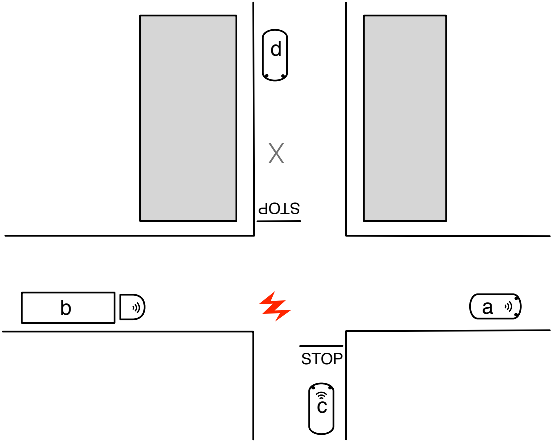
Second-Order Know-How Strategies

Pavel Naumov
Claremont McKenna College
Claremont, California, USA

Jia Tao
Lafayette College
Easton, Pennsylvania, USA

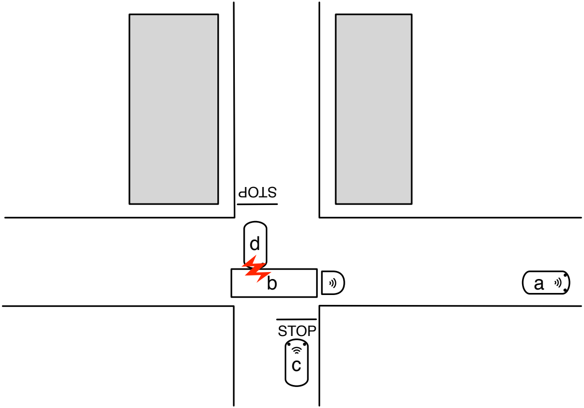
Consider the following traffic situation



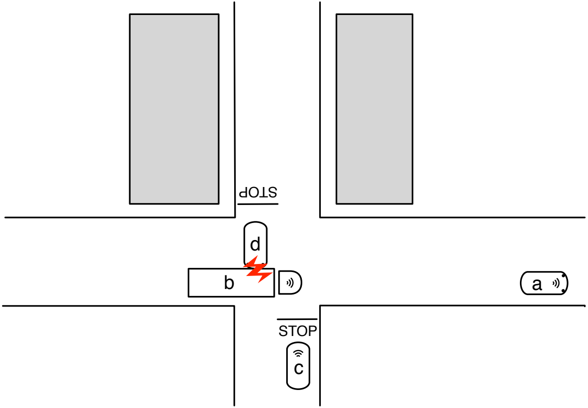
Assumption:

- The driver of car d does not notice the stop sign.

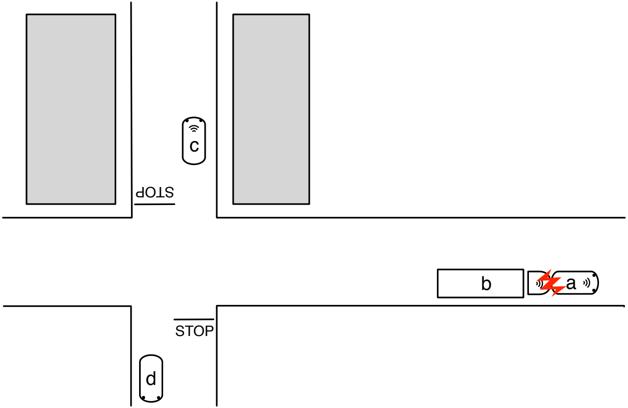
Truck b maintains the same speed



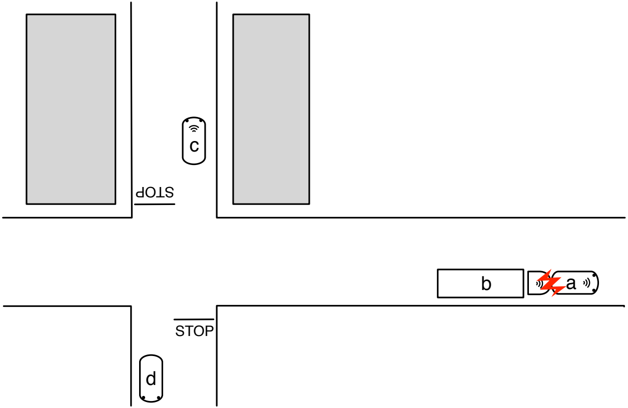
Truck b slows down



Truck b accelerates

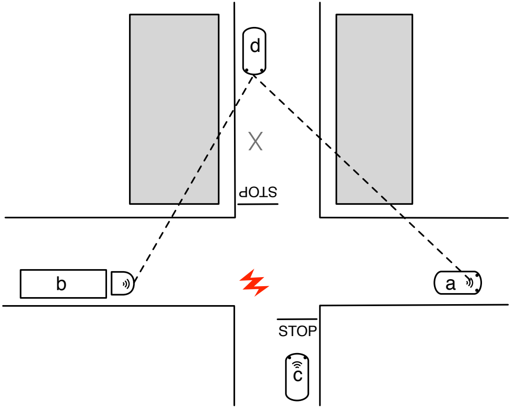


Truck b accelerates



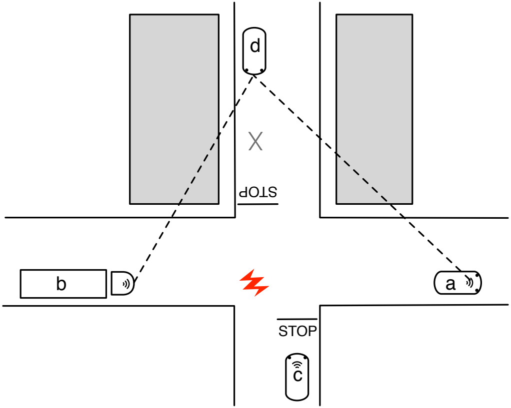
To prevent a collision, both truck b and car a must accelerate.

Second-order know-how strategies



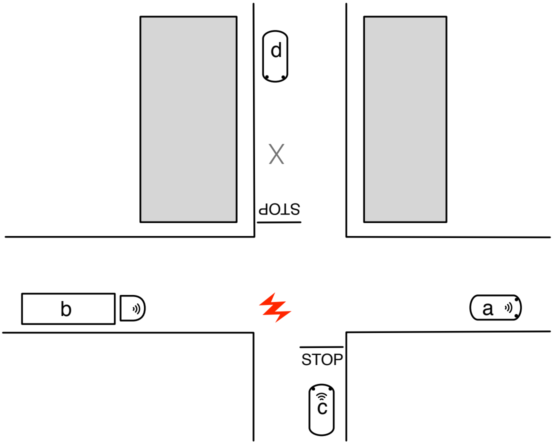
- Coalition $\{a, b\}$ has a *strategy* to prevent a collision.

Second-order know-how strategies



- Coalition $\{a, b\}$ has a *strategy* to prevent a collision.
- Coalition $\{a, b\}$ does not know such a strategy exists.
- Car c knows what is the strategy of coalition $\{a, b\}$ to avoid a collision.

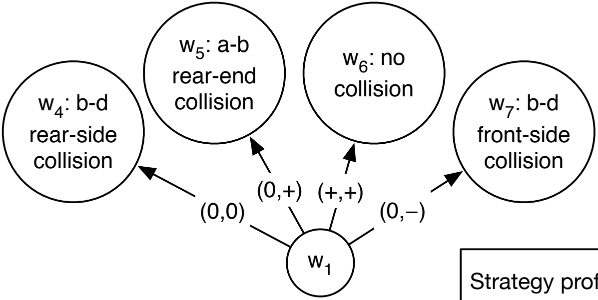
Second-order know-how strategies



$w_1 \Vdash \mathsf{H}^{\{a,b\}}_{\{c\}}$ ("avoid a collision")

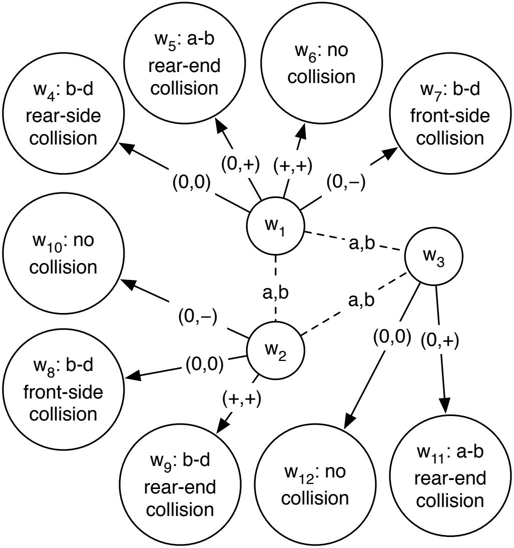
$w \Vdash \mathsf{H}^D_C\varphi$: if coalition C has distributed knowledge of how coalition D can achieve outcome φ from state w.

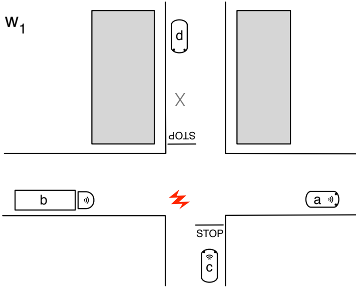
Epistemic Transition System



Strategy profile (s_a, s_b)
Domain of actions:
+: to accelerate
–: to slow down
0: to maintain speed

Epistemic Transition System





w_2: car d is at the spot X

w_3: car d is not present at the scene

Strategy profile (s_a, s_b)

Domain of actions:

+: to accelerate

–: to slow down

0: to maintain speed

Definition 1

A tuple $(W, \{\sim_a\}_{a\in\mathcal{A}}, \Delta, M, \pi)$ is an (epistemic) transition system, if

1. W is a set (of epistemic states),
2. $\sim_a$ is an indistinguishability equivalence relation on set W for each $a \in \mathcal{A}$,
3. Δ is a nonempty set, called the domain of actions,
4. $M \subseteq W \times \Delta^{\mathcal{A}} \times W$ is an aggregation mechanism,
5. π is a function from propositional variables to subsets of W.

Definition 1

A tuple $(W, \{\sim_a\}_{a\in\mathcal{A}}, \Delta, M, \pi)$ is an (epistemic) transition system, if

1. W is a set (of epistemic states),
2. $\sim_a$ is an indistinguishability equivalence relation on set W for each $a \in \mathcal{A}$,
3. Δ is a nonempty set, called the domain of actions,
4. $M \subseteq W \times \Delta^{\mathcal{A}} \times W$ is an aggregation mechanism,
5. π is a function from propositional variables to subsets of W.

For any states $w_1, w_2 \in W$ and any coalition C, let $w_1 \sim_C w_2$ if $w_1 \sim_a w_2$ for each agent $a \in C$.

Definition 1

A tuple $(W, \{\sim_a\}_{a\in\mathcal{A}}, \Delta, M, \pi)$ is an (epistemic) transition system, if

1. W is a set (of epistemic states),
2. $\sim_a$ is an indistinguishability equivalence relation on set W for each $a \in \mathcal{A}$,
3. Δ is a nonempty set, called the domain of actions,
4. $M \subseteq W \times \Delta^{\mathcal{A}} \times W$ is an aggregation mechanism,
5. π is a function from propositional variables to subsets of W.

For any states $w_1, w_2 \in W$ and any coalition C, let $w_1 \sim_C w_2$ if $w_1 \sim_a w_2$ for each agent $a \in C$.
A *strategy profile of a coalition* C is a tuple of values from Δ indexed by set C. A *complete* strategy profile is a strategy profile of the coalition $\mathcal{A}$.

Definition 2

Let Φ be the minimal set of formulae such that

1. $p \in \Phi$ for each propositional variable p,
2. $\neg\varphi, \varphi \to \psi \in \Phi$ for all formulae $\varphi, \psi \in \Phi$,
3. $\mathsf{K}_C\varphi, \mathsf{H}_C^D\varphi \in \Phi$ for each coalition C, each finite coalition D, and each formula $\varphi \in \Phi$.

Definition 3

For any epistemic state $w \in W$ of a transition system $(W, \{\sim_a\}_{a\in\mathcal{A}}, V, M, \pi)$ and any formula $\varphi \in \Phi$, let relation $w \Vdash \varphi$ be defined as follows:

1. $w \Vdash p$ if $w \in \pi(p)$, where p is a propositional variable,
2. $w \Vdash \neg\varphi$ if $w \nVdash \varphi$,
3. $w \Vdash \varphi \to \psi$ if $w \nVdash \varphi$ or $w \Vdash \psi$,
4. $w \Vdash \mathsf{K}_C\varphi$ if $w' \Vdash \varphi$ for each $w' \in W$ such that $w \sim_C w'$,
5. $w \Vdash \mathsf{H}_C^D\varphi$ if there is a strategy profile $\mathbf{s} \in V^D$ such that for any two states $w', u \in W$ and any complete strategy profile $\mathbf{s}'$, if $w \sim_C w'$, $\mathbf{s} =_D \mathbf{s}'$, and $(w', \mathbf{s}', u) \in M$, then $u \Vdash \varphi$.

Axioms

1. Truth: $\mathsf{K}_C\varphi \to \varphi$,
2. Negative Introspection: $\neg\mathsf{K}_C\varphi \to \mathsf{K}_C\neg\mathsf{K}_C\varphi$,
3. Distributivity: $\mathsf{K}_C(\varphi \to \psi) \to (\mathsf{K}_C\varphi \to \mathsf{K}_C\psi)$,
4. Monotonicity: $\mathsf{K}_C\varphi \to \mathsf{K}_D\varphi$, if $C \subseteq D$,

Axioms

1. Truth: $\mathsf{K}_C\varphi \to \varphi$,
2. Negative Introspection: $\neg\mathsf{K}_C\varphi \to \mathsf{K}_C\neg\mathsf{K}_C\varphi$,
3. Distributivity: $\mathsf{K}_C(\varphi \to \psi) \to (\mathsf{K}_C\varphi \to \mathsf{K}_C\psi)$,
4. Monotonicity: $\mathsf{K}_C\varphi \to \mathsf{K}_D\varphi$, if $C \subseteq D$,
5. Cooperation: $\mathsf{H}^{D_1}_{C_1}(\varphi \to \psi) \to (\mathsf{H}^{D_2}_{C_2}\varphi \to \mathsf{H}^{D_1\cup D_2}_{C_1\cup C_2}\psi)$, where $D_1 \cap D_2 = \varnothing$.

Axioms

1. Truth: $\mathsf{K}_C\varphi \to \varphi$,
2. Negative Introspection: $\neg\mathsf{K}_C\varphi \to \mathsf{K}_C\neg\mathsf{K}_C\varphi$,
3. Distributivity: $\mathsf{K}_C(\varphi \to \psi) \to (\mathsf{K}_C\varphi \to \mathsf{K}_C\psi)$,
4. Monotonicity: $\mathsf{K}_C\varphi \to \mathsf{K}_D\varphi$, if $C \subseteq D$,
5. Cooperation: $\mathsf{H}^{D_1}_{C_1}(\varphi \to \psi) \to (\mathsf{H}^{D_2}_{C_2}\varphi \to \mathsf{H}^{D_1\cup D_2}_{C_1\cup C_2}\psi)$, where $D_1 \cap D_2 = \varnothing$.
6. Strategic Introspection: $\mathsf{H}^D_C\varphi \to \mathsf{K}_C\mathsf{H}^D_C\varphi$,

Axioms

1. Truth: $\mathsf{K}_C\varphi \to \varphi$,
2. Negative Introspection: $\neg\mathsf{K}_C\varphi \to \mathsf{K}_C\neg\mathsf{K}_C\varphi$,
3. Distributivity: $\mathsf{K}_C(\varphi \to \psi) \to (\mathsf{K}_C\varphi \to \mathsf{K}_C\psi)$,
4. Monotonicity: $\mathsf{K}_C\varphi \to \mathsf{K}_D\varphi$, if $C \subseteq D$,
5. Cooperation: $\mathsf{H}^{D_1}_{C_1}(\varphi \to \psi) \to (\mathsf{H}^{D_2}_{C_2}\varphi \to \mathsf{H}^{D_1\cup D_2}_{C_1\cup C_2}\psi)$, where $D_1 \cap D_2 = \varnothing$.
6. Strategic Introspection: $\mathsf{H}^D_C\varphi \to \mathsf{K}_C\mathsf{H}^D_C\varphi$,
7. Empty Coalition: $\mathsf{K}_\varnothing\varphi \to \mathsf{H}^\varnothing_\varnothing\varphi$.

Axioms

1. Truth: $\mathsf{K}_C\varphi \to \varphi$,
2. Negative Introspection: $\neg\mathsf{K}_C\varphi \to \mathsf{K}_C\neg\mathsf{K}_C\varphi$,
3. Distributivity: $\mathsf{K}_C(\varphi \to \psi) \to (\mathsf{K}_C\varphi \to \mathsf{K}_C\psi)$,
4. Monotonicity: $\mathsf{K}_C\varphi \to \mathsf{K}_D\varphi$, if $C \subseteq D$,
5. Cooperation: $\mathsf{H}^{D_1}_{C_1}(\varphi \to \psi) \to (\mathsf{H}^{D_2}_{C_2}\varphi \to \mathsf{H}^{D_1\cup D_2}_{C_1\cup C_2}\psi)$, where $D_1 \cap D_2 = \varnothing$.
6. Strategic Introspection: $\mathsf{H}^D_C\varphi \to \mathsf{K}_C\mathsf{H}^D_C\varphi$,
7. Empty Coalition: $\mathsf{K}_\varnothing\varphi \to \mathsf{H}^\varnothing_\varnothing\varphi$.
8. Knowledge of Unavoidability: $\mathsf{K}_A\mathsf{H}^\varnothing_B\varphi \to \mathsf{H}^\varnothing_A\varphi$.

Inference Rules

1. Necessitation: $\dfrac{\varphi}{\mathsf{K}_C\varphi}$
2. Strategic Necessitation: $\dfrac{\varphi}{\mathsf{H}_C^D\varphi}$
3. Modus Ponens: $\dfrac{\varphi, \quad \varphi \to \psi}{\psi}$

Conclusion

AAMAS (International Conference on Autonomous Agents and Multiagent Systems) 2018

- A language for modeling an interplay between the distributed knowledge modality and the second-order coalition know-how modality.
- A sound and complete logical framework.