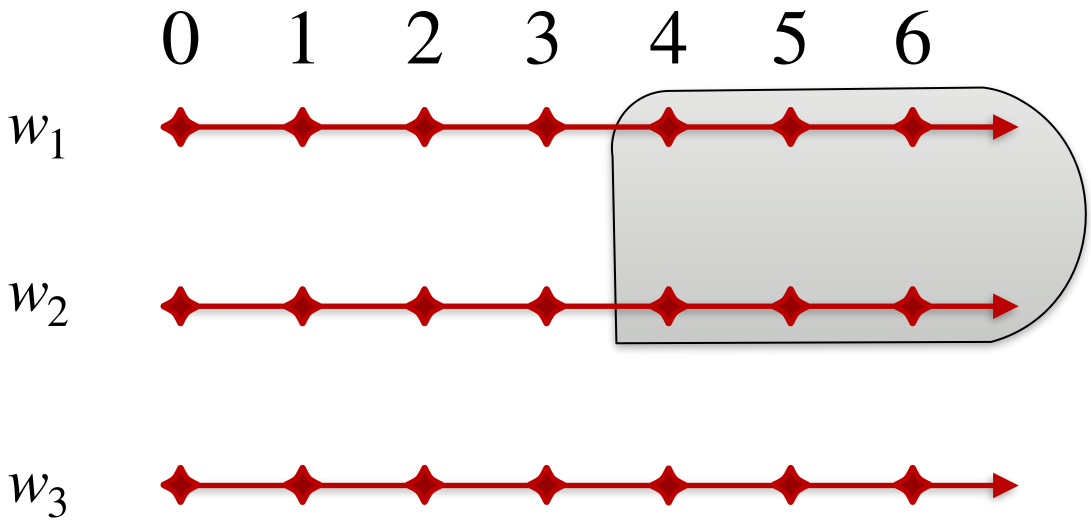
Example- (Everyone Knows)^5 P



Example- (Everyone Knows)^6 P

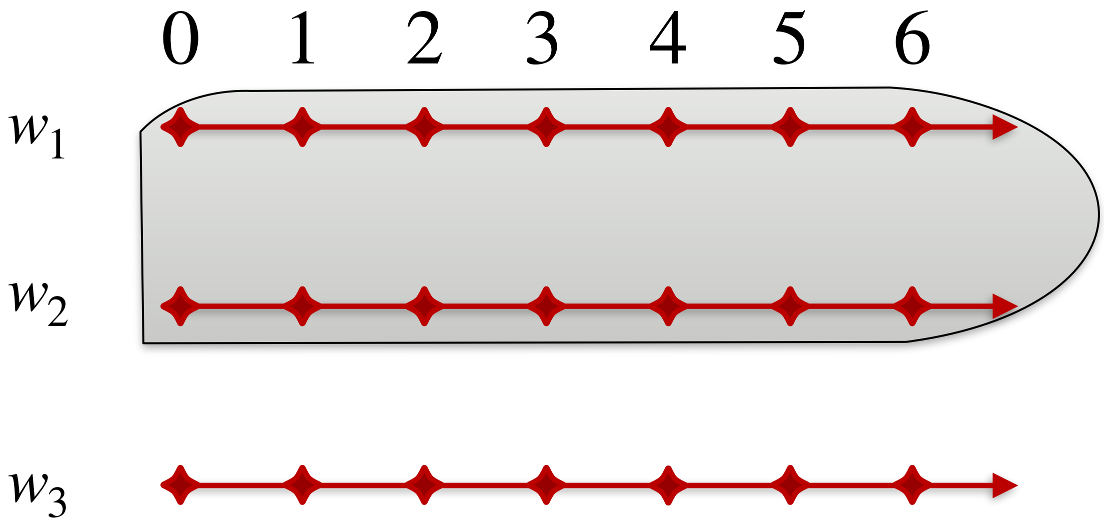


Example- (Everyone Knows)^6 P



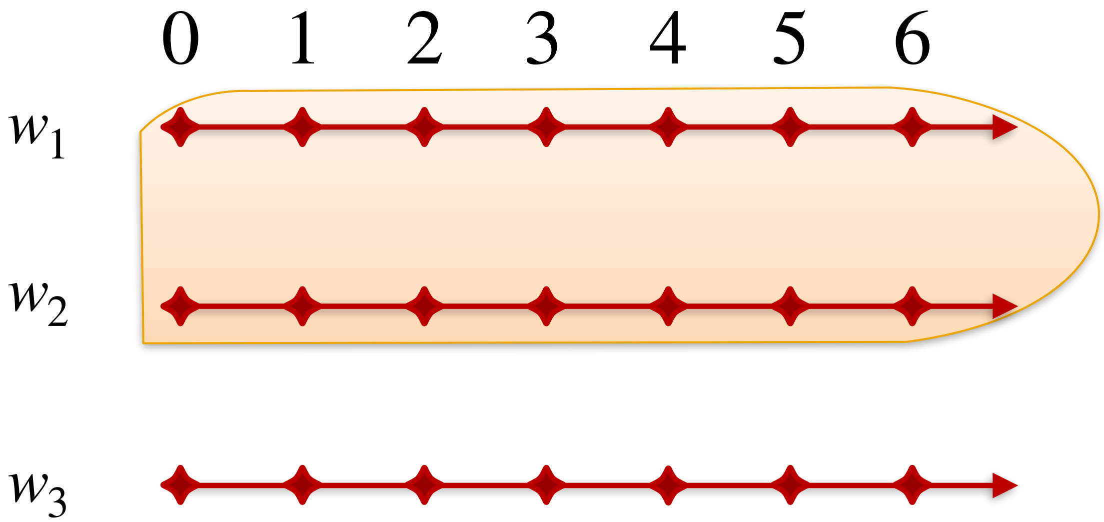
There are no states of affairs where P is common knowledge!

Example- W_a: Alice Sent The Message



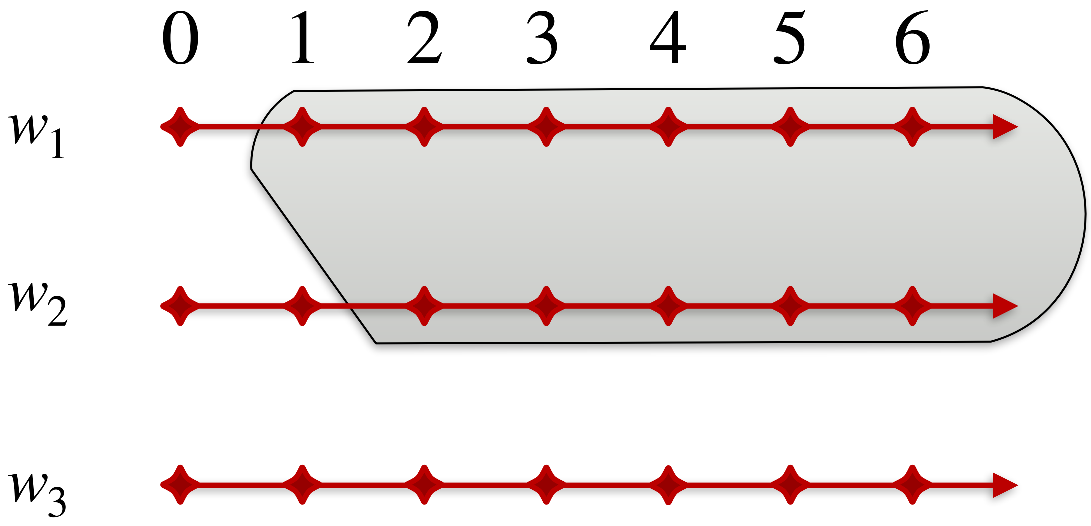
Carnegie Mellon University

Example- W_a: Alice Sent The Message



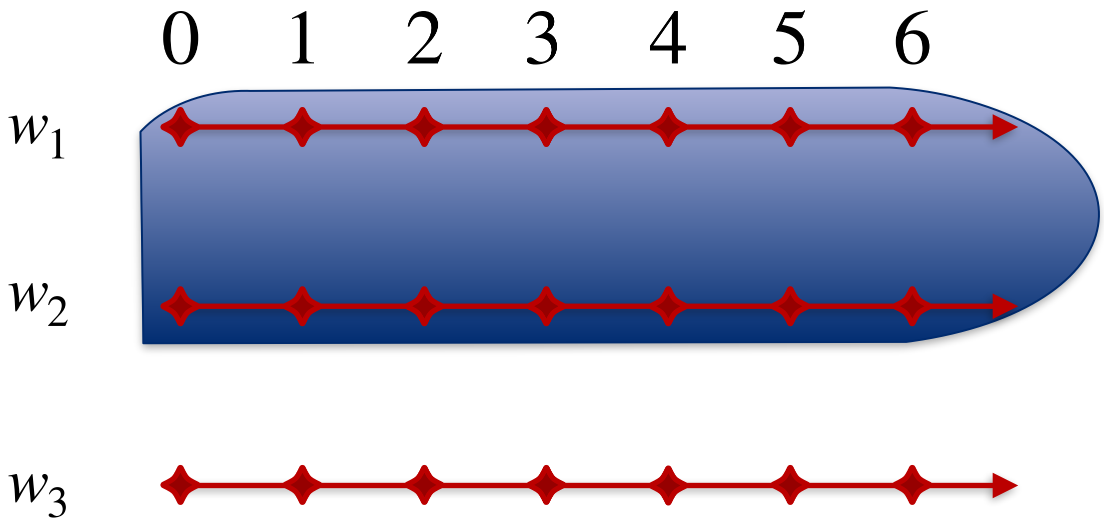
$K^*_{a@W_a}P$: Alice knows P upon certifying she sent the message

Example- W_b: Bob Received The Message



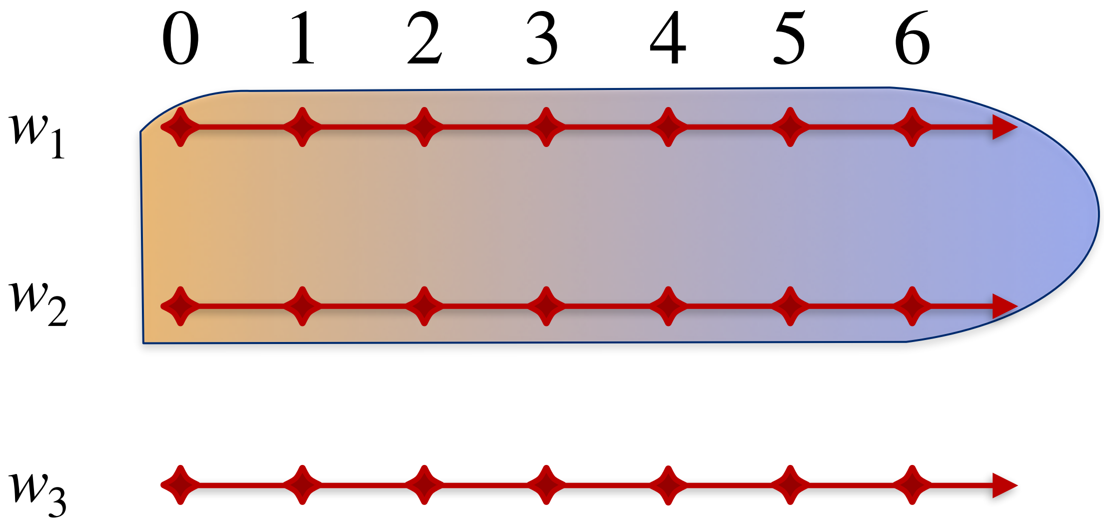
Carnegie Mellon University

Example- W_b: Bob Received The Message



$K^*_{b@W_b}P$: Bob knows P upon certifying he received the message

Example- $C^*_{N@\overline{W}}P$



$$K^*_{a@W_a}P = K^*_{b@W_b}P = P \rightarrow C^*_{N@\overline{W}}P = P$$

Detemporalizing G&M

- There is no need for explicit talk of time.
- Agents' payoffs do not depend on *when* their reports take place.
- Agents only care in what *worlds* they converge on reporting P is true.
- So, detemporalize G&M's semantics as follows:
- Take an agent $i \in N$ with information partition $\mathscr{P}_i$ over $\Omega \times \mathbb{N}$.
- $\forall k \in \mathscr{P}_i, \mathscr{C}_i(k) = \{w \in \Omega \mid \exists t \in \mathbb{N}, (w, t) \in k\}$ and set $\mathscr{C}_i = \{\mathscr{C}_i(k) \mid \forall k \in \mathscr{P}_i\}$.
- $\mathscr{C}_i$ is a *cover* over Ω.
- Elements of $\mathscr{C}_i$ correspond to information states.
- $\mathscr{C}_{i|w}$ is the collection of information states i passes through in world w.

Carnegie Mellon University

Detemporalizing G&M

- Given $X \subseteq \Omega$, $\text{cert}_i X$ denotes $\{w \mid \exists C \in \mathscr{C}_{i|w}, C \subseteq X\}$, i.e. the worlds where i at some information state certifies X.
- Given a profile $\overline{W} \subseteq \Omega^N$, define the following operators:

$$w, \mathscr{C}_i \vDash K_{i@W_i}P \text{ iff } w \in \text{cert}_i W_i \cap \{w' \mid \forall C \in \mathscr{C}_{i|w'}, C \subseteq W_i \to C \subseteq P\}$$

$$w, (\mathscr{C}_i)_{i \in N} \vDash E_{N@\overline{W}}P \text{ iff } w, \mathscr{C}_i \vDash K_{i@W_i}P \text{ for all } i \in N$$

$$w, (\mathscr{C}_i)_{i \in N} \vDash M^1_{N@\overline{W}}P \text{ iff } w, (\mathscr{C}_i)_{i \in N} \vDash E_{N@\overline{W}}P$$

$$w, (\mathscr{C}_i)_{i \in N} \vDash M^k_{N@\overline{W}}P \text{ iff } w, (\mathscr{C}_i)_{i \in N} \vDash E_{N@\overline{W}}M^{k-1}_{N@\overline{W}}P$$

$$w, (\mathscr{C}_i)_{i \in N} \vDash C_{N@\overline{W}}P \text{ iff } w, (\mathscr{C}_i)_{i \in N} \vDash M^k_{N@\overline{W}}P \text{ for all } k \in \mathbb{N} \backslash \{0\}$$

Carnegie Mellon University

Detemporalizing G&M

Indeed, if $\forall X \subseteq \Omega$ we let $[X]$ denote $X \times \mathbb{N}$, then we have the following translation:

$$w, \mathscr{C}_i \vDash K_{i@W_i}P \text{ iff } \exists t \in \mathbb{N} \text{ so that } (w,t), \mathscr{P}_i \vDash K^*_{i@K_i^*[W_i]}[P]$$

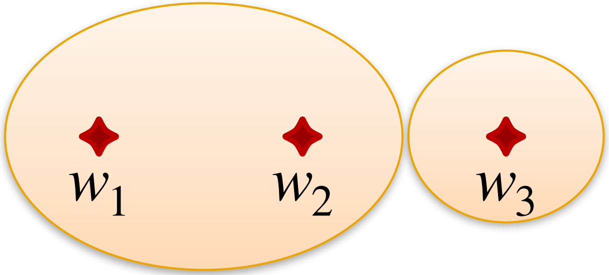
$$w, (\mathscr{C}_i)_{i\in N} \vDash E_{N@\overline{W}}P \text{ iff } \exists t \in \mathbb{N} \text{ so that } (w,t), (\mathscr{P}_i)_{i\in N} \vDash E^*_{N@(K_i^*[W_i])_{i\in N}}[P]$$

$$w, (\mathscr{C}_i)_{i\in N} \vDash M^k_{N@\overline{W}}P \text{ iff } \exists t \in \mathbb{N} \text{ so that } (w,t), (\mathscr{P}_i)_{i\in N} \vDash M^{*k}_{N@(K_i^*[W_i])_{i\in N}}[P]$$

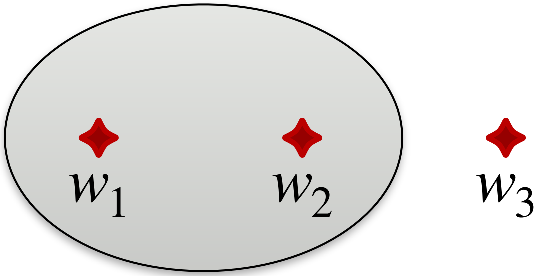
$$w, (\mathscr{C}_i)_{i\in N} \vDash C_{N@\overline{W}}P \text{ iff } \exists t \in \mathbb{N} \text{ so that } (w,t), (\mathscr{P}_i)_{i\in N} \vDash C^*_{N@(K_i^*[W_i])_{i\in N}}[P]$$

Carnegie Mellon University

Example- Alice's Cover

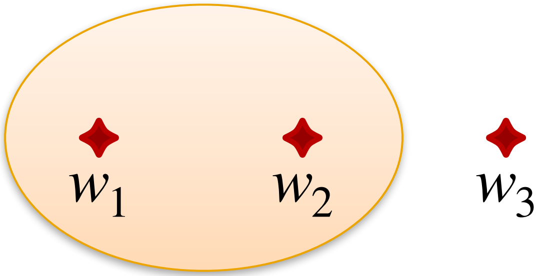


Example- W_a: Alice Sends The Message



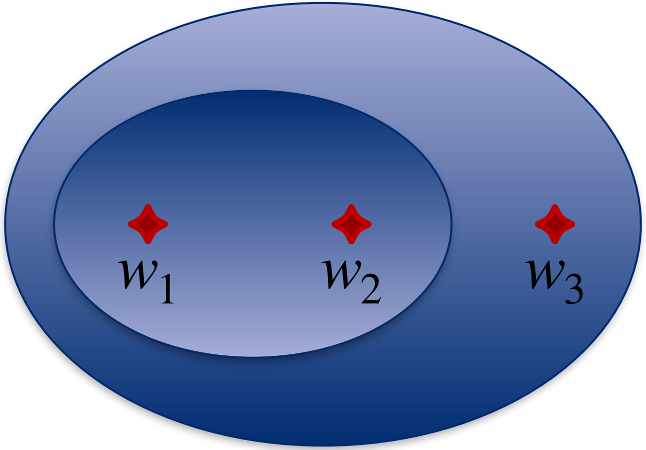
Carnegie Mellon University

Example- W_a: Alice Sends The Message

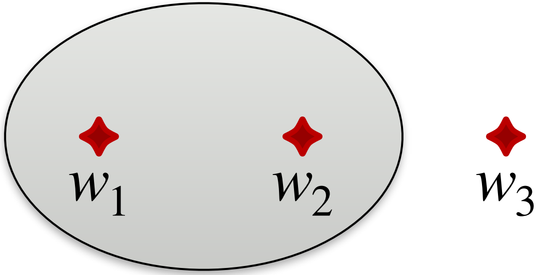


$K_{a@W_a}P$: Alice knows P upon certifying she sends the message

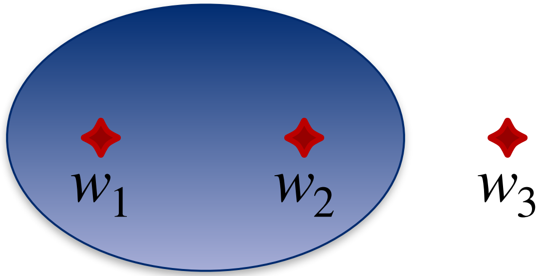
Example- Bob's Cover



Example- W_b: Bob Receives The Message

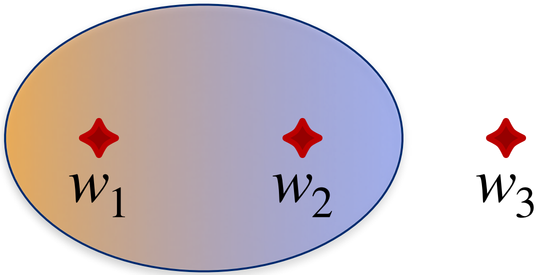


Example- W_b: Bob Receives The Message



$K_{b@W_b}P$: Bob knows P upon certifying he receives the message

Example- $C_{N@\overline{W}}P$



$$K_{a@W_a}P = K_{b@W_b}P = P \rightarrow C_{N@\overline{W}}P = P$$

ACK$_N P$

- Is P commonly verifiable iff P can become asynchronous common knowledge?
- We can capture the worlds where P can become asynchronous common knowledge by defining:

$w, (\mathscr{C}_i)_{i \in N} \vDash \text{ACK}_N P$ iff $\exists \overline{W} \subseteq \Omega^N$ so that $w, (\mathscr{C}_i)_{i \in N} \vDash C_{N @ \overline{W}} P$

- So, our question is whether $\text{int}_N P = \text{ACK}_N P$.
- However, our question is still not well formed!

Carnegie Mellon University

$\mathsf{cert}_N P$

- Why?
- $\mathscr{C}_i$ is a cover, not a basis.
- So, interior isn't defined.
- But, $\mathrm{cert}_i P$ naturally coincides with $\mathrm{int}_i P$ when $\mathscr{C}_i$ is a basis.
- So, we can define $\mathrm{cert}_N P$ as the greatest fixed point of $X \mapsto \cap_{i \in N} \mathrm{cert}_i(P \cap X)$.
- Does it hold that $\mathrm{cert}_N P = \mathrm{ACK}_N P$?

Carnegie Mellon University

$\mathbf{cert}_N P = \mathbf{ACK}_N P$

- Yes!
- So, P is commonly *certifiable* iff P can become asynchronous common knowledge.
- In summary, denote the map $X \mapsto \{C_{N@\overline{W}} X \mid \overline{W} \subseteq \Omega^N\}$ by $C_{N@\_}$.
- The following diagram commutes:

$$\begin{tikzcd}
\mathscr{P}(\Omega) \arrow[r, "C_{N@_}"] \arrow[dr, "\mathrm{cert}_N"'] & \mathscr{P}^2(\Omega) \arrow[d, "\cup"] \\
& \mathscr{P}(\Omega)
\end{tikzcd}$$

Carnegie Mellon University

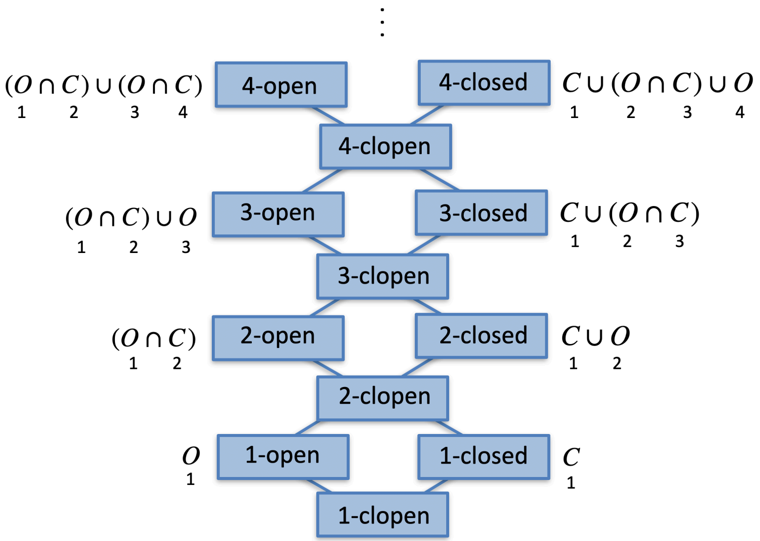
0-Retraction Equilibria Revisited

- Agents have perfect recall in our game.
- So each agent's cover is actually a basis $\mathscr{B}_i$.
- Fact: X is a subset of P which is open in each $\mathscr{T}_i$ iff $\exists \overline{W} \subseteq \Omega^N$ so that $C_{N@\overline{W}}P = X$.
- So, $C_{N@\_}P$ characterizes all equilibria in the 0-retraction game.
- Also, $\mathrm{int}_N P$ captures the worlds where agents can successfully coordinate in some equilibrium of the 0-retraction game.
- Can we do something similar for the n-retraction game when $n > 0$?

Carnegie Mellon University

n-Retraction Equilibria

- Fact: There exists a n-retraction strategy s_i so that $\sigma(s_i, \cdot)^{-1}(\text{Yes}) = X$ iff X is $2n+1$-*open* in the topology $\mathcal{T}_i$.
- So, we can characterize equilibria by finding subsets of P which are $2n+1$-open in each $\mathcal{T}_i$.
- But how do we find these sets?



(Image courtesy of Kevin Kelly)

Carnegie Mellon University

Asynchronous Common Belief

- $\text{Ver}_i^n(X \mid B)$ holds iff X is n-open in the subspace topology $\mathscr{T}_i | B$.
- So, i can decide X in n switches starting with no in light of evidence B.
- $\text{Ref}_i^n(X \mid B)$ holds iff X is n-closed in the subspace topology $\mathscr{T}_i | B$.
- So, i can decide X in n switches starting with yes in light of evidence B.
- Define $\text{Yes}_i^1(X \mid B)$ as $\text{Ref}_i^1(X \mid B) \wedge (X \cap B \neq \emptyset)$.
- Define $\text{Yes}_i^n(X \mid B)$ as $(\text{Ref}_i^n(X \mid B) \wedge \neg\text{Ver}_i^{n-1}(X \mid B)) \vee \text{Yes}_i^{n-1}(X \mid B)$.
- $\text{Yes}_i^n(X \mid B)$ expresses "i says yes to X in light of evidence B when their switching tolerance is n."

Carnegie Mellon University

Asynchronous Common Belief

- Given a witness profile $\overline{W} \subseteq \Omega^N$, $\mathrm{B}^n_{i@W_i}P$ denotes "i believes P upon n-affirming W_i", i.e.:

$$w, \mathscr{B}_i \vDash \mathrm{B}^n_{i@W_i}P \text{ iff } w \in \{w' \mid \exists B \in \mathscr{B}_{i|w'}, \mathrm{Yes}^n_i(W_i \mid B)\} \cap \{w' \mid \forall B \in \mathscr{B}_{i|w'}, \mathrm{Yes}^n_i(W_i \mid B) \to B \cap W_i \subseteq P\}$$

- Similarly, $\mathrm{EB}^n_{N@\overline{W}}P$ denotes "everyone believes P upon asynchronous n-affirmation of $\overline{W}$", i.e. :

$$w, (\mathscr{B}_i)_{i\in N} \vDash \mathrm{EB}^n_{N@\overline{W}}P \text{ iff } w, \mathscr{B}_i \vDash \mathrm{B}^n_{i@W_i}P \text{ for all } i \in N$$

- Next, $\mathrm{CB}^n_{N@\overline{W}}P$ denotes "P is common belief upon asynchronous n-affirmation of $\overline{W}$", i.e.:

$$w, (\mathscr{B}_i)_{i\in N} \vDash \mathrm{MB}^{n,1}_{N@\overline{W}}P \text{ iff } w, (\mathscr{B}_i)_{i\in N} \vDash \mathrm{EB}^n_{N@\overline{W}}P$$

$$w, (\mathscr{B}_i)_{i\in N} \vDash \mathrm{MB}^{n,k}_{N@\overline{W}}P \text{ iff } w, (\mathscr{B}_i)_{i\in N} \vDash \mathrm{EB}^n_{N@\overline{W}}\mathrm{MB}^{n,k-1}_{N@\overline{W}}P$$

$$w, (\mathscr{B}_i)_{i\in N} \vDash \mathrm{CB}^n_{N@\overline{W}}P \text{ iff } w, (\mathscr{B}_i)_{i\in N} \vDash \mathrm{MB}^{n,k}_{N@\overline{W}}P \text{ for all } k \in \mathbb{N}\backslash\{0\}$$

- Finally, $\mathrm{ACB}^n_N P$ denotes "P can become asynchronous common belief when everyone's switching tolerance is n", i.e.:

$$w, (\mathscr{B}_i)_{i\in N} \vDash \mathrm{ACB}^n_N P \text{ iff } \exists \overline{W} \subseteq \Omega^N \text{ so that } w, (\mathscr{B}_i)_{i\in N} \vDash \mathrm{CB}^n_{N@\overline{W}}P$$

Carnegie Mellon University

Asynchronous Common Knowledge

- Given a witness profile $\overline{W} \subseteq \Omega^N$, $K_{i@W_i}P$ denotes "i knows P upon certifying W_i", i.e.:

$$w, \mathscr{C}_i \vDash K_{i@W_i}P \text{ iff } w \in \text{cert}_i W_i \cap \{w' \mid \forall C \in \mathscr{C}_{i|w'}, C \subseteq W_i \to C \subseteq P\}$$

- Similarly, $E_{N@\overline{W}}P$ denotes "everyone knows P upon asynchronous certification of $\overline{W}$", i.e. :

$$w, (\mathscr{C}_i)_{i\in N} \vDash E_{N@\overline{W}}P \text{ iff } w, \mathscr{C}_i \vDash K_{i@W_i}P \text{ for all } i \in N$$

- Next, $C_{N@\overline{W}}P$ denotes "P is common knowledge upon asynchronous certification of $\overline{W}$", i.e.:

$$w, (\mathscr{C}_i)_{i\in N} \vDash M^1_{N@\overline{W}}P \text{ iff } w, (\mathscr{C}_i)_{i\in N} \vDash E_{N@\overline{W}}P$$

$$w, (\mathscr{C}_i)_{i\in N} \vDash M^k_{N@\overline{W}}P \text{ iff } w, (\mathscr{C}_i)_{i\in N} \vDash E_{N@\overline{W}}M^{k-1}_{N@\overline{W}}P$$

$$w, (\mathscr{C}_i)_{i\in N} \vDash C_{N@\overline{W}}P \text{ iff } w, (\mathscr{C}_i)_{i\in N} \vDash M^k_{N@\overline{W}}P \text{ for all } k \in \mathbb{N}\backslash\{0\}$$

- Finally, $\text{ACK}_N P$ denotes "P can become asynchronous common knowledge", i.e.:

$$w, (\mathscr{C}_i)_{i\in N} \vDash \text{ACK}_N P \text{ iff } \exists \overline{W} \subseteq \Omega^N \text{ so that } w, (\mathscr{C}_i)_{i\in N} \vDash C_{N@\overline{W}}P$$

Carnegie Mellon University

Asynchronous Common Belief

- Fact: X is a subset of P which is $n+1$-open in each $\mathscr{T}_i$ iff $\exists \overline{W} \subseteq \Omega^N$ so that $(\cap_{i \in N} W_i) \cap \mathrm{CB}^n_{N@\overline{W}} P = X$.
- Denote the map $X \mapsto \{(\cap_{i\in N} W_i) \cap \mathrm{CB}^n_{N@\overline{W}} X \mid \overline{W} \subseteq \Omega^N\}$ by $\mathbb{C}^n_{N@\_}$.
- $\mathbb{C}^{2n}_{N@\_} P$ characterizes all equilibria in the n-retraction game.
- Let $\mathscr{T}_i^{O\cap C}$ denote the topology generated by taking $n+1$-open sets in $\mathscr{T}_i$ as basis elements (the resulting topology is independent of n).
- INT_i is interior with respect to $\mathscr{T}_i^{O\cap C}$ and INT_N is interior with respect to $\cap_{i\in N} \mathscr{T}_i^{O\cap C}$.
- The following diagram commutes:

$$\begin{tikzcd}
\mathscr{P}(\Omega) \arrow[r, "\mathbb{C}^n_{N@_}"] \arrow[dr, "\mathrm{INT}_N"'] & \mathscr{P}^2(\Omega) \arrow[d, "\cup"] \\
 & \mathscr{P}(\Omega)
\end{tikzcd}$$

Switching Tolerance Invariance

- $\text{INT}_N P$ does not depend on n.
- Further, we can show $\text{INT}_N P = P \cap \text{ACB}^n_N P$.
- So the worlds where P can become true asynchronous common belief does not depend on agents' switching tolerances so long as $n > 0$?
- Shouldn't agents be able to successfully coordinate reporting on larger and larger sets in equilibrium as their retraction budgets increase?

Switching Tolerance Invariance

- Yes and yes.
- $\mathrm{INT}_N P$ does not necessarily *characterize* any equilibria.
- The worlds where P can become true asynchronous common belief are simply the worlds where agents can successfully coordinate in *some* equilibrium of the n-retraction game.
- But, there need not be any equilibria which successfully coordinate reporting in every world where P can become true asynchronous common belief.

Conclusion

- Topological bases provide a simple semantics for defining asynchronous common knowledge.
- Topological semantics also naturally admit a notion of asynchronous common belief.
- These concepts can be shown to characterize the equilibria of a Byzantine generals-like learning game.
- This tracks with our intuition that common knowledge ought to guide coordination!
- Future work should analyze the unbounded-retraction game and consider how to extend to when agents don't have perfect recall.
- Latter will require additional structure on covers…what's a retraction?

Q&A

Carnegie Mellon University