Topological Semantics for Common Inductive Knowledge

Siddharth Namachivayam
Department of Philosophy
Carnegie Mellon University*
Pittsburgh, Pennsylvania
snamachi@andrew.cmu.edu

Lewis' account of common knowledge in *Convention* describes the generation of higher-order expectations between agents as hinging upon agents' inductive standards and a shared witness. This paper attempts to draw from insights in learning theory to provide a formal account of common inductive knowledge and how it can be generated by a witness. Our language has a rather rich syntax in order to capture equally rich notions central to Lewis' account of common knowledge; for instance, we speak of an agent 'having some reason to believe' a proposition and one proposition 'indicating' to an agent that another proposition holds. A similar line of work was pursued by Cubitt & Sugden 2003; however, their account was left wanting for a corresponding semantics. Our syntax affords a novel topological semantics which, following Kelly 1996's approach in *The Logic of Reliable Inquiry*, takes as primitives agents' information bases. In particular, we endow each agent with a 'switching tolerance' meant to represent their personal inductive standards for learning. Curiously, when all agents are truly inductive learners (not choosing to believe only those propositions which are deductively verified), we show that the set of worlds where a proposition P is common inductive knowledge is invariant of agents' switching tolerances. Contrarily, the question of whether a specific witness W generates common inductive knowledge of P is sensitive to changing agents' switching tolerances. After establishing soundness of our proof system with respect to this semantics, we conclude by applying our logic to solve an 'inductive' variant of the coordinated attack problem.

1 Introduction

In *Convention*, David Lewis presents 3 conditions for a witness W to generate common knowledge of a proposition P.[1] These are:

1. Everyone[2] has reason to believe W.
2. W indicates to everyone that P.
3. W indicates to everyone that 1.

where W indicates to an agent i that P iff:

if i had reason to believe W, i would thereby have reason to believe P.

*The author would like to thank Adam Bjorndahl and Kevin Kelly for advising this project. Preliminary progress was presented in 2025 at the Indian Statistical Institute's conference on Recent Trends in Logic and Game Theory as well as the Progic Conference Series' 12th Workshop on Combining Logic and Probability. This work also received financial support in the summer of 2024 from Carnegie Mellon University's Institute for Complex Social Dynamics.

[1] Strictly speaking, Lewis defines what it means for W to be a 'basis' for common knowledge in P. We will instead call W a 'witness,' reserving the term 'basis' to be used in its topological sense.

[2] Throughout this paper, the set of agents is taken to be finite.

Submitted to:
arXiv

For Lewis, what W indicates to i (and whether W generates common knowledge of P) crucially depends on i's inductive standards. After making appropriate 'ancillary' assumptions about the nature of these inductive standards, Lewis uses his 3 conditions to derive that:

(1) Everyone has reason to believe P.

(2) Everyone has reason to believe (1).

(3) Everyone has reason to believe (2).

... ...

(n) Everyone has reason to believe ($n-1$).

However, Lewis is rather opaque about what his 'ancillary' assumptions exactly are. Understandably, subsequent formal treatments of common knowledge largely tended to overlook this critical inductive dimension in their accounts. Additionally, most formal treatments have also neglected the role of witnesses in generating common knowledge. For instance, at a pre-theoretic level, it is widespread in the game theory literature to speak of common knowledge/belief of P and mean no more than the obtaining of some infinite collection of propositions like:

(1') Everyone knows/believes P.

(2') Everyone knows/believes (1').

(3') Everyone knows/believes (2').

... ...

(n') Everyone knows/believes ($n-1$').

As a result, the original Lewisian picture of a witness generating common knowledge has largely been swept under the rug. Cubitt & Sugden 2003 are a notable exception; they attempt to painstakingly reconstruct Lewis' vision and provide an account of what his 'ancillary' assumptions about agents' inductive standards might be. While such an exercise is not without its merits, the primary purpose of this paper is not to argue about what Lewis did or did not think. Instead, we will aim to provide a logic for capturing how common inductive knowledge is generated by a witness that dispenses with the need for Lewis' 'ancillary' assumptions as much as possible while still maintaining the spirit of his original picture.

2 Pre-Theory

Central to our pre-theory will be the distinction between 'having *reason simpliciter* to believe' and 'having W *as reason* to believe.' Our primary thesis will be that i has W as reason to believe P iff:

- Some piece of evidence E gives i reason simpliciter to believe W.[3]
- If E is a piece of evidence which gives i reason simpliciter to believe W then $W \wedge E$ entails P.

In fact, we take this latter requirement to constitute the meaning of the proposition 'W indicates to i that P' since it can be glossed as saying 'if i had reason simpliciter to believe W, i would thereby have W as reason to believe P.' According to these definitions then, we have as primitive rule of inference:

$$\frac{}{(i \text{ has } A \text{ as reason to believe } B) \leftrightarrow [(i \text{ has reason simpliciter to believe } A) \wedge (A \text{ indicates to } i \text{ that } B)]}$$

[3]The crux of our semantics will be to properly cash out the phrase 'E is a piece of evidence which gives i reason simpliciter to believe W.' While the phrase will not be a part of our logical language, pre-theoretically, it will suffice to imagine that it is a proposition which either holds at no worlds or holds precisely at the worlds where E holds.

Next observe for any propositions W and E that $W \wedge E$ entails W. Hence, by the definition of indication, we can conclude W always indicates to i that W. Similarly if P is valid, since all propositions entail P, we can conclude any witness W indicates to i that P. Thus, we have as additional primitive rules:

$$\frac{}{A \text{ indicates to } i \text{ that } A}$$

$$\frac{B}{A \text{ indicates to } i \text{ that } B}$$

Note that by our first inference rule:

$$\frac{}{(i \text{ has } A \text{ as reason to believe } A) \leftrightarrow [(i \text{ has reason simpliciter to believe } A) \wedge (A \text{ indicates to } i \text{ that } A)]}$$

Therefore, by applying our second inference rule, we have the derived rule:

$$\frac{}{(i \text{ has } A \text{ as reason to believe } A) \leftrightarrow (i \text{ has reason simpliciter to believe } A)}$$

i.e., we can view the proposition 'i has reason simpliciter to believe W' as simply being shorthand for the proposition 'i has W as reason to believe W.' We insist the latter is pre-theoretically true of any valid proposition and so take as primitive rule of inference:

$$\frac{A}{i \text{ has reason simpliciter to believe } A}$$

We also take it that i has W as reason to believe W iff i has reason simpliciter to believe 'i has W as reason to believe W,' yielding the primitive rule:

$$\frac{}{\begin{array}{c} i \text{ has reason simpliciter to believe } A \\ \leftrightarrow \\ i \text{ has reason simpliciter to believe } (i \text{ has reason simpliciter to believe } A) \end{array}}$$

Now suppose E gives i reason simpliciter to believe W. Since $W \wedge E$ entails E, we have that $W \wedge E$ entails 'E gives i reason simpliciter to believe W', and so $W \wedge E$ entails 'i has reason simpliciter to believe W.' Hence, W always indicates to i that 'i has reason simpliciter to believe W,' giving us the primitive rule:

$$\frac{}{A \text{ indicates to } i \text{ that } (i \text{ has reason simpliciter to believe } A)}$$

Similarly, suppose W indicates to i that 'W indicates to i that P.' If E gives i reason simpliciter to believe W then $W \wedge E$ entails 'W indicates to i that P.' So assume $W \wedge E$ holds. Then W indicates to i that P. Further, since E gives i reason simpliciter to believe W and $W \wedge E$ holds, P must also hold. Hence, we can conclude $W \wedge E$ entails P, i.e. W indicates to i that P. Writing this as another primitive rule:

$$\frac{}{[A \text{ indicates to } i \text{ that } (A \text{ indicates to } i \text{ that } B)] \rightarrow (A \text{ indicates to } i \text{ that } B)}$$

Next, we pre-theoretically commit ourselves to the idea that if two propositions are logically equivalent in light of some evidence E then E either gives i reason simpliciter to believe both propositions or neither. To illustrate the principle, suppose W indicates to i that P and E is some piece of evidence which gives i reason simpliciter to believe W. Since $W \wedge E$ entails P we have that $W \wedge E$ is logically equivalent to $(W \wedge P) \wedge E$, i.e. W and $W \wedge P$ are logically equivalent in light of evidence E. Hence, E gives i reason simpliciter to believe $W \wedge P$.

This motivates us to assert that if W indicates to i that P and i has reason simpliciter to believe W then i has reason simpliciter to believe $W \wedge P$. In other words, we have the following primitive rule:

$$\frac{}{(i \text{ has } A \text{ as reason to believe } B) \rightarrow [i \text{ has reason simpliciter to believe } (A \wedge B)]}$$

Now suppose W indicates to i that P and $W \wedge P$ indicates to i that Q. If E is some piece of evidence which gives i reason simpliciter to believe W then, as we argued earlier, $W \wedge E$ is logically equivalent to $(W \wedge P) \wedge E$ and E gives i reason simpliciter to believe $W \wedge P$. Further, because $W \wedge P$ indicates to i that Q, we have $(W \wedge P) \wedge E$ entails Q and so $W \wedge E$ also entails Q. Thus, we conclude W indicates to i that Q and stipulate as yet another primitive rule of inference:

$$\frac{}{(A \text{ indicates to } i \text{ that } B) \wedge [(A \wedge B) \text{ indicates to } i \text{ that } C] \rightarrow (A \text{ indicates to } i \text{ that } C)}$$

One might have hoped that if W indicates to i that P and P indicates to i that Q then we could conclude W indicates to i that Q. However, if W indicates to i that P then a piece of evidence which gives i reason simpliciter to believe W will not necessarily give i reason simpliciter to believe P. Rather, a piece of evidence which gives i reason simpliciter to believe W will only give i reason simpliciter to believe $W \wedge P$. This is precisely why indication satisfies only the weak form of transitivity displayed in our last rule and not the stronger form of transitivity characteristic of the material conditional.

Another form of transitivity for indication arises from observing that the collection of propositions indicated to i by a fixed witness are closed under deduction. For suppose W indicates to i that P and W also indicates to i that $P \rightarrow Q$. If E is some piece of evidence which gives i reason simpliciter to believe W then $W \wedge E$ entails both P and $P \rightarrow Q$, i.e. $W \wedge E$ also entails Q. Hence, we have that W indicates to i that Q. This gives us the primitive rule:

$$\frac{}{[(A \text{ indicates to } i \text{ that } B) \wedge (A \text{ indicates to } i \text{ that } B \rightarrow C)] \rightarrow (A \text{ indicates to } i \text{ that } C)}$$

The above rule is quite powerful. For instance, along with just our first three primitive rules, it can be used to show all the following derived rules:

$$\frac{}{[(i \text{ has } A \text{ as reason to believe } B) \wedge (i \text{ has } A \text{ as reason to believe } B \rightarrow C)] \rightarrow (i \text{ has } A \text{ as reason to believe } C)}$$

$$\frac{}{[(A \text{ indicates to } i \text{ that } B) \wedge (A \text{ indicates to } i \text{ that } C)] \rightarrow (A \text{ indicates to } i \text{ that } B \wedge C)}$$

$$\frac{}{[(i \text{ has } A \text{ as reason to believe } B) \wedge (i \text{ has } A \text{ as reason to believe } C)] \rightarrow (i \text{ has } A \text{ as reason to believe } B \wedge C)}$$

$$\frac{A \rightarrow B}{A \text{ indicates to } i \text{ that } B}$$

$$\frac{B \rightarrow C}{(A \text{ indicates to } i \text{ that } B) \rightarrow (A \text{ indicates to } i \text{ that } C)}$$

$$\frac{B \rightarrow C}{(i \text{ has } A \text{ as reason to believe } B) \rightarrow (i \text{ has } A \text{ as reason to believe } C)}$$

Additionally, if we avail ourselves of the rule "A indicates to i that 'A indicates to i that B' implies A indicates to i that B," then we can establish a form of the *reflection principle*:

$$\frac{}{[i \text{ has } A \text{ as reason to believe } (i \text{ has } A \text{ as reason to believe } B)] \rightarrow (i \text{ has } A \text{ as reason to believe } B)}$$

We will also maintain throughout that our agents are *valid reasoners*, i.e. if W holds and i has W as reason to believe P then P holds. This is because whenever i has W as reason to believe P, there is some evidence E which gives i reason simpliciter to believe W. Furthermore, $W \wedge E$ entails P. As a result, if W holds then $W \wedge E$ holds and so P must also hold. Thus, we have as rule of inference:

$$\frac{}{[A \wedge (i \text{ has } A \text{ as reason to believe } B)] \to B}$$

However, we do not take the above rule to be primitive. For, next, we introduce into our syntax the phrase 'i has *some true reason* to believe P' so that if W holds and i has W as reason to believe P then i has some true reason to believe P. To ensure our agents are valid reasoners, we then maintain that P holds whenever i has some true reason to believe P. More compactly, this corresponds to the primitive rules:

$$\frac{}{[A \wedge (i \text{ has } A \text{ as reason to believe } B)] \to (i \text{ has some true reason to believe } B)}$$

$$\frac{}{(i \text{ has some true reason to believe } A) \to A}$$

Now, suppose W indicates to i that P and E gives i reason simpliciter to believe W. Then $W \wedge E$ entails P and so $W \wedge E$ indicates to i that P. Further, because W and $W \wedge E$ are logically equivalent in light of E, we know that E gives i reason simpliciter to believe $W \wedge E$. Hence i has $W \wedge E$ as reason to believe P. Thus, $W \wedge E$ entails 'i has some true reason to believe P' and so W indicates to i that 'i has some true reason to believe P,' giving us the primitive rule:

$$\frac{}{(W \text{ indicates to } i \text{ that } P) \to [W \text{ indicates to } i \text{ that } (i \text{ has some true reason to believe } P)]}$$

Similarly, suppose i has some true reason to believe P. This means for some 'witness' W that W holds and i has W as reason to believe P. In particular, because W indicates to i that P, the above rule implies W indicates to i that 'i has some true reason to believe P.' Since W holds and i has reason simpliciter to believe W, it follows that i has some true reason to believe 'i has some true reason to believe P.' As i is a valid reasoner, this allows us to write the biconditional primitive rule:

$$\frac{}{\begin{array}{c} i \text{ has some true reason to believe } A \\ \leftrightarrow \\ i \text{ has some true reason to believe } (i \text{ has some true reason to believe } A) \end{array}}$$

Next, imagine we have a (possibly infinite) collection of propositions $\mathscr{C}$ where for each proposition $P \in \mathscr{C}$ it holds that i has W as reason to believe P. Then it seems plausible to conclude that i has W as reason to believe every $P \in \mathscr{C}$ holds. On the other hand, if for each $P \in \mathscr{C}$ it can be said that i has some true reason to believe P then it seems *far less* plausible to conclude that i has some true reason to believe every $P \in \mathscr{C}$ holds. Couldn't i have a different reason for believing each $P \in \mathscr{C}$? Without a principled procedure for picking out and combining these reasons, to say that i has some true reason to believe every $P \in \mathscr{C}$ seems like a far stronger statement than saying for each $P \in \mathscr{C}$ that i has some true reason to believe P. Interestingly, in our semantics, it will turn out that if $\mathscr{C}$ is finite then such a principled procedure exists. Therefore, we take as primitive rule of inference:

$$\frac{}{\begin{array}{c} (i \text{ has some true reason to believe } A) \wedge (i \text{ has some true reason to believe } B) \\ \leftrightarrow \\ i \text{ has some true reason to believe } A \wedge B \end{array}}$$

Importantly though, if $\mathscr{C}$ is infinite and for each $P \in \mathscr{C}$ it can be said i has some true reason to believe P, we *do not* pre-theoretically commit ourselves to saying that i has some true reason to believe every $P \in \mathscr{C}$. This point will become relevant again when justifying our definition common inductive knowledge at the end of this section.

Our secondary thesis will be that the following infinite list of conditions **L1** constitute what it means for a witness W to generate common inductive knowledge of a proposition P:

1. Everyone has reason simpliciter to believe W.
2. W indicates to everyone that P.
3. W indicates to everyone that 1 and 2.

   ... ...

n. W indicates to everyone that 1 and $(n-1)$.

Using 1 and 2, we have as primitive rule:

$$\frac{}{(A \text{ generates common inductive knowledge of } B) \rightarrow (i \text{ has } A \text{ as reason to believe } B)}$$

Now, note 1 and 3 entail 'Everyone has W as reason to believe 1 and 2.' Since 1 and 2 entail 'Everyone has W as reason to believe P,' 1 and 3 also entail 'Everyone has W as reason to believe 'everyone has W reason to believe P.'' Similarly, starting with 1 and 4, one can recursively derive 'Everyone has W as reason to believe 'everyone has W as reason to believe 'everyone has W as reason to believe P.''' Hence, **L1** entails what we will call **L2**:

(1) Everyone has W as reason to believe P.

(2) Everyone has W as reason to believe (1).

(3) Everyone has W as reason to believe (2).

... ...

(n) Everyone has W as reason to believe $(n-1)$.

Next, note (2) entails 'W indicates to everyone that (1).' Since (1) entails 1 and 2, (2) also entails 'W indicates to everyone that 1 and 2,' i.e. (2) entails 1 and 3. Similarly, starting with (3), one can recursively derive 'W indicates to everyone that 1 and 3,' i.e. (3) entails 1 and 4. Thus, **L1** and **L2** are logically equivalent and so either can be taken to constitute what it means for W to generate common inductive knowledge of P. While **L1** evidently bears more structural similarity to Lewis' account of how common inductive knowledge is generated, we will work frequently with **L2** for purposes of clarity. For instance, **L2** makes it obvious that if W generates common inductive knowledge of P then for each $(n) \in \mathbf{L2}$ everyone has W as reason to believe (n), and so given our prior pre-theoretic commitments, everyone has W as reason to believe W generates common inductive knowledge of P. This motivates the primitive rule of inference:

$$\frac{}{\begin{array}{c} A \text{ generates common inductive knowledge of } B \\ \rightarrow \\ i \text{ has } A \text{ as reason to believe } (A \text{ generates common inductive knowledge of } B) \end{array}}$$

In fact, for any proposition Q, it can be seen that if Q entails 'everyone has W as reason to believe P' and Q entails 'everyone has W as reason to believe Q' then Q entails every proposition in **L2** holds.

To demonstrate, using the first entailment and the consequent of the second entailment, we have Q entails 'everyone has W as reason to believe 'everyone has W as reason to believe P.'' Similarly, using this new entailment with the consequent of the second entailment, one can recursively derive Q entails 'everyone has W as reason to believe 'everyone has W as reason to believe 'everyone has W as reason to believe P.''' Thus, we stipulate as yet another primitive rule:

$$\frac{C \to (\text{Everyone has } A \text{ as reason to believe } C) \quad C \to (\text{Everyone has } A \text{ as reason to believe } B)}{C \to (A \text{ generates common inductive knowledge of } B)}$$

How then ought we define common inductive knowledge? Lewis suggests we say P is common inductive knowledge if and only if :

some state of affairs holds such that it generates common inductive knowledge of P.[4]

While we agree with Lewis that this condition is sufficient for common inductive knowledge, we will insist that it is not necessary. To motivate why, suppose that W holds and generates common inductive knowledge of P as Lewis suggests. Then, using **L2**, we immediately have **L3**:

(1') Everyone has some true reason to believe P.

(2') Everyone has some true reason to believe (1').

(3') Everyone has some true reason to believe (2').

... ...

(n') Everyone has some true reason to believe $(n-1')$.

Recall Lewis seeks to show his 3 conditions for a witness to generate common inductive knowledge of P entail exactly this sort of infinite list. But **L3** is not equivalent to Lewis' definition of common inductive knowledge for each agent i could have a different true reason for believing each proposition in **L3**. On our view, Lewis' definition is unnecessarily stringent by his own lights and **L3** is closer to a desirable definition of common inductive knowledge. However, **L3** is not without its own shortcomings. For if every proposition in **L3** holds, although we can conclude for each $(n') \in \mathbf{L3}$ everyone has some true reason to believe (n'), we cannot further conclude everyone has some true reason to believe every $(n') \in \mathbf{L3}$. After all, since **L3** is infinite, agents may not have a principled procedure for creating a new 'combined' true reason to believe every $(n') \in \mathbf{L3}$. Thus, to guarantee our notion of common inductive knowledge is a 'fixed point,' our tertiary thesis will be that what it means for P to be common inductive knowledge is constituted by satisfying the primitive rules of inference:

$$\frac{}{(A \text{ is common inductive knowledge}) \to A}$$

$$\frac{}{\begin{gathered} A \text{ is common inductive knowledge} \\ \to \\ i \text{ has some true reason to believe } A \text{ is common inductive knowledge} \end{gathered}}$$

$$\frac{A \to (\text{Everyone has some true reason to believe } A) \quad A \to B}{A \to (B \text{ is common inductive knowledge})}$$

, i.e. 'P is common inductive knowledge' is the weakest proposition entailing P such that if it is true then everyone has some true reason to believe it. As a check, one can easily verify we have the derived rule:

$$\frac{}{[A \wedge (A \text{ generates common inductive knowledge of } B)] \to (B \text{ is common inductive knowledge})}$$

, delivering on our promise that Lewis' definition would entail ours.

[4]Again, strictly speaking, Lewis actually says that P is common knowledge iff some state of affairs holds such that it is a basis for common knowledge in P.

With pre-theoretic commitments out of the way, we will now present the syntax for our logic.

3 Syntax

Let PROP be a countable set of primitive propositions and N be a finite set of agents.

Our logical language $\mathscr{L}$ will be defined by the Backus-Naur form:

$$\varphi ::= p|\top|\bot|\neg\varphi|\varphi_1 \wedge \varphi_2|\varphi_1 \vee \varphi_2|\varphi_1 \rightarrow \varphi_2|\varphi_1 \leftrightarrow \varphi_2|\mathbb{R}_i\varphi|\mathbb{I}_{i@\varphi_1}\varphi_2|\mathbb{B}_{i@\varphi_1}\varphi_2|\mathbb{S}_i\varphi|\mathbb{G}_{\varphi_1}\varphi_2|\mathbb{C}\varphi$$

where $p \in$ PROP and $i \in N$.

The below table summarizes how the epistemic modalities above correspond to the pre-theoretic notions discussed in the prior section:

Notation	**Gloss**
$\mathbb{R}_i\varphi$	i has reason simpliciter to believe φ
$\mathbb{I}_{i@\varphi_1}\varphi_2$	φ_1 indicates to i that φ_2
$\mathbb{B}_{i@\varphi_1}\varphi_2$	i has φ_1 as reason to believe φ_2
$\mathbb{S}_i\varphi$	i has some true reason to believe φ
$\mathbb{G}_{\varphi_1}\varphi_2$	φ_1 generates common inductive knowledge of φ_2
$\mathbb{C}\varphi$	φ is common inductive knowledge

Table 1: Gloss of epistemic modalities

Accordingly, our proof system takes not only all propositional tautologies to be axioms, but also all instances of the following primitive schemes:

$\mathbf{Ax}_{\mathsf{R}}: \ \mathbb{R}_i\varphi \leftrightarrow \mathbb{R}_i\mathbb{R}_i\varphi$

$\mathbf{Ax}_{\mathsf{I1}}: \ \mathbb{I}_{i@\varphi}\varphi$

$\mathbf{Ax}_{\mathsf{I2}}: \ \mathbb{I}_{i@\varphi}\mathbb{R}_i\varphi$

$\mathbf{Ax}_{\mathsf{I3}}: \ \mathbb{I}_{i@\varphi_1}\mathbb{I}_{i@\varphi_1}\varphi_2 \rightarrow \mathbb{I}_{i@\varphi_1}\varphi_2$

$\mathbf{Ax}_{\mathsf{I4}}: \ (\mathbb{I}_{i@\varphi_1}\varphi_2 \wedge \mathbb{I}_{i@(\varphi_1\wedge\varphi_2)}\varphi_3) \rightarrow \mathbb{I}_{i@\varphi_1}\varphi_3$

$\mathbf{Ax}_{\mathsf{I5}}: \ (\mathbb{I}_{i@\varphi_1}\varphi_2 \wedge \mathbb{I}_{i@\varphi_1}(\varphi_2 \rightarrow \varphi_3)) \rightarrow \mathbb{I}_{i@\varphi_1}\varphi_3$

$\mathbf{Ax}_{\mathsf{B1}}: \ \mathbb{B}_{i@\varphi_1}\varphi_2 \leftrightarrow (\mathbb{R}_i\varphi_1 \wedge \mathbb{I}_{i@\varphi_1}\varphi_2)$

$\mathbf{Ax}_{\mathsf{B2}}: \ \mathbb{B}_{i@\varphi_1}\varphi_2 \rightarrow \mathbb{R}_i(\varphi_1 \wedge \varphi_2)$

$\mathbf{Ax}_{\mathsf{S1}}: \ (\varphi_1 \wedge \mathbb{B}_{i@\varphi_1}\varphi_2) \rightarrow \mathbb{S}_i\varphi_2$

$\mathbf{Ax}_{\mathsf{S2}}: \ \mathbb{S}_i\varphi \rightarrow \varphi$

$\mathbf{Ax}_{\mathsf{S3}}: \ \mathbb{I}_{i@\varphi_1}\varphi_2 \rightarrow \mathbb{I}_{i@\varphi_1}\mathbb{S}_i\varphi_2$

$\mathbf{Ax}_{\mathsf{S4}}: \ \mathbb{S}_i\varphi \leftrightarrow \mathbb{S}_i\mathbb{S}_i\varphi$

$\mathbf{Ax}_{\mathsf{S5}}: \ (\mathbb{S}_i\varphi_1 \wedge \mathbb{S}_i\varphi_2) \leftrightarrow \mathbb{S}_i(\varphi_1 \wedge \varphi_2)$

$\mathbf{Ax}_{\mathsf{G1}}: \ \mathbb{G}_{\varphi_1}\varphi_2 \rightarrow \mathbb{B}_{i@\varphi_1}\varphi_2$

$\mathbf{Ax}_{\mathsf{G2}}: \ \mathbb{G}_{\varphi_1}\varphi_2 \rightarrow \mathbb{B}_{i@\varphi_1}\mathbb{G}_{\varphi_1}\varphi_2$

$\mathbf{Ax}_{\mathsf{C1}}: \ \mathbb{C}\varphi \rightarrow \varphi$

$\mathbf{Ax}_{\mathsf{C2}}: \ \mathbb{C}\varphi \rightarrow \mathbb{S}_i\mathbb{C}\varphi$

We also take as primitive rules of inference:

$$\mathbf{R}_{\mathsf{R}}: \ \frac{\varphi}{\mathbb{R}_i\varphi} \quad \mathbf{R}_{\mathsf{I}}: \ \frac{\varphi_2}{\mathbb{I}_{i@\varphi_1}\varphi_2}$$

$$\mathbf{R}_{\mathsf{G}}: \ \frac{\varphi_3 \rightarrow \bigwedge_{i\in N}\mathbb{B}_{i@\varphi_1}\varphi_3 \quad \varphi_3 \rightarrow \bigwedge_{i\in N}\mathbb{B}_{i@\varphi_1}\varphi_2}{\varphi_3 \rightarrow \mathbb{G}_{\varphi_1}\varphi_2} \quad \mathbf{R}_{\mathsf{C}}: \ \frac{\varphi_1 \rightarrow \bigwedge_{i\in N}\mathbb{S}_i\varphi_1 \quad \varphi_1 \rightarrow \varphi_2}{\varphi_1 \rightarrow \mathbb{C}\varphi_2}$$

In the coming section, we provide the reader with an overview of concepts from topological learning theory which will feature prominently in our semantics. We begin by reviewing standard results from the field. This review is intended to fix notation and recall basic facts. Readers who are already familiar with the topological difference hierarchy and its epistemic interpretation may wish to skip ahead. Such readers can turn directly to page 12 where we begin introducing several new pieces of terminology. These new terms and the theorems we prove about them will be essential for understanding the rest of the paper.

4 Topology & Learning

Let $\Omega \neq \varnothing$ be the set of all possible worlds. We say that a *piece of evidence E* for agent i is a non-empty subset of possible worlds in Ω which eliminates all worlds outside E as possibilities. Let $\mathscr{E}_i$ consist of all pieces of evidence agent i might ever learn. Further, given a world $w \in \Omega$, let $\mathscr{E}_i(w)$ denote those elements of $\mathscr{E}_i$ which contain w. Intuitively, $\mathscr{E}_i(w)$ represents exactly those pieces of evidence which agent i eventually learns in world w. We will assume that $\mathscr{E}_i$ forms an *information basis* for i, that is:

1. $\forall w \in \Omega, \mathscr{E}_i(w) \neq \varnothing$.
2. $\forall w \in \Omega$ and $\forall E_1, E_2 \in \mathscr{E}_i(w)$ there exists $E_3 \in \mathscr{E}_i(w)$ such that $E_3 \subseteq E_1 \cap E_2$.

The first requirement above simply says that agent i learns some piece of evidence in every world. In our view, this is a very weak assumption. For even if agent i never rules out any possible world in Ω, then Ω itself must be a piece of evidence for i. The second requirement is more substantive and corresponds to the assumption that evidence is cumulative. Namely, this requirement says that if agent i learns both E_1 and E_2 as evidence then they will also learn some piece of evidence E_3 at least as strong as $E_1 \cap E_2$. We now note that our definition of an *information* basis is identical to that of a *topological* basis. Hence, the collection $\mathscr{T}_i$ generated by taking arbitrary unions of pieces of evidence in $\mathscr{E}_i$ forms a topology over Ω.

In particular, it turns out the limit decidability of a set for agent i will be closely related to whether the set can be expressed as the nested difference of a certain number of descending open sets in the topology $\mathscr{T}_i$.[5] However, to illustrate why, we will need to introduce quite a bit of terminology. For starters, say that:

- A *decision method* for i is a map $m_i : \mathscr{E}_i \to \{\text{Yes}, \text{No}\}$.
- A *t-switching sequence starting with 'Yes'* for m_i is a finite downward sequence $E_0 \supseteq ... \supseteq E_t$ of pieces of evidence in $\mathscr{E}_i$ such that $m_i(E_{2k}) = \text{Yes}$ and $m_i(E_{2k+1}) = \text{No}$.
- A *t-switching sequence starting with 'No'* for m_i is a finite downward sequence $E_0 \supseteq ... \supseteq E_t$ of pieces of evidence in $\mathscr{E}_i$ such that $m_i(E_{2k}) = \text{No}$ and $m_i(E_{2k+1}) = \text{Yes}$.
- A decision method for *i has at most n switches after saying 'Yes'* if it is a decision method which contains no t-switching sequences starting with 'Yes' for any $t > n$.
- A decision method for *i has at most n switches after saying 'No'* if it is a decision method which contains no t-switching sequences starting with 'No' for any $t > n$.
- A decision method for *i has a bounded number of switches* if it is a decision method which either has at most n switches after saying 'Yes' or has at most n switches after saying 'No' for some n.

[5] We define the nested difference of a finite sequence of sets $S_0,...,S_n$ to be given by $S_0 \setminus (S_1 \setminus (...S_n)...)$.

Lemma 4.0: Given a decision method m_i for agent i which has a bounded number of switches, there exists a function $\sigma_{m_i} : \Omega \to \{\text{Yes}, \text{No}\}$ which $\forall w \in \Omega$ sets:

$$\sigma_{m_i}(w) = \begin{cases} \text{Yes} & \exists E \in \mathscr{E}_i(w) \forall E' \in \mathscr{E}_i(w),\ E' \subseteq E \to m_i(E') = \text{Yes} \\ \text{No} & \exists E \in \mathscr{E}_i(w) \forall E' \in \mathscr{E}_i(w),\ E' \subseteq E \to m_i(E') = \text{No} \end{cases}$$

Proof

First we show σ_{m_i} is a well-defined partial function. So suppose towards a contradiction it were not. Then for some $w^* \in \Omega$ both piecewise conditions above are satisfied. Then there $\exists E_1, E_2 \in \mathscr{E}_i(w^*)$ such that $\forall E' \in \mathscr{E}_i(w^*),\ E' \subseteq E_1 \to m_i(E') = \text{Yes}$ and $E' \subseteq E_2 \to m_i(E') = \text{No}$. Further, since $\mathscr{E}_i$ is an information basis, $\exists E_3 \in \mathscr{E}_i(w^*)$ such that $E_3 \subseteq E_1 \cap E_2$. But this implies that both $m_i(E_3) = \text{Yes}$ and $m_i(E_3) = \text{No}$, giving us our desired contradiction.

Now we show σ_{m_i} is total. So suppose towards a contradiction it were not. Then for some $w^* \in \Omega$ neither piecewise condition above is satisfied. Further, since $\mathscr{E}_i$ is an information basis, we know $\exists E_0 \in \mathscr{E}_i(w^*)$. WLOG suppose $m_i(E_0) = \text{Yes}$. Because the first piecewise is not satisfied, $\exists E_1 \in \mathscr{E}_i(w^*)$ so that $E_1 \subseteq E_0$ and $m_i(E_1) = \text{No}$. However, because the second piecewise condition is not satisfied, $\exists E_2 \in \mathscr{E}_i(w^*)$ so that $E_2 \subseteq E_1$ and $m_i(E_2) = \text{Yes}$. In this way we can create an infinite downward sequence $E_0 \supseteq E_1 \supseteq E_2 \ldots$ of pieces of evidence in $\mathscr{E}_i$ such that $m_i(E_{2k}) = \text{Yes}$ and $m_i(E_{2k+1}) = \text{No}$. This contradicts the fact that m_i is a decision method which has a bounded number of switches, as desired. $\square$

Conceptually speaking, given a decision method m_i for agent i which has a bounded number of switches, $\sigma_{m_i}(w)$ returns the output that m_i converges to in world w. Hence, say that:

- Agent *i can limit decide W in n-switches after saying 'Yes'* just in case there is a decision method m_i for i such that m_i has at most n switches after saying 'Yes' and $\sigma_{m_i}^{-1}(\text{Yes}) = W$.
- Agent *i can limit decide W in n-switches after saying 'No'* just in case there is a decision method m_i for i such that m_i has at most n switches after saying 'No' and $\sigma_{m_i}^{-1}(\text{Yes}) = W$.

Theorem 4.0: Agent i can limit decide W in n switches after saying 'Yes' if and only if W can be expressed as the nested difference of $n+1$ descending open sets in $\mathscr{T}_i$.

Proof

First we prove the 'only if' direction. So assume agent i can limit decide W in n switches after saying 'Yes.' Then there is a decision method m_i for i such that m_i has at most n switches after saying 'Yes' and $\sigma_{m_i}^{-1}(\text{Yes}) = W$. Let $I_0 = \{E \in \mathscr{E}_i : m_i(E) = \text{Yes}\}$. Next, recursively define I_k by setting:

$$I_k = \begin{cases} \{E' \in \mathscr{E}_i : \exists E \in I_{k-1}, E' \subseteq E \text{ and } m_i(E') = \text{No}\} & k \text{ is odd} \\ \{E' \in \mathscr{E}_i : \exists E \in I_{k-1}, E' \subseteq E \text{ and } m_i(E') = \text{Yes}\} & k \text{ is even} \end{cases}$$

Note that $I_{n+1} = \varnothing$ since if $\exists E_{n+1} \in I_{n+1}$ then we could find a $n+1$ switching sequence starting with 'Yes' for m_i of the form $E_0 \supseteq \ldots \supseteq E_{n+1}$ where each $E_k \in I_k$. Now let $O_k = \bigcup_{E \in I_K} E$. Each O_k is open as it is a union of basis elements. In particular, $\{O_k\}_{k \in \mathbb{N}}$ forms a descending sequence of open sets with $O_{n+1} = \varnothing$ and so it suffices show $O_0 \backslash (O_1 \backslash (\ldots O_n) \ldots) = W$.

Suppose $w \in O_0 \backslash (O_1 \backslash (\ldots O_n) \ldots)$. Let k^* be the largest k for which $w \in O_k$. Since O_{k^*} is not empty, we know $k^* < n+1$. Additionally, since $w \in O_0 \backslash (O_1 \backslash (\ldots O_n) \ldots)$, we know k^* is even. Hence, $\exists E^* \in I_{k^*}$ such that $w \in E^*$ and $m_i(E^*) = \text{Yes}$. Assume for sake of contradiction that $\exists E' \in \mathscr{E}_i(w)$ such that $E' \subseteq E^*$ and $m_i(E') = \text{No}$. Then $E' \in I_{k^*+1}$ and so $w \in E' \subseteq O_{k^*+1}$, contradicting maximality of k^*. Thus, E^* is an element of $\mathscr{E}_i(w)$ such that $\forall E' \in \mathscr{E}_i(w)$ we have $E' \subseteq E^* \to m_i(E') = \text{Yes}$, i.e. $w \in \sigma_{m_i}^{-1}(\text{Yes}) = W$.

Now suppose $w \in W = \sigma_{m_i}^{-1}(\text{Yes})$. Then $\exists E_0 \in \mathscr{E}_i(w)$ such that $m_i(E_0) = \text{Yes}$. In particular, $E_0 \in I_0$ and so $w \in E_0 \subseteq O_0$. Let k^* be the largest k for which $w \in O_k$. Since O_{k^*} is not empty we know $k^* < n+1$. Suppose towards a contradiction that k^* is odd. Then, $\exists E^* \in I_{k^*}$ such that $w \in E^*$ and $m_i(E^*) = \text{No}$. Further E^* is an element of $\mathscr{E}_i(w)$ such that $\forall E' \in \mathscr{E}_i(w)$, if $E' \subseteq E^*$ then $m_i(E') = \text{No}$ (otherwise we would have $E' \in I_{k^*+1}$ and $w \in E' \subseteq O_{k^*+1}$, contradicting maximality of k^*). But this means $w \in \sigma_{m_i}^{-1}(\text{No})$. Contradiction. Thus k^* is even and so we can conclude $w \in O_0 \backslash (O_1 \backslash (...O_n)...)$.

Next we prove the 'if' direction. So suppose W can be expressed as the nested difference of $n+1$ descending open sets $O_0 \supseteq ... \supseteq O_n$ in $\mathscr{T}_i$. Given $E \in \mathscr{E}_i$, define k_E to be the largest k for which $E \subseteq O_k$ (if there is no such k set $k_E = -1$). Let m_i be the decision method for i defined by $\forall E \in \mathscr{E}_i$ setting:

$$m_i(E) = \begin{cases} \text{No} & k_E \text{ is odd} \\ \text{Yes} & k_E \text{ is even} \end{cases}$$

Assume for sake of contradiction there is a $n+1$ switching sequence starting with 'Yes' for m_i of the form $E_0 \supseteq ... \supseteq E_{n+1}$. Since k_{E_0} is at least 0 and $k_{E_0} < k_{E_1} < ... < k_{E_{n+1}}$, we know $k_{E_{n+1}} \geq n+1$. However, we also know $k_{E_{n+1}} \leq n$ by definition. Contradiction. Hence, m_i has at most n switches after saying 'Yes' and so it suffices to show $O_0 \backslash (O_1 \backslash (...O_n)...) = \sigma_{m_i}^{-1}(\text{Yes})$.
Suppose $w \in O_0 \backslash (O_1 \backslash (...O_n)...)$. Let k^* be the largest k for which $w \in O_k$. Then $\exists E^* \in \mathscr{E}_i(w)$ such that $E^* \subseteq O_{k^*}$. Further, E^* is an element $\mathscr{E}_i(w)$ such that $\forall E' \in \mathscr{E}_i(w)$, if $E' \subseteq E^*$ then $k_{E'} = k^*$ (otherwise we would have $k_{E'} > k^*$ and $w \in E' \subseteq O_{k_{E'}}$, contradicting maximality of k^*). Since $w \in O_0 \backslash (O_1 \backslash (...O_n)...)$, we know k^* is even. Thus $\forall E' \in \mathscr{E}_i(w)$, if $E' \subseteq E^*$ then $m_i(E') = \text{Yes}$, i.e. $w \in \sigma_{m_i}^{-1}(\text{Yes})$.
Now suppose $w \in \sigma_{m_i}^{-1}(\text{Yes})$. Then $\exists E_0 \in \mathscr{E}_i(w)$ such that $m_i(E_0) = \text{Yes}$. In particular, $w \in E_0 \subseteq O_0$. Let k^* be the largest k for which $w \in O_k$. Suppose for contradiction that k^* is odd. We know $\exists E^* \in \mathscr{E}_i(w)$ such that $E^* \subseteq O_{k^*}$. Further, E^* is an element $\mathscr{E}_i(w)$ such that $\forall E' \in \mathscr{E}_i(w)$, if $E' \subseteq E^*$ then $k_{E'} = k^*$ (otherwise we would have $k_{E'} > k^*$ and $w \in E' \subseteq O_{k_{E'}}$, contradicting maximality of k^*). But since k^* is odd this means $\forall E' \in \mathscr{E}_i(w)$, if $E' \subseteq E^*$ then $m_i(E') = \text{No}$, i.e. $w \in \sigma_{m_i}^{-1}(\text{No})$. Contradiction. Thus k^* is even and so we can conclude $w \in O_0 \backslash (O_1 \backslash (...O_n)...)$. □

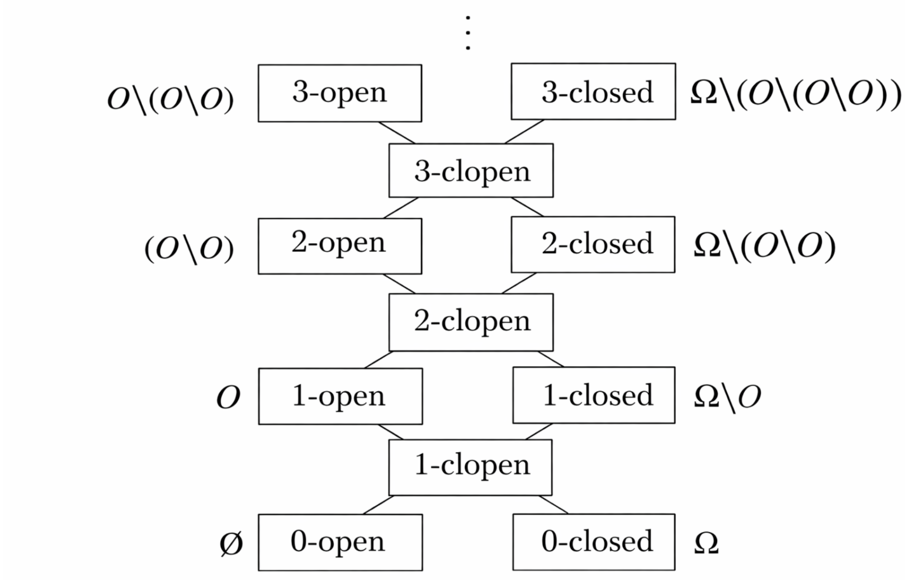
Lemma 4.1: Agent i can limit decide W in n switches after saying 'No' if and only if agent i can limit decide $\Omega \backslash W$ in n switches after saying 'Yes.'
<u>Proof</u>
We only show one direction as the other direction follows from a parallel argument. So assume agent i can limit decide W in n switches after saying 'No.' Then there is a decision method m_i for i such that m_i has at most n switches after saying 'No' and $\sigma_{m_i}^{-1}(\text{Yes}) = W$. Consider the decision method m_i' obtained by $\forall E \in \mathscr{E}_i$ setting:

$$m_i'(E) = \begin{cases} \text{Yes} & m_i(E) = \text{No} \\ \text{No} & m_i(E) = \text{Yes} \end{cases}$$

Note that m_i' has at most n switches after saying 'Yes.' Further $\sigma_{m_i'}^{-1}(\text{Yes}) = \sigma_{m_i}^{-1}(\text{No}) = \Omega \backslash W$. Thus agent i can limit decide $\Omega \backslash W$ in n switches after saying 'Yes.' □

Corollary 4.0: Agent i can limit decide W in n switches after saying 'No' if and only if W can be expressed as the complement of a nested difference of $n+1$ descending open sets in $\mathscr{T}_i$.
<u>Proof</u>
Immediate by the Theorem 4.0 and Lemma 4.1. □

Going forward, we will say a set is *n-open* in $\mathscr{T}_i$ iff it is the nested difference of n descending open sets in $\mathscr{T}_i$. Additionally, say a set is *n-closed* in $\mathscr{T}_i$ iff it is the complement of a set which is n-open in $\mathscr{T}_i$. Finally, say a set is *n-clopen* in $\mathscr{T}_i$ iff it is both n-open and n-closed in $\mathscr{T}_i$. These topological complexity classes are then ordered by inclusion according to the below *topological difference hierarchy*:[6]



We have already seen an epistemic interpretation for when $n > 0$ of what it means to be n-open, n-closed, and n-clopen in $\mathscr{T}_i$. But what does it mean for a set to be 0-open or 0-closed in $\mathscr{T}_i$? One interpretation is as follows: say an information basis $\mathscr{E}_i$ for *i has a starting point* if $\Omega \in \mathscr{E}_i$, i.e. agent i receives Ω as a piece of evidence before learning anything stronger. If $\mathscr{E}_i$ has a starting point, a decision method m_i for i must make a choice of whether $m_i(\Omega) = \text{No}$ or $m_i(\Omega) = \text{Yes}$. Intuitively then, a set W is 0-open in $\mathscr{T}_i$ if and only if agent i can limit decide W in 0 switches *starting from* 'No,' i.e. $W = \varnothing$. Similarly, a set W is 0-closed in $\mathscr{T}_i$ if and only if agent i can limit decide W in 0 switches *starting from* 'Yes,' i.e. $W = \Omega$.

To cash out this intuition, temporarily assume agent i's information basis $\mathscr{E}_i$ has a starting point and extend our prior terminology by saying:

- A decision method m_i for *i starts from 'Yes' ('No')* if $m_i(\Omega) = \text{Yes}$ (No).
- A decision method for *i has at most n switches starting from 'Yes' ('No')* if it is a decision method which starts from 'Yes' ('No') and has at most n switches after saying 'Yes' ('No').
- A decision method for *i has at most n switches* if it is a decision method which has at most n switches starting from 'Yes' or at most n switches starting from 'No.'
- Agent *i can limit decide W in n switches starting from 'Yes' ('No')* if there is a decision method m_i for i such that m_i has at most n switches starting from 'Yes' ('No') and $\sigma_{m_i}^{-1}(\text{Yes}) = W$.
- Agent *i can limit decide W in n switches* if agent i can limit decide W in n switches starting from 'Yes' or in n switches starting from 'No.'

[6] For historical context, Putnam's 1965 paper was the first to try and characterize the complexity of what he called 'trial and error predicates,' albeit in a computability theoretic context. Kevin Kelly later recast the problem in general topological terms...

The above notions are also closely related to the topological difference hierarchy. To see precisely how, given a set W, let $n_{i,W}$ be the smallest n for which agent i can limit decide W in n switches (if there is no such n, let $n_{i,W} = \infty$). Additionally, assume agent i has finite *switching tolerance* $n_i \geq 0$ so that i only adopts those decision methods which have at most n_i switches. Because i can 'feasibly' limit decide W iff $n_{i,W} \leq n_i$, we say that:

- Agent i *has reason to limit decide* W *starting from 'Yes' ('No')* if $n_{i,W} \leq n_i$ and agent i can limit decide W in $n_{i,W}$ switches starting from 'Yes' ('No').

Theorem 4.1: Agent i has reason to limit decide W starting from 'Yes' if and only if $\exists k \leq n_i$ such that W is k-closed in $\mathscr{T}_i$ but not $(k-1)$-open.[7]

Proof

First we show the 'if' direction. Suppose $\exists k \leq n_i$ so that W is k-closed in $\mathscr{T}_i$ but not $(k-1)$-open.
If $k = 0$ we know W is 0-closed, and hence, i can limit decide W in 0 switches starting from 'Yes,' i.e. $n_{i,W} = 0$. Since $0 = n_{i,W} \leq n_i$, we can conclude agent i has reason to limit decide W starting from 'Yes.'
If $k = 1$ we know W is not 0-open, and hence, i cannot limit decide W in 0 switches starting from 'No.' However, because W is 1-closed, i can limit decide W in 0 switches after saying 'No.' Thus, there is a method m_i for i such that m_i has at most 0 switches after saying 'No' and $\sigma_{m_i}^{-1}(\text{Yes}) = W$. Further, we know $m_i(\Omega) \neq$ 'No' as otherwise i could limit decide W in 0 switches starting from 'No.' Hence $m_i(\Omega) =$ 'Yes' and so i can limit decide W in 1 switch starting from 'Yes.' Therefore, either $n_{i,W} = 0$ or $n_{i,W} = 1$. Further $1 = k \leq n_i$ and so $n_{i,W} \leq n_i$. If $n_{i,W} = 1$ then we can conclude agent i has reason to limit decide W starting from 'Yes.' Similarly, if $n_{i,W} = 0$ then i can limit decide W in 0 switches starting from 'Yes' (as i can't limit decide W in 0 switches starting from 'No'). So, in this case as well, we can conclude agent i has reason to limit decide W starting from 'Yes.'
If $k \geq 2$ we know W is not $(k-1)$-open, and hence, i cannot limit decide W in $k-2$ switches after saying 'Yes.' However because W is k-closed, i can limit decide W in $k-1$ switches after saying 'No.' Thus, there is a method m_i for i such that m_i has at most $k-1$ switches after saying 'No' and $\sigma_{m_i}^{-1}(\text{Yes}) = W$. Further, for any such m_i, we know $m_i(\Omega) \neq$ 'No' as otherwise i could limit decide W is $k-2$ switches after saying 'Yes.' Hence $m_i(\Omega) = \text{Yes}$ and so i can limit decide W in k switches starting from 'Yes' but cannot in $(k-1)$-switches starting from 'No.' Therefore, $n_{i,W} \leq k \leq n_i$. If $n_{i,W} = k$ then we can conclude agent i has reason to limit decide W starting from 'Yes.' Similarly, if $n_{i,W} < k$ then i can limit decide W in $n_{i,W}$ switches starting from 'Yes' (as i can't limit decide W in $n_{i,W}$ switches starting from 'No'). So in these cases as well, we can conclude i has reason to limit decide W starting from 'Yes.'

Now we show the 'only if' direction. Suppose agent i has reason to limit decide W starting from 'Yes.' Then $n_{i,W} \leq n_i$. Further, i can limit decide W in $n_{i,W}$ switches starting from 'Yes' and, hence, after saying 'Yes'. Thus, W is $n_{i,W}$-closed in $\mathscr{T}_i$ and so it suffices to show W is not $(n_{i,W} - 1)$-open in $\mathscr{T}_i$.
If $n_{i,W} = 0$ then, since W is never (-1)-open, we are done.
If $n_{i,W} = 1$ then, by definition, i cannot limit decide W in 0 switches. In particular, i cannot limit decide W in 0 switches starting from 'No.' Hence $W \neq \varnothing$ and so W is not 0-open as desired.
If $n_{i,W} \geq 2$ then, by definition, i cannot limit decide W in $(n_{i,W} - 1)$-switches. Suppose towards a contradiction W is $(n_{i,W} - 1)$-open. Then i can limit decide W in $n_{i,W} - 2$ switches after saying 'Yes.' Thus, there is a method m_i for i such that m_i has at most $n_{i,W} - 2$ switches after saying 'Yes' and $\sigma_{m_i}^{-1}(\text{Yes}) = W$. Further, we know $m_i(\Omega) \neq \text{Yes}$ as otherwise i could limit decide W in $(n_{i,W} - 2)$-switches starting from 'Yes.' Hence $m_i(\Omega) = \text{No}$. But this means agent i can limit decide W in $(n_{i,W} - 1)$-switches starting from 'No.' Contradiction. Hence W is not $(n_{i,W} - 1)$-open, as desired. □

[7] No set is (-1)-open.

Lemma 4.2: Agent i has reason to limit decide W starting from 'No' if and only if agent i has reason to limit decide $\Omega\backslash W$ starting from 'Yes.'
Proof
We only show one direction as the other direction follows from a parallel argument. So assume agent i has reason to limit decide W starting from 'No.' Then $n_{i,W} \leq n_i$ and there is a decision method m_i for i such that m_i has at most $n_{i,W}$ switches starting from 'No.' and $\sigma_{m_i}^{-1}(\text{Yes}) = W$. Consider the decision m_i' obtain by $\forall E \in \mathscr{E}_i$ setting:

$$m_i'(E) = \begin{cases} \text{Yes} & m_i(E) = \text{No} \\ \text{No} & m_i(E) = \text{Yes} \end{cases}$$

Note that m_i' has at most $n_{i,W}$ switches starting from 'Yes.' Further $\sigma_{m_i'}^{-1}(\text{Yes}) = \sigma_{m_i}^{-1}(\text{No}) = \Omega\backslash W$. Hence $n_{i,\Omega\backslash W} \leq n_{i,W} \leq n_i$. Suppose towards a contradiction $n_{i,\Omega\backslash W} < n_{i,W}$. Then we can construct a decision method m_i'' for i such that m_i'' has at most $n_{i,\Omega\backslash W}$ switches and $\sigma_{m_i''}^{-1}(\text{Yes}) = \Omega\backslash W$. Further, we can construct a decision method m_i''' from m_i'' (analogously to how we constructed m_i' from m_i) such that m_i''' has at most $n_{i,\Omega\backslash W}$ switches and $\sigma_{m_i'''}^{-1}(\text{Yes}) = W$. Contradiction. Hence $n_{i,\Omega\backslash W} = n_{i,W}$ and we can conclude agent i has reason to limit decide $\Omega\backslash W$ starting from 'Yes.' $\square$

Corollary 4.1: Agent i has reason to limit decide W starting from 'No' if and only if $\exists k \leq n_i$ such that W is k-open in $\mathscr{T}_i$ but not $(k-1)$-closed.[8]
Proof
Immediate by Theorem 4.1 and Lemma 4.2. $\square$

We now no longer assume $\mathscr{E}_i$ has a starting point. Instead, fix $E \in \mathscr{E}_i$ and let $\mathscr{E}_i|E = \{E' \in \mathscr{E}_i : E' \subseteq E\}$. If, in some state of affairs, E comprises the set of all worlds i considers possible then $\mathscr{E}_i|E$ consists of all pieces of evidence i might de facto learn going forward. In particular, $\mathscr{E}_i|E$ forms a basis for the subspace topology $\mathscr{T}_i|E$ over E such that $E \in \mathscr{E}_i|E$. Therefore, we can consistently relativize every bit of terminology we have introduced thus far to E by replacing Ω with E as the set of all possible worlds and replacing $\mathscr{E}_i$ with $\mathscr{E}_i|E$ so that i's information basis has a starting point. In this context, it immediately follows agent i has reason to limit decide W starting from 'Yes' if and only if $\exists k \leq n_i$ such that W is k-closed in $\mathscr{T}_i|E$ but not $(k-1)$-open. Similarly, i has reason to limit decide W starting from 'No' if and only if $\exists k \leq n_i$ such that W is k-open in $\mathscr{T}_i|E$ but not $(k-1)$-closed. Implicit in both these statements however is that W is a subset of E. So, given an arbitrary subset W of Ω, one must take its intersection with E for such statements to even make sense (after all $W \cap E$ consists of those worlds where W is true and i learns evidence E). As a consequence, in our original context, we say:

- Agent *i has reason to limit decide W starting from 'Yes' in light of evidence E* if $\exists k \leq n_i$ such that $W \cap E$ is k-closed in $\mathscr{T}_i|E$ but not $(k-1)$-open.
- Agent *i has reason to limit decide W starting from 'No' in light of evidence E* if $\exists k \leq n_i$ such that $W \cap E$ is k-open in $\mathscr{T}_i|E$ but not $(k-1)$-closed.

For concision, we abbreviate the above phrases in natural language by saying:

- *E gives i reason simpliciter to believe W.*
- *E gives i reason simpliciter to disbelieve W.*

and denote them by the predicates $\text{Yes}_i(W|E)$ and $\text{No}_i(W|E)$ respectively.

[8] Again, no set is (-1)-closed.

Lemma 4.3: Suppose $(X_j)_{j\in J}$ is a collection of $(n+1)$-open sets in $\mathscr{T}_i$. Furthermore, suppose that each X_j is the nested difference of a descending sequence of open sets $O_0^j \supseteq ... \supseteq O_n^j$. If for each $m \in J$ we have $X_m = O_0^m \cap \bigcup_{j\in J} X_j$ then $\bigcup_{j\in J} X_j$ is the nested difference of the descending sequence of open sets $\bigcup_{j\in J} O_0^j \supseteq ... \supseteq \bigcup_{j\in J} O_n^j$ and so $\bigcup_{j\in J} X_j$ is $(n+1)$-open in $\mathscr{T}_i$.
Proof
We proceed by induction on n. When $n=0$, each $X_j = O_0^j$ and so we can immediately establish the base case as $\bigcup_{j\in J} X_j = \bigcup_{j\in J} O_0^j$. For the inductive step, assume the claim holds when $n=k-1$. Take $(X_j)_{j\in J}$ to be a collection of $(k+1)$-open sets in $\mathscr{T}_i$ so that each X_j is the nested difference of a descending sequence of open sets $O_0^j \supseteq .... \supseteq O_k^j$. Suppose for each $m \in J$ we have $X_m = O_0^m \cap \bigcup_{j\in J} X_j$. Next, for each $j \in J$, let X'_j be the nested difference of the descending sequence of open sets $O_1^j \supseteq ... \supseteq O_k^j$.

We claim now that:

$$\bigcup_{j\in J} X_j = \left(\bigcup_{j\in J} O_0^j\right) \setminus \left(\bigcup_{j\in J} X'_j\right).$$

Suppose $w \in \left(\bigcup_{j\in J} O_0^j\right) \setminus \left(\bigcup_{j\in J} X'_j\right)$ then there is some $s \in J$ such that $w \in O_0^s$. Since we are guaranteed $w \notin X'_s$ this means $w \in X_s$ and so $w \in \bigcup_{j\in J} X_j$. Now suppose $w \in \bigcup_{j\in J} X_j$. Then there is some $s \in J$ such that $w \in X_s$. In particular, we know $w \in O_0^s$. Entertain for contradiction there exists some $m \in J$ so that $w \in X'_m$. Since $X_m = O_0^m \cap \bigcup_{j\in J} X_j$ we know $O_0^m \cap X_s \subseteq X_m$. Since $w \in X'_m$ we have $w \notin X_m$ and hence $w \notin O_0^m \cap X_s$. However, since $X'_m \subseteq O_0^m$, this is only possible if $w \notin X_s$. Contradiction. Therefore $w \notin \bigcup_{j\in J} X'_j$ and so $w \in \left(\bigcup_{j\in J} O_0^j\right) \setminus \left(\bigcup_{j\in J} X'_j\right)$.

Next we claim for each $m \in J$ that:

$$X'_m = O_0^m \cap \left(\bigcup_{j\in J} X_j\right)^c.$$

Suppose $w \in O_0^m \cap \left(\bigcup_{j\in J} X_j\right)^c$. Then $w \in O_0^m$ and $w \notin \bigcup_{j\in J} X_j$. Since $X_m = O_0^m \cap \bigcup_{j\in J} X_j$, we know $w \notin X_m$. Further, because $w \in O_0^m$, this is only possible if $w \in X'_m$. Now suppose $w \in X'_m$. Then $w \in O_0^m$ and $w \notin X_m$. Since $X_m = O_0^m \cap \bigcup_{j\in J} X_j$, this is only possible if $w \notin \bigcup_{j\in J} X_j$. Hence, $w \in O_0^m \cap \left(\bigcup_{j\in J} X_j\right)^c$.

Now, using the fact $X'_m \subseteq O_1^m$ and $X'_m \subseteq O_0^m \subseteq \bigcup_{j\in J} O_0^j$, we have for each $m \in J$ that:

$$X'_m = \left(O_0^m \cap \left(\bigcup_{j\in J} X_j\right)^c\right) \cap \left(O_1^m \cap \bigcup_{j\in J} O_0^j\right).$$

Since $O_1^m \subseteq O_0^m$ and $\left(\bigcup_{j\in J} X_j\right)^c \cap \left(\bigcup_{j\in J} O_0^j\right) = \bigcup_{j\in J}\left(O_0^j \cap \left(\bigcup_{j'\in J} X_{j'}\right)^c\right) = \bigcup_{j\in J} X'_j$, we can rearrange the above to obtain for each $m \in J$ that:

$$X'_m = O_1^m \cap \bigcup_{j\in J} X'_j.$$

By hypothesis then, we have $\bigcup_{j\in J} X'_j$ is the nested difference of the descending sequence of open sets $\bigcup_{j\in J} O_1^j \supseteq ... \supseteq \bigcup_{j\in J} O_k^j$. Since $\bigcup_{j\in J} X_j = \left(\bigcup_{j\in J} O_0^j\right) \setminus \left(\bigcup_{j\in J} X'_j\right)$, we can conclude $\bigcup_{j\in J} X_j$ is the nested difference of the descending sequence of open sets $\bigcup_{j\in J} O_0^j \supseteq ... \supseteq \bigcup_{j\in J} O_k^j$. □

Lemma 4.4: If W is $(n+1)$-open in $\mathscr{T}_i$ then $\forall w \in W$ there $\exists E \in \mathscr{E}_i(w)$ such that $W \cap E$ is n-closed in the subspace topology $\mathscr{T}_i|E$.
Proof
Since W is $(n+1)$-open there exists a descending sequence of open sets $O_0 \supseteq ... \supseteq O_n$ in $\mathscr{T}_i$ such that $W = O_0 \backslash (O_1 \backslash (...O_n)...)$. Suppose $w \in W$. Then $w \in O_0$ and so there exists $E^* \in \mathscr{E}_i(w)$ such that $E^* \subseteq O_0$. Additionally:

$$W \cap E^* = (O_0 \cap E^*) \backslash ((O_1 \cap E^*) \backslash (...(O_n \cap E^*)...) = E^* \backslash ((O_1 \cap E^*) \backslash (...(O_n \cap E^*))...)$$

, i.e. $W \cap E^*$ is n-closed in $\mathscr{T}_i|E^*$. $\square$

Theorem 4.2: W is (n_i+1)-open in $\mathscr{T}_i$ iff $\forall w \in W$ there $\exists E \in \mathscr{E}_i(w)$ such that $\text{Yes}_i(W|E)$ holds.
Proof
First we show the 'if' direction. Suppose $\forall w \in W$ there $\exists E \in \mathscr{E}_i(w)$ such that $\text{Yes}_i(W|E)$ holds. In particular, for each $w \in W$, take E_w to be an element of $\mathscr{E}_i(w)$ for which $\text{Yes}_i(W|E_w)$ holds so that $W = \bigcup_{w \in W} W \cap E_w$. Note for each $w \in W$ there is a decision method $m_{i,w} : \mathscr{E}_i|E_w \to \{\text{Yes}, \text{No}\}$ such that $m_{i,w}$ has at most n_i switches starting from 'Yes' and $\sigma^{-1}_{m_{i,w}}(\text{Yes}) = W \cap E_w$. Therefore, $W \cap E_w$ is the nested difference of n_i+1 descending open sets $O^w_0 \supseteq ...O^w_{n_i}$ in $\mathscr{T}_i|E_w$ with $O^w_0 = E_w$ (as per the construction in Theorem 4.0). Thus, for each $w' \in W$, we have $W \cap E_{w'} = O^{w'}_0 \cap W = O^{w'}_0 \cap \bigcup_{w \in W}(W \cap E_w)$. Applying Lemma 4.3 then, we can conclude $W = \bigcup_{w \in W} W \cap E_w$ is (n_i+1)-open in $\mathscr{T}_i$.

Now we show the 'only if' direction. Suppose W is (n_i+1)-open and $w \in W$. By Lemma 4.4, we can define $k_w \le n_i$ to be the least k for which $\exists E \in \mathscr{E}_i(w)$ so that $W \cap E$ is k-closed in $\mathscr{T}_i|E$ and let $E_w \in \mathscr{E}_i(w)$ be an information state for which $W \cap E_w$ is k_w-closed in $\mathscr{T}_i|E_w$.
If $k_w = 0$ then $W \cap E_w$ is 0-closed in $\mathscr{T}_i|E_w$ but not (-1)-open and so $\text{Yes}_i(W|E_w)$ holds.
If $k_w = 1$ then $W \cap E_w$ is 1-closed in $\mathscr{T}_i|E_w$ but not 0-open (as $w \in W \cap E_w$) and so $\text{Yes}_i(W|E_w)$ holds.
If $k_w \ge 2$ then $W \cap E_w$ is k_w-closed in $\mathscr{T}_i|E_w$. Suppose for contradiction that $W \cap E_w$ is (k_w-1)-open in $\mathscr{E}_i|E_w$. Then there is a decision method $m_i : \mathscr{E}_i|E_w \to \{\text{Yes}, \text{No}\}$ such that m_i has at most $k_w - 2$ switches after saying 'Yes' and $\sigma^{-1}_{m_i}(\text{Yes}) = W \cap E_w$. Because $w \in W \cap E_w$, there $\exists E' \in (\mathscr{E}_i|E_w)(w)$ for which $m_i(E') = \text{Yes}$. Consider the decision method $m'_i : \mathscr{E}_i|E' \to \{\text{Yes}, \text{No}\}$ which $\forall E \in \mathscr{E}_i|E'$ sets $m'_i(E) = m_i(E)$. Clearly, m'_i has at most $k^* - 2$ switches starting from 'Yes.' Consider arbitrary $w \in E'$. If $\sigma_{m_i}(w) = \text{Yes}$ then $\exists E'' \in (\mathscr{E}_i|E_w)(w)$ so that $\forall E \in (\mathscr{E}_i|E'')(w)$ we have $m_i(E) = \text{Yes}$. In particular, if we take E''' to be a subset $E' \cap E''$ containing w, then this means $\forall E \in (\mathscr{E}_i|E''')(w)$ we have $m'_i(E) = m_i(E) = \text{Yes}$. Hence, $\sigma_{m'_i}(w) = \text{Yes}$. Symmetrically, if $\sigma_{m_i}(w) = \text{No}$ then $\sigma_{m'_i}(w) = \text{No}$. Thus, it follows $\forall w \in E'$ that $\sigma_{m'_i}(w) = \sigma_{m_i}(w)$. Namely, $\sigma^{-1}_{m'_i}(\text{Yes}) = W \cap E'$ and so $W \cap E'$ is the nested difference of $k_w - 1$ descending open sets $O_0 \supseteq ... \supseteq O_{k_w-2}$ in $\mathscr{T}_i|E'$ with $O_0 = E'$ (as per the construction in Theorem 4.0), i.e. $W \cap E'$ is (k_w-2)-closed in $\mathscr{E}_i|E'$. However, this contradicts minimality of k_w. Therefore, $W \cap E_w$ is k_w-closed in $\mathscr{T}_i|E_w$ but not (k_w-1)-open and so $\text{Yes}_i(W|E_w)$ holds. $\square$

Corollary 4.2: W is (n_i+1)-closed in $\mathscr{T}_i$ iff $\forall w \in \Omega \backslash W$ there $\exists E \in \mathscr{E}_i(w)$ such that $\text{No}_i(W|E)$ holds.
Proof
Immediate by Theorem 4.2 and the fact $\text{No}_i(W|E)$ holds iff $\text{Yes}_i(\Omega \backslash W|E)$ holds. $\square$

In the next section, we will semantically interpret each of the pre-theoretic phrases we introduced earlier. For now, this concludes our overview of topological learning theory. We refer the curious reader to Kelly 1996 for a more comprehensive treatment of the subject as well as its relationship to computability.

5 Semantics

Define a frame $\mathscr{F}$ to be a pair $(\Omega,(\mathscr{E}_i,n_i)_{i\in N})$ where $\Omega\neq\varnothing$ and for each $i\in N$ we have that $(\mathscr{E}_i,n_i)$ consists of an information basis together with a switching tolerance for agent i.

Next, given a frame $\mathscr{F}=(\Omega,(\mathscr{E}_i,n_i)_{i\in N})$, for each $i\in N$ define the operator $\mathbf{R}_i:2^\Omega\to 2^\Omega$ so that $\forall W\subseteq\Omega$:

$$\mathbf{R}_i(W)=\{w\in\Omega:\exists E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\}$$

, i.e. $\mathbf{R}_i(W)$ corresponds to those worlds where i has reason simpliciter to believe W.

Additionally, mirroring our discussion in the pre-theory section, given $W\subseteq\Omega$ define the operators $\mathbf{I}_{i@W}:2^\Omega\to 2^\Omega$ and $\mathbf{B}_{i@W}:2^\Omega\to 2^\Omega$ by $\forall P\subseteq\Omega$ setting:

$$\mathbf{I}_{i@W}(P)=\{w\in\Omega:\forall E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\to W\cap E\subseteq P\}\text{ and }\mathbf{B}_{i@W}(P)=\mathbf{R}_i(W)\cap\mathbf{I}_{i@W}(P)$$

,i.e. $\mathbf{I}_{i@W}(P)$ and $\mathbf{B}_{i@W}(P)$ correspond to the worlds where W indicates to i that P and i has W as reason to believe P respectively.

Now, define $\mathbf{S}_i:2^\Omega\to 2^\Omega$ so that $\forall P\subseteq\Omega$:

$$\mathbf{S}_i(P)=\{w\in\Omega:\exists W\subseteq\Omega,w\in W\cap\mathbf{B}_{i@W}(P)\}$$

, i.e. $\mathbf{S}_i(P)$ corresponds to those worlds where i has some true reason to believe P.

Further, $\forall k\in\mathbb{N}^+$, let $\mathbf{E}^k_W(P)$ denote the result of applying the map $X\mapsto\bigcap_{i\in N}\mathbf{B}_{i@W}(X)$ iteratively k times on P. Then, given $W\subseteq\Omega$, define $\mathbf{G}_W:2^\Omega\to 2^\Omega$ by $\forall P\subseteq\Omega$ setting:

$$\mathbf{G}_W(P)=\bigcap_{k\in\mathbb{N}^+}\mathbf{E}^k_W(P)$$

, i.e. $\mathbf{G}_W(P)$ corresponds to those worlds where W generates common inductive knowledge of P.

Finally, recall from the pre-theory section 'P is common inductive knowledge' is supposed to be the weakest proposition entailing P such that if it holds then everyone has some true reason to believe it. However, this notion is well-defined relative to our frame $\mathscr{F}$ if and only if we can guarantee the existence of a greatest prefixed point for the map $X\mapsto P\cap\bigcap_{i\in N}\mathbf{S}_i(X)$. We prove this is the case below.

Lemma 5.0: $\mathbf{S}_i$ is monotone.
<u>Proof</u>
Suppose $P_1\subseteq P_2$ and $w\in\mathbf{S}_i(P_1)$. Then for some $W\subseteq\Omega$ there exists $E^*\in\mathscr{E}_i(w)$ such that $\mathrm{Yes}_i(W|E)$. Further, $\forall E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\to W\cap E\subseteq P_1$. In particular, $\forall E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\to W\cap E\subseteq P_2$ as $P_1\subseteq P_2$. Hence, $w\in\mathbf{S}_i(P_2)$. $\square$

Theorem 5.0: The map $X\mapsto P\cap\bigcap_{i\in N}\mathbf{S}_i(X)$ has a greatest (pre)fixed point.
<u>Proof</u>
Since each $\mathbf{S}_i$ is monotone, the map $X\mapsto P\cap\bigcap_{i\in N}\mathbf{S}_i(X)$ is monotone and so Tarski's theorem guarantees the existence of a greatest fixed point which is equal to its greatest prefixed point. $\square$

But what can be said of this fixed point beyond its existence? In order to provide a more 'constructive' characterization of what it means for P to be common inductive knowledge, we now study how each operator $\mathbf{S}_i$ relates to the topology $\mathscr{T}_i$ generated by $\mathscr{E}_i$.

Lemma 5.1: Let int_i be the interior operator with respect to the topology $\mathscr{T}_i$. If $n_i = 0$ then $\forall P \subseteq \Omega$ we have that $\mathbf{S}_i(P) = \text{int}_i(P)$.
Proof
Firstly, note $\text{Yes}_i(W|E)$ holds for some piece of evidence $E \in \mathscr{E}_i$ iff $W \cap E$ is 0-closed in $\mathscr{T}_i|E$, i.e. $E \subseteq W$. Now suppose $w \in \text{int}_i(P)$. Then $\exists W \in \mathscr{E}_i(w)$ such that $W \subseteq P$. Thus we immediately have that $\forall E \in \mathscr{E}_i(w), \text{Yes}_i(W|E) \to W \cap E \subseteq P$. Further, since $w \in W$ and $\text{Yes}_i(W|W)$ holds, we can conclude $\exists E \in \mathscr{E}_i(w)$ such that $\text{Yes}_i(W|E)$ holds and therefore $w \in \mathbf{S}_i(P)$. Next suppose $w \in \mathbf{S}_i(P)$. Then $\exists W \subseteq \Omega$ containing w for which $\exists E^* \in \mathscr{E}_i(w), E^* \subseteq W$ and $\forall E \in \mathscr{E}_i(w), E \subseteq W \to W \cap E \subseteq P$. In particular, this means $W \cap E^* \subseteq P$ and so $E^* \subseteq P$, i.e. $w \in \text{int}_i(P)$. □

This leads us to ask whether, in general, $\mathbf{S}_i$ acts like an 'interior' operator with respect to some topology over Ω. Towards that end, the Kuratowski interior axioms say a map $\mathbf{i} : 2^\Omega \to 2^\Omega$ can be viewed as the interior operator with respect to some topology $\mathscr{T}_\mathbf{i}$ over Ω iff:

1. $\mathbf{i}$ *preserves the total space*, i.e. we have $\mathbf{i}(\Omega) = \Omega$
2. $\mathbf{i}$ *is intensive*, i.e. $\forall P \subseteq \Omega$ we have $\mathbf{i}(P) \subseteq P$
3. $\mathbf{i}$ *is idempotent*, i.e. $\forall P \subseteq \Omega$ we have $\mathbf{i}(P) = \mathbf{i}(\mathbf{i}(P))$
4. $\mathbf{i}$ *preserves binary intersections*, i.e. $\forall P_1, P_2 \subseteq \Omega$ we have $\mathbf{i}(P_1) \cap \mathbf{i}(P_2) = \mathbf{i}(P_1 \cap P_2)$

Indeed, it turns out that $\mathbf{S}_i$ satisfies all four of Kuratowski's axioms.

Theorem 5.1: $\mathbf{S}_i$ preserves the total space.
Proof
It suffices to show $\Omega \subseteq \mathbf{S}_i(\Omega)$. Suppose $w \in \Omega$. Since $\mathscr{E}_i(w)$ is non-empty, $\exists E^* \in \mathscr{E}_i(w)$. In particular, $\Omega \cap E^* = E^*$ is 0-closed in $\mathscr{T}_i|E^*$ and so $\text{Yes}_i(\Omega|E^*)$ holds. Further $\forall E \in \mathscr{E}_i(w), \text{Yes}_i(\Omega|E) \to \Omega \cap E \subseteq \Omega$. Since $\Omega \subseteq \Omega$, we can conclude $w \in \mathbf{S}_i(\Omega)$. □

Theorem 5.2: $\mathbf{S}_i$ is intensive.[9]
Proof
Suppose $w \in \mathbf{S}_i(P)$. Then for some $W \subseteq \Omega$ containing w, we have that $\exists E^* \in \mathscr{E}_i(w)$ such that $\text{Yes}_i(W|E^*)$ holds. Further, $\forall E \in \mathscr{E}_i(w), \text{Yes}_i(W|E) \to W \cap E \subseteq P$. Thus $W \cap E^* \subseteq P$. Since $w \in E^*$, this means we can conclude $w \in P$. □

Theorem 5.3: $\mathbf{S}_i$ is idempotent.
Proof
Because $\mathbf{S}_i$ is intensive we know $\mathbf{S}_i(\mathbf{S}_i(P)) \subseteq \mathbf{S}_i(P)$ and so it suffices to show $\mathbf{S}_i(P) \subseteq \mathbf{S}_i(\mathbf{S}_i(P))$. Suppose $w \in \mathbf{S}_i(P)$. Then there exists $W \subseteq \Omega$ containing w such that $\forall E \in \mathscr{E}_i(w), \text{Yes}_i(W|E) \to W \cap E \subseteq P$. Take arbitrary $E' \in \mathscr{E}_i(w)$ for which $\text{Yes}_i(W|E')$ holds. Then, $\exists k \leq n_i$ so that $W \cap E'$ is k_i-closed in $\mathscr{T}_i|E'$ but not $(k_i - 1)$-open. Hence, $\text{Yes}_i(W \cap E'|E')$ also holds. Take $w' \in W \cap E'$. Because $E' \in \mathscr{E}_i(w')$, we know

[9] By Theorems 5.1 and 5.2, we can transfinitely iterate the map $X \mapsto P \cap \bigcap_{i \in N} \mathbf{S}_i(X)$ starting with Ω and obtain its greatest fixed point is the transfinite limit of the statements 'Everyone has some true reason to believe P,' 'Everyone has some true reason to believe 'Everyone has some true reason to believe P.''

$\exists E \in \mathscr{E}_i(w')$ for which $\text{Yes}_i(W \cap E'|E)$ holds. Further, $\forall E \in \mathscr{E}_i(w'), \text{Yes}_i(W \cap E'|E) \to (W \cap E') \cap E \subseteq P$ as $W \cap E' \subseteq P$. Hence, $w' \in \mathbf{S}_i(P)$ and so $W \cap E' \subseteq \mathbf{S}_i(P)$. Since E' was arbitrary, this means that $\forall E \in \mathscr{E}_i(w), \text{Yes}_i(W|E) \to W \cap E \subseteq \mathbf{S}_i(P)$. Also, because $w \in \mathbf{S}_i(P)$, there exists $E^* \in \mathscr{E}_i(w)$ for which $\text{Yes}_i(W|E^*)$ holds. Therefore, we can conclude $w \in \mathbf{S}_i(\mathbf{S}_i(P))$. □

By monotonicity of $\mathbf{S}_i$, we already have $\forall P_1, P_2 \subseteq \Omega$ that $\mathbf{S}_i(P_1 \cap P_2) \subseteq \mathbf{S}_i(P_1)$ and $\mathbf{S}_i(P_1 \cap P_2) \subseteq \mathbf{S}_i(P_2)$, i.e. $\mathbf{S}_i(P_1 \cap P_2) \subseteq \mathbf{S}_i(P_1) \cap \mathbf{S}_i(P_2)$. Hence, for $\mathbf{S}_i$ to preserves intersections, it suffices to show $\forall P_1, P_2 \subseteq \Omega$ that $\mathbf{S}_i(P_1) \cap \mathbf{S}_i(P_2) \subseteq \mathbf{S}_i(P_1 \cap P_2)$. Intuitively then, whenever $\mathbf{S}_i(P_1)$ and $\mathbf{S}_i(P_2)$ hold, we want a principled procedure which picks out some true reason i has to believe P_1 as well as some true reason i has to believe P_2 and combines them to produce some true reason i has for believing $P_1 \cap P_2$. We now prove a few lemmas that will help construct exactly this procedure.

Lemma 5.2: $\mathbf{S}_i(P)$ is the union of all (n_i+1)-open sets in $\mathscr{T}_i$ which entail P.
Proof
Let $P(n_i)$ be the collection of all (n_i+1)-open sets in $\mathscr{T}_i$ which entail P. Suppose $w \in \bigcup_{W \in P(n_i)} W$. Then $\exists W \subseteq \Omega$ containing w such that W is (n_i+1)-open in $\mathscr{T}_i$ and $W \subseteq P$. Further, as $W \subseteq P$, we immediately have $\forall E \in \mathscr{E}_i(w), \text{Yes}_i(W|E) \to W \cap E \subseteq P$. In particular, by Theorem 4.2, $\exists E^* \in \mathscr{E}_i(w)$ for which $\text{Yes}_i(W|E)$ holds. Hence, $w \in \mathbf{S}_i(P)$. Now suppose $w \in \mathbf{S}_i(P)$. Then for some $W' \subseteq \Omega$ containing w there exists $E^* \in \mathscr{E}_i(w)$ such that $\text{Yes}_i(W'|E^*)$ holds and $W' \cap E^* \subseteq P$. In particular, $W' \cap E^*$ is n_i-closed in $\mathscr{T}_i|E^*$ and so it is (n_i+1)-open in $\mathscr{T}_i$. Finally, because $w \in W' \cap E^*$, we have that $W' \cap E^* \in P(n_i)$ and so $w \in \bigcup_{W \in P(n_i)} W$. Thus, $\mathbf{S}_i(P) = \bigcup_{W \in P(n_i)} W$ is the union of all (n_i+1)-open sets in $\mathscr{T}_i$ entailing P. □

Lemma 5.3: $\forall n > 0$, P is a union of $(n+1)$-open sets in $\mathscr{T}_i$ iff P is a union of 2-open sets in $\mathscr{T}_i$.
Proof
Since the 'if' direction is trivial, we focus on the 'only if' direction. Suppose P is a union of $(n+1)$-open sets in $\mathscr{T}_i$, i.e. there exists a collection $\mathscr{C}$ of $(n+1)$-open sets in $\mathscr{T}_i$ such that $P = \cup_{W \in \mathscr{C}} W$. Fix $W \in \mathscr{C}$. Since W is $(n+1)$-open in $\mathscr{T}_i$ there exists a descending sequence of open sets $W_0 \supseteq \ldots \supseteq W_n$ in $\mathscr{T}_i$ such that $W = W_0 \backslash (W_1 \backslash (\ldots W_n) \ldots)$. For ease of argument, let $W_{n+1} = \varnothing$. Note $w \in W$ iff there is an even $k \neq n+1$ for which $w \in W_k \backslash W_{k+1}$, i.e. $W = \cup_{k \leq n : k \text{ is even}} W_k \backslash W_{k+1}$. Since each $W_k \backslash W_{k+1}$ is 2-open in $\mathscr{T}_i$, it follows that P is a union of 2-open sets in $\mathscr{T}_i$. □

Lemma 5.4: If $n_i > 0$ then $\mathbf{S}_i(P)$ is the union of all 2-open sets in $\mathscr{T}_i$ which entail P.
Proof
Let $P(2)$ be the collection of all 2-open sets in $\mathscr{T}_i$ which entail P. Suppose $w \in \bigcup_{W \in P(2)} W$. Then $\exists W \subseteq \Omega$ containing w such that W is 2-open in $\mathscr{T}_i$ and $W \subseteq P$. Note W is certainly (n_i+1)-open in $\mathscr{T}_i$ as it is 2-open. Hence, by Lemma 5.2, $w \in \mathbf{S}_i(P)$. Now suppose $w \in \mathbf{S}_i(P)$. Then, by Lemma 5.2, w belongs to a set W' which is (n_i+1)-open $\mathscr{T}_i$ and entails P. Additionally, by Lemma 5.3, w belongs to a set W which is 2-open in $\mathscr{T}_i$ and entails W'. Finally, because $W \subseteq P$, we can conclude $w \in \bigcup_{W \in P(2)} W$. Thus, $\mathbf{S}_i(P) = \bigcup_{W \in P(2)} W$ is the union of all 2-open sets in $\mathscr{T}_i$ entailing P. □

The above implies that whenever $n_i > 0$ if $w \in \mathbf{S}_i(P)$ then w is contained in a 2-open set W of $\mathscr{T}_i$ which entails P. Thus, if i has some true reason to believe P then i also has some true 2-open set as reason to believe P. We will think of the principled procedure we are constructing as picking out such 2-open sets as reasons in case $n_i > 0$. In the next lemma, we show that if one 'combines' two 2-open reasons by taking their intersection then the resulting reason is also 2-open in $\mathscr{T}_i$.

Lemma 5.5: If W_1 and W_2 are 2-open in $\mathscr{T}_i$ then so is $W_1 \cap W_2$.
Proof
Since W_1 and W_2 are 2-open in $\mathscr{T}_i$, there exists descending sequences of open sets $O_0 \supseteq O_1$ and $O'_0 \supseteq O'_1$ such that $W_1 = O_0 \backslash O_1$ and $W_2 = O'_0 \backslash O'_1$. In particular:

$$W_1 \cap W_2 = (O_0 \cap O'_0) \backslash (O_1 \cup O'_1) = (O_0 \cap O'_0) \backslash ((O_0 \cap O'_0) \cap (O_1 \cup O'_1))$$

Since $(O_0 \cap O'_0) \supseteq ((O_0 \cap O'_0) \cap (O_1 \cup O'_1))$, we can conclude $W_1 \cap W_2$ is 2-open in $\mathscr{T}_i$. $\square$

Conveniently then, using Lemmas 5.4 and 5.5, we can show...

Theorem 5.4: $\mathbf{S}_i$ preserves binary intersections.
Proof
By monotonicity, it suffices to show $\mathbf{S}_i(P_1) \cap \mathbf{S}_i(P_2) \subseteq \mathbf{S}_i(P_1 \cap P_2)$. When $n_i = 0$ we know $\mathbf{S}_i$ is just int_i and so we are done. Hence, suppose $n_i > 0$ and $w \in \mathbf{S}_i(P_1) \cap \mathbf{S}_i(P_2)$. By Lemma 5.4, for some $W_1, W_2 \subseteq \Omega$ containing w which are 2-open in $\mathscr{T}_i$, we have that $W_1 \subseteq P_1$ and $W_2 \subseteq P_2$. By Lemma 5.5, $W_1 \cap W_2$ is 2-open in $\mathscr{T}_i$. Thus w is contained in a 2-open set of $\mathscr{T}_i$ which entails $P_1 \cap P_2$, i.e. $w \in \mathbf{S}_i(P_1 \cap P_2)$. $\square$

As promised then, by Theorems 5.1-5.4, $\mathbf{S}_i$ satisfies the Kuratowski interior axioms. If agent i is a deductive learner with $n_i = 0$ then the topology $\mathscr{T}_{\mathbf{S}_i}$ over Ω generated by $\mathbf{S}_i$ is just $\mathscr{T}_i$. Alternatively, if agent i is a truly inductive learner with $n_i > 0$, Lemma 5.4 and Lemma 5.5 tell us $\mathscr{T}_{\mathbf{S}_i}$ can be generated by taking the 2-open sets of $\mathscr{T}_i$ as basis elements. In particular, so long as $n_i > 0$, the topology $\mathscr{T}_{\mathbf{S}_i}$ is invariant of n_i. Having detailed the properties of $\mathbf{S}_i$, we now return to the question of how we can more constructively characterize the worlds where P is common inductive knowledge.

Theorem 5.5: Let int_N be the interior operator with respect to the intersection topology $\bigcap_{i \in N} \mathscr{T}_{\mathbf{S}_i}$. The greatest fixed point of the map $X \mapsto P \cap \bigcap_{i \in N} \mathbf{S}_i(X)$ is $\text{int}_N(P)$.
Proof
Suppose X is a fixed point of the map $X \mapsto P \cap \bigcap_{i \in N} \mathbf{S}_i(X)$, i.e. $X = P \cap \bigcap_{i \in N} \mathbf{S}_i(X)$. Clearly X is a subset of P. Additionally for each $i \in N$, we have $X \subseteq \mathbf{S}_i(X)$ and so $X = \mathbf{S}_i(X)$ by intensivity. But this means X is open in each $\mathscr{T}_{\mathbf{S}_i}$, i.e. X is open in $\bigcap_{i \in N} \mathscr{T}_{\mathbf{S}_i}$. Hence, it follows that $X \subseteq \text{int}_N(P)$. Next, since $\text{int}_N(P)$ is a subset of P and open in each $\mathscr{T}_i$, we know $\text{int}_N(P) = P \cap \bigcap_{i \in N} \mathbf{S}_i(\text{int}_N(P))$. As a result, $\text{int}_N(P)$ is the greatest fixed point of the map $X \mapsto P \cap \bigcap_{i \in N} \mathbf{S}_i(X)$. $\square$

Therefore, define the operator $\mathbf{C} : 2^{\Omega} \to 2^{\Omega}$ so that $\forall P \subseteq \Omega$:

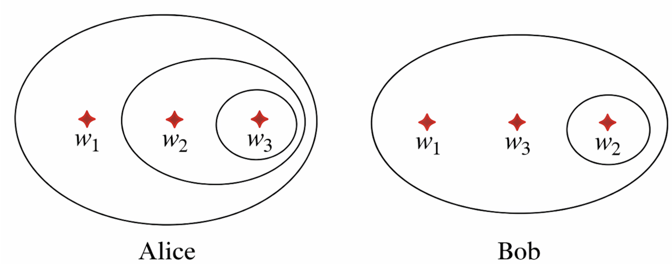
$$\mathbf{C}(P) = \text{int}_N(P)$$

, i.e. $\mathbf{C}(P)$ corresponds to those worlds where P is common inductive knowledge. Note if each $n_i > 0$ then $\bigcap_{i \in N} \mathscr{T}_{\mathbf{S}_i}$ is invariant of the n_i and so $\mathbf{C}(P)$ is as well. Thus, if agents are truly inductive learners, then the worlds where P is common inductive knowledge is invariant of agents' switching tolerances. We view this property of our semantics as a feature and not a bug. After all, a good notion of common inductive knowledge should be fairly robust to changes in agents' inductive standards.

On the other hand, recall from the pre-theory section that Lewis says P is common inductive knowledge if and only if:

> some state of affairs holds such that it generates common inductive knowledge of P.

Lewis' account lacks the robustness of our definition as, for any witnessing state of affairs which holds, whether that witness generates common inductive knowledge substantially varies with agents' switching tolerances. To illustrate, suppose $\Omega = \{w_1, w_2, w_3\}$ and take $P = \{w_1, w_3\}$. Further, suppose we have $|N| = 2$ agents named Alice and Bob (who we abbreviate by writing a and b respectively). Finally, suppose Alice and Bob each have information bases over Ω given by the below diagrams:



If $n_a = n_b = 1$ it is easy to see the only witness W that Bob has as reason to believe P in any world is $W = \{w_1, w_3\}$. One can also compute straightforwardly that $\mathbf{B}_{a@W}(P) = \{w_3\}$. However, this means $\mathbf{B}_{b@W}(\mathbf{B}_{a@W}(P)) = \varnothing$ and so $\mathbf{G}_W(P) = \varnothing$. Hence, under Lewis' definition, P is not common inductive knowledge in any world. Alternatively if $n_a = 2$, one can compute that $\mathbf{B}_{a@W}(P) = \mathbf{B}_{b@W}(P) = \Omega$ and so $\mathbf{G}_W(P) = \Omega$. Moreover, on Lewis' account, we suddenly have P is common inductive knowledge in every world P holds as $W \cap \mathbf{G}_W(P) = P$. In fact, one can generalize the above example so Lewis' definition shifts from saying P is not common inductive knowledge in any world to saying P is common inductive knowledge in every world P holds at arbitrarily high switching tolerances for Alice. By contrast, for all values of $n_a, n_b > 0$, our definition of common inductive knowledge stably maintains that $\mathbf{C}(P) = P$. This can be seen visually using the topological characterization of $\mathbf{C}$ in Theorem 5.5 almost immediately. For a more epistemic argument, note $W_1 = \{w_1\} = \Omega \backslash \{w_2, w_3\}$ and $W_3 = \{w_3\}$ are 2-open sets in $\mathscr{T}_a$ which entail P. Hence, Alice has W_1 as true reason to believe P at w_1 and W_3 as true reason to believe P at w_3, i.e. $\mathbf{S}_a(P) = P$. Similarly, $W_1 \cup W_3 = \Omega \backslash \{w_2\}$ is a 2-open set in $\mathscr{T}_b$ which entails P. Therefore, Bob has $W_1 \cup W_3$ as true reason to believe P at both w_1 and w_3, i.e. $\mathbf{S}_b(P) = P$. Thus, P is the weakest proposition entailing P such that if it is true then Alice and Bob have some true reason to believe it, i.e. $\mathbf{C}(P) = P$.

Though likely obvious, we now conclude by stating how to interpret all the formulas φ in our logical language $\mathscr{L}$. Namely, say a model $\mathscr{M}$ is a pair $(\mathscr{F}, v)$ where $\mathscr{F}$ is a frame over Ω and $v : \mathrm{PROP} \to 2^{\Omega}$ is valuation function. Then $\forall w \in \Omega$, given a primitive proposition $p \in \mathrm{PROP}$ we write $(\mathscr{M}, w) \vDash p$ iff $w \in v(p)$, i.e. $v(p)$ yields an interpretation of those worlds where p holds in $\mathscr{M}$. Next, we extend v so that its domain consists of all formulas φ in $\mathscr{L}$ by recursively writing:[10][11]

$$
\begin{aligned}
&(\mathscr{M}, w) \vDash \mathbb{R}_i \varphi \text{ iff } w \in \mathbf{R}_i(v(\varphi)) \\
&(\mathscr{M}, w) \vDash \mathbb{I}_{i@\varphi_1} \varphi_2 \text{ iff } w \in \mathbf{I}_{i@v(\varphi_1)}(v(\varphi_2)) \\
&(\mathscr{M}, w) \vDash \mathbb{B}_{i@\varphi_1} \varphi_2 \text{ iff } w \in \mathbf{B}_{i@v(\varphi_1)}(v(\varphi_2)) \\
&(\mathscr{M}, w) \vDash \mathbb{S}_i \varphi \text{ iff } w \in \mathbf{S}_i(v(\varphi)) \\
&(\mathscr{M}, w) \vDash \mathbb{G}_{\varphi_1} \varphi_2 \text{ iff } w \in \mathbf{G}_{v(\varphi_1)}(v(\varphi_2)) \\
&(\mathscr{M}, w) \vDash \mathbb{C} \varphi \text{ iff } w \in \mathbf{C}(v(\varphi))
\end{aligned}
$$

and $\forall \varphi \in \mathscr{L}$ setting $v(\varphi) = \{w \in \Omega : (\mathscr{M}, w) \vDash \varphi\}$.

[10] Classical propositional connectives are interpreted in their standard manner.
[11] All operators on the right hand side are interpreted relative to the frame $\mathscr{F}$.

In the sequel, we establish soundness of our proof system with respect to this semantics. Readers eager to see the relevance of our logic to distributed consensus should feel free to skip this section.

6 Soundness

$\mathbf{Ax}_{\mathsf{R}}:\ \mathbb{R}_i\varphi \leftrightarrow \mathbb{R}_i\mathbb{R}_i\varphi$

Proof

Let $\mathscr{F}=(\Omega,(\mathscr{E}_i,n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W be any subset of Ω. It suffices to show $\mathbf{R}_i(W)=\mathbf{R}_i(\mathbf{R}_i(W))$. Note that $\mathbf{R}_i(W)=\{w\in\Omega:\exists E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\}$ is the union of all information states E in $\mathscr{E}_i$ for which $\mathrm{Yes}_i(W|E)$ holds. Hence, $\mathbf{R}_i(W)$ is 1-open in $\mathscr{T}_i$ and therefore also (n_i+1)-open. By Theorem 4.2, this means $\mathbf{R}_i(W)\subseteq\mathbf{R}_i(\mathbf{R}_i(W))=\{w\in\Omega:\exists E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\}$. We claim also that $\mathbf{R}_i(\mathbf{R}_i(W))\subseteq\mathbf{R}_i(W)$. So suppose $w\in\mathbf{R}_i(\mathbf{R}_i(W))$. Then $\exists E^*\in\mathscr{E}_i(w)$ so that $\mathrm{Yes}_i(\mathbf{R}_i(W)|E^*)$ holds, i.e. $\exists k\leq n_i$ so that $\mathbf{R}_i(W)\cap E^*$ is k-closed but not $(k-1)$-open in $\mathscr{T}_i|E^*$. Since $\mathbf{R}_i(W)$ is 1-open in $\mathscr{T}_i$, we know $\mathbf{R}_i(W)\cap E^*$ is also 1-open in $\mathscr{T}_i|E^*$. Therefore, it must be that $\mathbf{R}_i(W)\cap E^*$ is 0-closed in $\mathscr{T}_i|E^*$, i.e. $E^*\subseteq\mathbf{R}_i(W)$. Thus, since $w\in E^*$, we conclude $w\in\mathbf{R}_i(W)$. □

$\mathbf{Ax}_{\mathsf{I1}}:\ \mathbb{I}_{i@\varphi}\varphi$

Proof

Let $\mathscr{F}=(\Omega,(\mathscr{E}_i,n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W be any subset of Ω. It suffices to show $\forall w\in\Omega$ that $w\in\mathbf{I}_{i@W}(W)$. Take arbitrary $w\in\Omega$. Note $\forall E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\to W\cap E\subseteq W$ as we always know $W\cap E\subseteq W$. By definition then, $w\in\mathbf{I}_W(W)$. □

$\mathbf{Ax}_{\mathsf{I2}}:\ \mathbb{I}_{i@\varphi}\mathbb{R}_i\varphi$

Proof

Let $\mathscr{F}=(\Omega,(\mathscr{E}_i,n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W be any subset of Ω. It suffices to show $\forall w\in\Omega$ that $w\in\mathbf{I}_{i@W}(\mathbf{R}_i(W))$. Take arbitrary $w\in\Omega$ and arbitrary $E^*\in\mathscr{E}_i(w)$. If $\mathrm{Yes}_i(W|E^*)$ holds then note $\forall w'\in E^*$ that $w'\in\mathbf{R}_i(W)$, i.e. $E^*\subseteq\mathbf{R}_i(W)$. Therefore is also the case $W\cap E^*\subseteq\mathbf{R}_i(W)$. Since E^* was arbitrary, this means $\forall E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\to W\cap E\subseteq\mathbf{R}_i(W)$. By definition then, $w\in\mathbf{I}_{i@W}(\mathbf{R}_i(W))$. □

$\mathbf{Ax}_{\mathsf{I3}}:\ \mathbb{I}_{i@\varphi_1}\mathbb{I}_{i@\varphi_1}\varphi_2\to\mathbb{I}_{i@\varphi_1}\varphi_2$

Proof

Let $\mathscr{F}=(\Omega,(\mathscr{E}_i,n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W,P be any subsets of Ω. It suffices to show $\mathbf{I}_{i@W}(\mathbf{I}_{i@W}(P))\subseteq\mathbf{I}_{i@W}(P)$. Take arbitrary $w\in\mathbf{I}_{i@W}(\mathbf{I}_{i@W}(P))$ and arbitrary $E^*\in\mathscr{E}_i(w)$. If $\mathrm{Yes}_i(W|E^*)$ holds then $W\cap E^*\subseteq\mathbf{I}_{i@W}(P)$. Then $\forall w'\in W\cap E^*$, since $w'\in\mathbf{I}_{i@W}(P)$ and $E^*\in\mathscr{E}_i(w')$, it must be that $w'\in P$ and so $W\cap E^*\subseteq P$. Since E^* was arbitrary, $\forall E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\to W\cap E\subseteq P$. By definition then, $w\in\mathbf{I}_{i@W}(P)$. □

$\mathbf{Ax}_{\mathsf{I4}}:\ (\mathbb{I}_{i@\varphi_1}\varphi_2\wedge\mathbb{I}_{i@(\varphi_1\wedge\varphi_2)}\varphi_3)\to\mathbb{I}_{i@\varphi_1}\varphi_3$

Proof

Let $\mathscr{F}=(\Omega,(\mathscr{E}_i,n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W,P,Q be any subsets of Ω. It suffices to show $\mathbf{I}_{i@W}(P)\cap\mathbf{I}_{i@(W\cap P)}(Q)\subseteq\mathbf{I}_{i@W}(Q)$. Take arbitrary $w\in\mathbf{I}_{i@W}(P)\cap\mathbf{I}_{i@(W\cap P)}(Q)$ and arbitrary $E^*\in\mathscr{E}_i(w)$. If $\mathrm{Yes}_i(W|E^*)$ holds then $W\cap E^*\subseteq P$ and so $W\cap E^*=(W\cap P)\cap E^*$. Hence, if $\mathrm{Yes}_i(W|E^*)$ holds then $\mathrm{Yes}_i(W\cap P|E^*)$ also holds and so $W\cap E^*=(W\cap P)\cap E^*\subseteq Q$. Since E^* was arbitrary, this means $\forall E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\to W\cap E\subseteq Q$. By definition then, $w\in\mathbf{I}_{i@W}(Q)$. □

$\mathbf{Ax}_{\mathsf{I5}}:\ (\mathbb{I}_{i@\varphi_1}\varphi_2 \wedge \mathbb{I}_{i@\varphi_1}(\varphi_2 \to \varphi_3)) \to \mathbb{I}_{i@\varphi_1}\varphi_3$

<u>Proof</u>

Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W, P, Q be any subsets of Ω. It suffices to show $\mathbf{I}_{i@W}(P) \cap \mathbf{I}_{i@W}(P \to Q) \subseteq \mathbf{I}_{i@W}(Q)$. Take arbitrary $w \in \mathbf{I}_{i@W}(P) \cap \mathbf{I}_{i@W}(P \to Q)$ and arbitrary $E^* \in \mathscr{E}_i(w)$. If $\text{Yes}_i(W|E^*)$ holds then $W \cap E^* \subseteq P$ and $W \cap E^* \subseteq (\Omega \backslash P) \cup Q$. Hence, $W \cap E^* \subseteq P \cap ((\Omega\backslash P) \cup Q) = Q$. Since E^* was arbitrary, this means $\forall E \in \mathscr{E}_i(w), \text{Yes}_i(W|E) \to W \cap E \subseteq Q$. By definition then, $w \in \mathbf{I}_{i@W}(Q)$. $\square$

$\mathbf{Ax}_{\mathsf{B1}}:\ \mathbb{B}_{i@\varphi_1}\varphi_2 \leftrightarrow (\mathbb{R}_i\varphi_1 \wedge \mathbb{I}_{i@\varphi_1}\varphi_2)$

<u>Proof</u>

Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W, P be any subsets of Ω. It suffices to show $\mathbf{B}_{i@W}(P) = \mathbf{R}_i(W) \cap \mathbf{I}_{i@W}(P)$, which holds by definition. $\square$

$\mathbf{Ax}_{\mathsf{B2}}:\ \mathbb{B}_{i@\varphi_1}\varphi_2 \to \mathbb{R}_i(\varphi_1 \wedge \varphi_2)$

<u>Proof</u>

Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W, P be any subsets of Ω. It suffices to show $\mathbf{B}_{i@W}(P) \subseteq \mathbf{R}_i(W \cap P)$. Take arbitrary $w \in \mathbf{B}_{i@W}(P)$. Then $w \in \mathbf{R}_i(P)$ and so $\exists E^* \in \mathscr{E}_i(w)$ for which $\text{Yes}_i(W|E^*)$ holds. Since $w \in \mathbf{I}_{i@W}(P)$, this means $W \cap E^* \subseteq P$ and thus $W \cap E^* = (W \cap P) \cap E^*$. Hence, $\text{Yes}_i(W \cap P|E^*)$ also holds and so $w \in \mathbf{R}_i(W \cap P)$. $\square$

$\mathbf{Ax}_{\mathsf{S1}}:\ (\varphi_1 \wedge \mathbb{B}_{i@\varphi_1}\varphi_2) \to \mathbb{S}_i\varphi_2$

<u>Proof</u>

Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W, P be any subsets of Ω. It suffices to show $W \cap \mathbf{B}_{i@W}(P) \subseteq \mathbf{S}_i(P)$, which holds by definition. $\square$

$\mathbf{Ax}_{\mathsf{S2}}:\ \mathbb{S}_i\varphi \to \varphi$

<u>Proof</u>

Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and P be any subset of Ω. It suffices to show $\mathbf{S}_i(P) \subseteq P$. We proved exactly this in Theorem 5.2. $\square$

$\mathbf{Ax}_{\mathsf{S3}}:\ \mathbb{I}_{i@\varphi_1}\varphi_2 \to \mathbb{I}_{i@\varphi_1}\mathbb{S}_i\varphi_2$

<u>Proof</u>

Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W, P be any subsets of Ω. It suffices to show $\mathbf{I}_{i@W}(P) \subseteq \mathbf{I}_{i@W}(\mathbf{S}_i(P))$. Take arbitrary $w \in \mathbf{I}_{i@W}(P)$ and $E^* \in \mathscr{E}_i(w)$. Suppose $\text{Yes}_i(W|E^*)$ holds. Note then $\mathbf{I}_{i@(W\cap E^*)}(P) = \{w'' \in \Omega : \forall E \in \mathscr{E}_i(w''), \text{Yes}_i(W \cap E^*|E) \to (W \cap E^*) \cap E \subseteq P\} = \Omega$ as $W \cap E^* \subseteq P$. Furthermore, we know $\text{Yes}_i(W \cap E^*|E^*)$ holds as we are guaranteed there $\exists k \leq n_i$ so that $(W \cap E^*) \cap E^* = W \cap E^*$ is k-closed in $\mathscr{T}_i|E^*$ but not $(k-1)$-open in $\mathscr{T}_i|E^*$. Consider $w' \in W \cap E^*$. Since $E^* \in \mathscr{E}_i(w')$ it follows $w' \in \mathbf{R}_i(W \cap E^*)$. Additionally, $w' \in \mathbf{I}_{i@(W\cap E^*)}(P)$ and so $w' \in \mathbf{B}_{i@(W\cap E^*)}(P)$. By our proof $\mathbf{Ax}_{\mathsf{S1}}$, we know $(W \cap E^*) \cap \mathbf{B}_{i@(W\cap E^*)}(P) \subseteq \mathbf{S}_i(P)$. Thus, $w' \in \mathbf{S}_i(P)$ and so $W \cap E^* \subseteq \mathbf{S}_i(P)$. Namely, we can conclude $w \in \mathbf{I}_{i@W}(\mathbf{S}_i(P))$ as desired. $\square$

$\mathbf{Ax}_{\mathsf{S4}}:\ \mathbb{S}_i\varphi \leftrightarrow \mathbb{S}_i\mathbb{S}_i\varphi$

<u>Proof</u>

Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and P be any subset of Ω. It suffices to show $\mathbf{S}_i(P) = \mathbf{S}_i(\mathbf{S}_i(P))$. We proved exactly this in Theorem 5.3. $\square$

$\mathbf{Ax}_{\mathsf{S5}}: \ (\mathbb{S}_i\varphi_1 \wedge \mathbb{S}_i\varphi_2) \leftrightarrow \mathbb{S}_i(\varphi_1 \wedge \varphi_2)$
Proof
Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and P_1, P_2 be any subsets of Ω. It suffices to show $\mathbf{S}_i(P_1) \cap \mathbf{S}_i(P_2) = \mathbf{S}_i(P_1 \cap P_2)$. We proved exactly this in Theorem 5.4. □

$\mathbf{Ax}_{\mathsf{G1}}: \ \mathbb{G}_{\varphi_1}\varphi_2 \to \mathbb{B}_{i@\varphi_1}\varphi_2$
Proof
Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W, P be any subsets of Ω. It suffices to show $\mathbf{G}_W(P) \subseteq \mathbf{B}_{i@W}(P)$, which holds by definition. □

$\mathbf{Ax}_{\mathsf{G2}}: \ \mathbb{G}_{\varphi_1}\varphi_2 \to \mathbb{B}_{i@\varphi_1}\mathbb{G}_{\varphi_1}\varphi_2$
Proof
Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and W, P be any subsets of Ω. It suffices to show $\mathbf{G}_W(P) \subseteq \mathbf{B}_{i@W}(\mathbf{G}_W(P))$. Take arbitrary $w \in \mathbf{G}_W(P) = \bigcap_{k\in\mathbb{N}^+}\mathbf{E}^k_W(P)$. Since we know $w \in \mathbf{E}^1_W(P) = \bigcap_{j\in N}\mathbf{B}_{j@W}(P)$ it follows $w \in \mathbf{B}_{i@W}(P) \subseteq \mathbf{R}_i(W)$. Additionally, since $\forall k \in \mathbb{N}^+$ we know $w \in \mathbf{E}^{k+1}_W(P) = \bigcap_{j\in N}\mathbf{B}_{j@W}(\mathbf{E}^k_W(P))$ it follows $w \in \mathbf{B}_{i@W}(\mathbf{E}^k_W(P)) \subseteq \mathbf{I}_{i@W}(\mathbf{E}^k_W(P))$. Take arbitrary $E^* \in \mathscr{E}_i(w)$. If $\mathrm{Yes}_i(W|E^*)$ holds then $\forall k \in \mathbb{N}^+$ we know $W \cap E^* \subseteq \mathbf{E}^k_W(P)$, i.e. we have $W \cap E^* \subseteq \mathbf{G}_W(P)$. Hence, as E^* was arbitrary, $w \in \mathbf{I}_{i@W}(\mathbf{G}_W(P))$ and so $w \in \mathbf{B}_{i@W}(\mathbf{G}_W(P))$. □

$\mathbf{Ax}_{\mathsf{C1}}: \ \mathbb{C}\varphi \to \varphi$
Proof
Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame and P be any subset of Ω. It suffices to show $\mathbf{C}(P) \subseteq P$. Recall $\mathbf{C}(P)$ is a fixed point of the map $X \mapsto P \cap \bigcap_{i\in N}\mathbf{S}_i(X)$. Thus, $\mathbf{C}(P) = P \cap \bigcap_{i\in N}\mathbf{S}_i(\mathbf{C}(P)) \subseteq P$. □

$\mathbf{Ax}_{\mathsf{C2}}: \ \mathbb{C}\varphi \to \mathbb{S}_i\mathbb{C}\varphi$
Proof
Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame, i be an agent in N, and P be any subset of Ω. It suffices to show $\mathbf{C}(P) \subseteq \mathbf{S}_i(\mathbf{C}(P))$. Recall $\mathbf{C}(P)$ is a fixed point of the map $X \mapsto P \cap \bigcap_{j\in N}\mathbf{S}_j(X)$. Thus, $\mathbf{C}(P) = P \cap \bigcap_{j\in N}\mathbf{S}_j(\mathbf{C}(P)) \subseteq \mathbf{S}_i(\mathbf{C}(P))$ as desired. □

$\mathbf{R}_{\mathsf{R}}: \ \frac{\varphi}{\mathbb{R}_i\varphi}$
Proof
Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame and i be an agent in N. It suffices to show $\forall w \in \Omega$ that $w \in \mathbf{R}_i(\Omega)$. Take arbitrary $w \in \Omega$. Then there exists some $E^* \in \mathscr{E}_i(w)$. Further, since $\Omega \cap E^* = E^*$ is 0-closed in $\mathscr{T}_i|E^*$, we know $\mathrm{Yes}_i(W|E^*)$ holds. Therefore $w \in \mathbf{R}_i(\Omega)$. □

$\mathbf{R}_{\mathsf{I}}: \ \frac{\varphi_2}{\mathbb{I}_{i@\varphi_1}\varphi_2}$
Proof
Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame and i be an agent in N and W be any subset of Ω. It suffices to show $\forall w \in \Omega$ that $w \in \mathbf{I}_{i@W}(\Omega)$. Note $\forall E \in \mathscr{E}_i(w), \mathrm{Yes}_i(W|E) \to W \cap E \subseteq \Omega$ as we always know $W \cap E \subseteq \Omega$. By definition then, $w \in \mathbf{I}_W(\Omega)$. □

While all the proofs up to this point have been relatively direct, showing that the rule $\mathbf{R}_{\mathsf{G}}$ is sound ends up being rather involved. To simplify our argument, we now establish two key lemmas...

Lemma 6.0: Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame and W, P be any subsets of Ω. Then the map $X \mapsto \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X))$ is monotone.
Proof
Suppose $X_1 \subseteq X_2$. Take arbitrary $w \in \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X_1))$. Since $\forall i \in N$ we know $w \in \mathbf{B}_{i@W}(P)$ it follows $\forall i \in N$ that $w \in \mathbf{R}_i(W)$. Take arbitrary $E^* \in \mathscr{E}_i(w)$. If $\text{Yes}_i(W|E^*)$ holds then $W \cap E^* \subseteq X_1$ and hence $W \cap E^* \subseteq X_2$. Hence, as E^* was arbitrary, we have $\forall i \in N$ that $w \in \mathbf{I}_{i@W}(X_2)$. Thus, $\forall i \in N$ it follows $w \in \mathbf{B}_{i@W}(X_2)$. In particular, $w \in \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X_2))$. Therefore, the map $X \mapsto \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X))$ is monotone. $\square$

Lemma 6.1: Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame and W, P be any subsets of Ω. Then the map $X \mapsto \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X))$ commutes with intersection.
Proof
Let $(X_j)_{j\in J}$ be an arbitrary collection of subsets of Ω. Note first that $\bigcap_{j\in J} \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X_j)) = \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \bigcap_{j\in J} \mathbf{B}_{i@W}(X_j)) = \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \bigcap_{j\in J} \mathbf{I}_{i@W}(X_j))$. Now observe $w \in \cap_{j\in J} \mathbf{I}_{i@W}(X_j)$ iff $\forall j \in J$ we have $\forall E \in \mathscr{E}_i(w), \text{Yes}_i(W|E) \to W \cap E \subseteq X_j$. Therefore, $w \in \bigcap_{j\in J} \mathbf{I}_{i@W}(X_j)$ just in case it holds that $\forall E \in \mathscr{E}_i(w), \text{Yes}_i(W|E) \to W \cap E \subseteq \cap_{j\in J} X_j$, i.e. iff $w \in \mathbf{I}_{i@W}(\cap_{j\in J} X_j)$. Hence, we have $\bigcap_{j\in J} \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X_j)) = \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{I}_{i@W}(\cap_{j\in J} X_j)) = \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(\cap_{j\in J} X_j))$. Thus, the map $X \mapsto \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X))$ commutes with intersection. $\square$

$\mathbf{R}_{\mathsf{G}}: \frac{\varphi_3 \to \bigwedge_{i\in N} \mathbb{B}_{i@\varphi_1} \varphi_3 \quad \varphi_3 \to \bigwedge_{i\in N} \mathbb{B}_{i@\varphi_1} \varphi_2}{\varphi_3 \to \mathbb{G}_{\varphi_1} \varphi_2}$
Proof
Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame and W, P, Q be any subsets of Ω. It suffices to show if $Q \subseteq \bigcap_{i\in N} \mathbf{B}_{i@W}(Q)$ and $Q \subseteq \bigcap_{i\in N} \mathbf{B}_{i@W}(P)$ then $Q \subseteq \mathbf{G}_W(P)$. By Lemma 6.0, Tarski's implies the map $X \mapsto \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X))$ has a greatest (pre)fixed point; by Lemma 6.1, if we let $\mathbf{e}^k_W(P)$ be the result of applying the map $X \mapsto \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X))$ iteratively k times on Ω, Kleene's theorem implies this greatest (pre)fixed point is given by $\bigcap_{k\in\mathbb{N}^+} \mathbf{e}^k_W(P)$. Observe, by monotonicity, that $\Omega \supseteq \mathbf{e}^1(P) \supseteq \mathbf{e}^2(P)...$ forms a descending sequence of subsets of Ω. We will now show $\forall k \in \mathbb{N}^+$ that $\mathbf{e}^k_W(P) = \mathbf{E}^k_W(P)$ and so $\mathbf{G}_W(P) = \bigcap_{k\in\mathbb{N}^+} \mathbf{e}^k_W(P)$. We proceed by induction on k. When $k = 1$, we know $\mathbf{e}^1_W(P) = \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(P)) = \mathbf{E}^1_W(P)$ establishing the base case. For the inductive step, assume $\forall k \leq t$ we have $\mathbf{e}^k_W(P) = \mathbf{E}^k_W(P)$. Note $\mathbf{e}^{t+1}_W(P) = \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(\mathbf{e}^t_W(P))) = \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(\mathbf{E}^t_W(P))) = \mathbf{E}^1_W(P) \cap \mathbf{E}^{t+1}_W(P)$. Take arbitrary $w \in \mathbf{E}^{t+1}_W(P) = \bigcap_{i\in N} \mathbf{B}_{i@W}(\mathbf{E}^t_W(P))$. Then $\forall i \in N$ we have $w \in \mathbf{R}_i(W)$ and $w \in \mathbf{I}_{i@W}(\mathbf{E}^t_W(P))$. Take arbitrary $E^* \in \mathscr{E}_i(w)$. If $\text{Yes}_i(W|E^*)$ holds then $W \cap E^* \subseteq \mathbf{E}^t_W(P)$. By hypothesis, we know $\mathbf{E}^t_W(P) \subseteq ... \subseteq \mathbf{E}^1_W(P)$. Hence, if $\text{Yes}_i(W|E^*)$ holds then $W \cap E^* \subseteq \mathbf{E}^1_W(P)$. Namely $\forall w' \in W \cap E^*$, since $w' \in \mathbf{E}^1_W(P) \subseteq \mathbf{I}_{i@W}(P)$ and $E^* \in \mathscr{E}_i(w')$, it must be that $w' \in P$ and so $W \cap E^* \subseteq P$. Hence, $w \in \mathbf{I}_{i@W}(P)$ and therefore $w \in \mathbf{E}^1(P) = \bigcap_{i\in N} \mathbf{B}_{i@W}(P)$. In particular, as $\mathbf{E}^{t+1}(P) \subseteq \mathbf{E}^1(P)$, this means $\mathbf{e}^{t+1}_W(P) = \mathbf{E}^{t+1}_W(P)$ thereby completing the inductive step. Since $Q \subseteq \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(Q))$, it is a prefixed point of $X \mapsto \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X))$. Since the greatest (pre)fixed point of $X \mapsto \bigcap_{i\in N} (\mathbf{B}_{i@W}(P) \cap \mathbf{B}_{i@W}(X))$ is $\mathbf{G}_W(P)$, it follows $Q \subseteq \mathbf{G}_W(P)$. $\square$

$\mathbf{R}_{\mathsf{C}}: \frac{\varphi_1 \to \bigwedge_{i\in N} \mathbb{S}_i \varphi_1 \quad \varphi_1 \to \varphi_2}{\varphi_1 \to \mathbb{C} \varphi_2}$
Proof
Let $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$ be an arbitrary frame and W, P be any subsets of Ω. It suffices to show if $W \subseteq \bigcap_{i\in N} \mathbf{S}_i(W)$ and $W \subseteq P$ then $W \subseteq \mathbf{C}(P)$. If $W \subseteq \bigcap_{i\in N} \mathbf{S}_i(W)$ and $W \subseteq P$, by Theorem 5.2, $W = P \cap \bigcap_{i\in N} \mathbf{S}_i(W)$, i.e. it is a fixed point of $X \mapsto P \cap \bigcap_{i\in N} \mathbf{S}_i(X)$. By definition then, $W \subseteq \mathbf{C}(P)$. $\square$

7 Inductive Coordinated Attack

Consider the following 'inductive' variant of the coordinated attack problem:[12]

> Each agent $i \in N$ is deciding whether to attest P is true or defer judgement. The maximum number of times agent i is willing change their mind after first saying 'Yes' to P is n_i.
>
> - If P is true, each agent i ideally wants everyone to converge on attesting P is true.
> - However, if P is false or some agent i does not converge on attesting P is true, then anyone who converges on attesting P is true is perpetually shamed.
>
> What non-trivial attestation protocols are available to the agents, if any, so that they can guarantee none of them will be perpetually shamed?

More formally, suppose we have a frame $\mathscr{F} = (\Omega, (\mathscr{E}_i, n_i)_{i\in N})$. An *attestation strategy* for i is a map $s_i : \mathscr{E}_i \to \{\text{Yes}, ?\}$ such that the decision method $m_i : \mathscr{E}_i \to \{\text{Yes}, \text{No}\}$ which $\forall E \in \mathscr{E}_i$ sets $m_i(E) = \text{Yes}$ iff $s_i(E) = \text{Yes}$ has at most n_i switches after saying 'Yes.' Additionally, as with decision methods, define $\sigma_{s_i} : \Omega \to \{\text{Yes}, ?\}$ to be the function which for each $w \in \Omega$ returns $\sigma_{s_i}(w)$ as the output s_i converges to in world w. Finally, an *attestation protocol* for N is a tuple $(s_i)_{i\in N}$ of attestation strategies for each $i \in N$. We say an attestation protocol $(s_i)_{i\in N}$ *solves coordinated attack for* P iff it satisfies:

1. **Validity-** $\forall i \in N$, $\sigma_{s_i}^{-1}(\text{Yes}) \subseteq P$.
2. **Agreement-** $\forall i, j \in N$, $\sigma_{s_i}^{-1}(\text{Yes}) = \sigma_{s_j}^{-1}(\text{Yes})$.
3. **Nontriviality-** $\bigcap_{i\in N} \sigma_{s_i}^{-1}(\text{Yes}) \neq \varnothing$.

Validity ensures that agents will never falsely converge on attesting P is true. Agreement ensures that agents will all converge on the same output in every world. Validity and agreement together ensure that no one is perpetually shamed. A trivial way to satisfy validity and agreement is to take the attestation protocol where each agent always defers judgement on P. To rule out such pathological protocols, the nontriviality condition further ensures there is at least one world where everyone converges on attesting P is true. In particular, given an attestation protocol $(s_i)_{i\in N}$ which solves consensus for P, we will say that $\bigcap_{i\in N} \sigma_{s_i}^{-1}(\text{Yes})$ is the *success set* of $(s_i)_{i\in N}$.

Theorem 7.0: There exists an attestation protocol $(s_i)_{i\in N}$ which solves coordinated attack for P with success set W if and only if W is a non-empty (n_i+1)-open subset of P in each $\mathscr{T}_i$.

<u>Proof</u>

First we show the 'if' direction. So suppose W is a non-empty (n_i+1)-open subset of P in each $\mathscr{T}_i$. By Theorem 4.0, $\forall i \in N$ there is a decision method m_i has at most n_i-switches after saying 'Yes' and $\sigma_{\mathrm{m}_i}^{-1}(\text{Yes}) = W$. Consider the attestation protocol $(s_i)_{i\in N}$ which $\forall i \in N$ sets $s_i(E) = \text{Yes}$ iff $\mathrm{m}_i(E) = \text{Yes}$. Clearly, $\forall i \in N$ we have $\sigma_{s_i}^{-1}(\text{Yes}) = \sigma_{\mathrm{m}_i}^{-1}(\text{Yes}) = W \subseteq P$. Hence, because W is non-empty, it immediately follows that $(s_i)_{i\in N}$ solves coordinated attack for P with success set W.

Now we show the 'only if' direction. Suppose there exists an attestation protocol $(s_i)_{i\in N}$ which solves coordinated attack for P with success set W. Agreement guarantees $\forall i \in N$ that $\sigma_{s_i}^{-1}(\text{Yes}) = W$. Nontriviality and validity further guarantee W is a non-empty subset of P. Additionally, $\forall i \in N$ the decision method $m_i : \mathscr{E}_i \to \{\text{Yes}, \text{No}\}$ which sets $m_i(E) = \text{Yes}$ iff $s_i(E) = \text{Yes}$ has at most n_i switches after saying 'Yes.' Since $\sigma_{m_i}^{-1}(\text{Yes}) = \sigma_{s_i}^{-1}(\text{Yes}) = W$, this means W is (n_i+1)-open in each $\mathscr{T}_i$. □

[12] Halpern & Moses 1990 were the first to analyze coordinated attack using a logic for common knowledge. To our awareness, this variant (despite its clear applications for designing distributed consensus protocols) has not received any such treatment.

Theorem 7.0 implies that the possible success sets of an attestation protocol solving coordinated attack for P are precisely those non-empty subsets of P which are (n_i+1)-open in each $\mathscr{T}_i$. Additionally, given any such subset W of P, Theorem 4.0 allows us to explicitly construct at least one attestation protocol solving coordinated attack for P with success set W. In effect then, if one regards two attestation protocols solving coordinated attack for P as equally successful whenever they have the same success sets, the problem of solving coordinated attack for P can be reduced to the task of simply finding those non-empty subsets of P which are (n_i+1)-open in each $\mathscr{T}_i$.

Theorem 7.1: W is a (n_i+1)-open subset of P in each $\mathscr{T}_i$ if and only if W is a fixed point of the map $X \mapsto X \cap \mathbf{G}_X(P)$. Further, $\forall W \subseteq \Omega$, $W \cap \mathbf{G}_W(P)$ is itself such a fixed point.

Proof

Suppose W is a (n_i+1)-open subset of P in each $\mathscr{T}_i$. By our proof of $\mathbf{Ax}_{\mathsf{I}1}$, we know $\forall i \in N$ that $\mathbf{I}_{i@W}(W)=\Omega$. Additionally, by Theorem 4.2, we have $W \subseteq \mathbf{R}_i(W)=\{w\in\Omega : \exists E \in \mathscr{E}_i(w), \mathrm{Yes}_i(W|E)\}$. Hence, $W \subseteq \mathbf{R}_i(W)\cap \mathbf{I}_{i@W}(W)=\mathbf{B}_{i@W}(W)$ and so $W \subseteq \bigcap_{i\in N}\mathbf{B}_{i@W}(W)$. Next, $\forall i \in N$ we have that $\mathbf{I}_{i@W}(P)=\{w\in\Omega : \forall E\in\mathscr{E}_i(w), \mathrm{Yes}_i(W|E)\to W\cap E\subseteq P\}=\Omega$ as $W\subseteq P$. In particular, it follows $W\subseteq \mathbf{R}_i(W)\cap\mathbf{I}_{i@W}(P)=\mathbf{B}_{i@W}(P)$ and so $W\subseteq\bigcap_{i\in N}\mathbf{B}_{i@W}(P)$. Therefore, using our proof of $\mathbf{R}_{\mathsf{G}}$, this means $W\subseteq\mathbf{G}_W(P)$, i.e. $W=W\cap\mathbf{G}_W(P)$.
Now suppose $W=W\cap\mathbf{G}_W(P)$. By our proof of $\mathbf{Ax}_{\mathsf{G}1}$, we know $\forall i\in N$ that $\mathbf{G}_W(P)\subseteq\mathbf{B}_{i@W}(P)$. Hence, $W\subseteq W\cap\mathbf{B}_{i@W}(P)$. By our proof of $\mathbf{Ax}_{\mathsf{S}1}$, we have $W\cap\mathbf{B}_{i@W}(P)\subseteq\mathbf{S}_i(P)$. Further, by Theorem 5.2, $\mathbf{S}_i(P)\subseteq P$. Therefore, $W\subseteq P$. Also by definition, we know that $\mathbf{B}_{i@W}(P)\subseteq\mathbf{R}_i(W)$. Hence, $W\subseteq\mathbf{R}_i(W)=\{w\in\Omega:\exists E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W|E)\}$ and so W is (n_i+1)-open in $\mathscr{T}_i$ by Theorem 4.2. Thus, W is a (n_i+1)-open subset of P in each $\mathscr{T}_i$.
Finally consider arbitrary $W\subseteq\Omega$. We claim $W\cap\mathbf{G}_W(P)=(W\cap\mathbf{G}_W(P))\cap\mathbf{G}_{W\cap\mathbf{G}_W(P)}(P)$. Since the RHS is clearly a subset of the LHS, it suffices to show the LHS is a subset of the RHS. By our proof of $\mathbf{Ax}_{\mathsf{G}2}$, we know $\forall i\in N$ that $\mathbf{G}_W(P)\subseteq\mathbf{B}_{i@W}(\mathbf{G}_W(P))$. Also, we know $\mathbf{B}_{i@W}(\mathbf{G}_W(P))\subseteq\mathbf{R}_i(W\cap\mathbf{G}_W(P))$ by our proof of $\mathbf{Ax}_{\mathsf{B}2}$. Thus, $W\cap\mathbf{G}_W(P)\subseteq\mathbf{R}_i(W\cap\mathbf{G}_W(P))$. Next, by our proof of $\mathbf{Ax}_{\mathsf{I}1}$, we know $\forall i\in N$ that $\mathbf{I}_{i@(W\cap\mathbf{G}_W(P))}(W\cap\mathbf{G}_W(P))=\Omega$. Since $\mathbf{R}_i(W\cap\mathbf{G}_W(P))\cap\mathbf{I}_{i@(W\cap\mathbf{G}_W(P))}(W\cap\mathbf{G}_W(P))=\mathbf{B}_{i@(W\cap\mathbf{G}_W(P))}(W\cap\mathbf{G}_W(P))$, we have $W\cap\mathbf{G}_W(P)\subseteq\mathbf{B}_{i@(W\cap\mathbf{G}_W(P))}(W\cap\mathbf{G}_W(P))$. In particular, this means $W\cap\mathbf{G}_W(P)\subseteq\bigcap_{i\in N}\mathbf{B}_{i@(W\cap\mathbf{G}_W(P))}(W\cap\mathbf{G}_W(P))$. By our proof of $\mathbf{Ax}_{\mathsf{B}1}$, we know $\mathbf{G}_W(P)\subseteq\mathbf{B}_{i@W}(P)$ and so $W\cap\mathbf{G}_W(P)\subseteq W\cap\mathbf{B}_{i@W}(P)$. By our proof of $\mathbf{Ax}_{\mathsf{S}1}$, $W\cap\mathbf{B}_{i@W}(P)\subseteq\mathbf{S}_i(P)$. Further, by Theorem 5.2, we have $\mathbf{S}_i(P)\subseteq P$. Thus, $W\cap\mathbf{G}_W(P)\subseteq P$. Using this fact then, $\forall i\in N$ it follows that we have $\mathbf{I}_{i@(W\cap\mathbf{G}_W(P))}(P)=\{w\in\Omega:\forall E\in\mathscr{E}_i(w),\mathrm{Yes}_i(W\cap\mathbf{G}_W(P)|E)\to(W\cap\mathbf{G}_W(P))\cap E\subseteq P\}=\Omega$. Since $\mathbf{R}_i(W\cap\mathbf{G}_W(P))\cap\mathbf{I}_{i@(W\cap\mathbf{G}_W(P))}(P)=\mathbf{B}_{i@(W\cap\mathbf{G}_W(P))}(P)$, we know $W\cap\mathbf{G}_W(P)\subseteq\mathbf{B}_{i@(W\cap\mathbf{G}_W(P))}(P)$. In particular, this means $W\cap\mathbf{G}_W(P)\subseteq\bigcap_{i\in N}\mathbf{B}_{i@(W\cap\mathbf{G}_W(P))}(P)$. Therefore, appealing to our proof of $\mathbf{R}_{\mathsf{G}}$, we have $W\cap\mathbf{G}_W(P)\subseteq\mathbf{G}_{W\cap\mathbf{G}_W(P)}(P)$ and so $W\cap\mathbf{G}_W(P)=(W\cap\mathbf{G}_W(P))\cap\mathbf{G}_{W\cap\mathbf{G}_W(P)}(P)$. $\square$

Hence, we can use the map $X\mapsto X\cap\mathbf{G}_X(P)$ to pick out any subset of P which is (n_i+1)-open in each $\mathscr{T}_i$ and thereby construct (modulo success sets) all those attestation protocols which solve coordinated attack for P. Additionally, if one defines the operator $\mathbf{L}: 2^\Omega\to 2^\Omega$ by $\forall P\subseteq\Omega$ setting:

$$\mathbf{L}(P)=\{w\in\Omega:\exists W\subseteq\Omega, w\in W\cap\mathbf{G}_W(P)\}$$

then $\mathbf{L}(P)$ corresponds to those worlds which belong to the success set of some attestation protocol that solves coordinated attack for P. Furthermore, note these are exactly those worlds where some state of affairs holds such that it generates common inductive knowledge of P! Thus, $\mathbf{L}(P)$ also corresponds to those worlds where P is *Lewisian* common inductive knowledge.

Lemma 7.0: $\mathbf{L}(P) \subseteq \mathbf{C}(P)$.
Proof
Suppose $w \in \mathbf{L}(P)$. Then $\exists W \subseteq \Omega$ for which $w \in W \cap \mathbf{G}_W(P)$. By Theorem 7.1, $\forall i \in N$ we have $W \cap \mathbf{G}_W(P)$ is (n_i+1)-open in each $\mathscr{T}_i$. Hence, by Lemma 5.2, $W \cap \mathbf{G}_W(P) \subseteq \mathbf{S}_i(W \cap \mathbf{G}_W(P))$. In particular, $W \cap \mathbf{G}_W(P) \subseteq \bigcap_{i \in N} \mathbf{S}_i(W \cap \mathbf{G}_W(P))$. While proving soundness of $\mathbf{Ax}_{\mathsf{G1}}$, we showed that $\forall i \in N$ we have $\mathbf{G}_W(P) \subseteq \mathbf{B}_{i@W}(P)$. Hence, $W \cap \mathbf{G}_W(P) \subseteq W \cap \mathbf{B}_{i@W}(P)$. By our proof of $\mathbf{Ax}_{\mathsf{S1}}$, we know $W \cap \mathbf{B}_{i@W}(P) \subseteq \mathbf{S}_i(P)$. Further, by Theorem 5.2, $\mathbf{S}_i(P) \subseteq P$. Thus, $W \cap \mathbf{G}_W(P) \subseteq P$. By our proof of $\mathbf{R}_{\mathsf{C}}$, this means $W \cap \mathbf{G}_W(P) \subseteq \mathbf{C}(P)$ and therefore $w \in \mathbf{C}(P)$ as desired. $\square$

By Lemma 7.0, since the success set of any attestation protocol which solves coordinated attack for P is a subset of $\mathbf{L}(P)$, we have that $\mathbf{C}(P)$ is also an upper bound for any such success set. When $\mathbf{C}(P)$ is non-empty, if Ω is finite and each n_i is sufficiently high, we can show that there is an attestation protocol solving coordinated attack for P whose success set attains this upper bound. Hence, asymptotically speaking, an attestation protocol solving coordinated attack for P is 'welfare-maximizing' iff its success set is $\mathbf{C}(P)$ as the worlds outside $\mathbf{C}(P)$ do not belong to the success set of any such attestation protocol.

We now prove an important lemma necessary to establish this result...

Lemma 7.1: Suppose Ω is finite. For each i there exists an n so that $\mathbf{C}(P)$ is $(n+1)$-open in $\mathscr{T}_i$. In particular, we can define $n^*_{i,P}$ to be the least such n for every agent $i \in N$.
Proof
Since Ω is finite $\mathscr{T}_i$ only has finitely many 2-open sets. By Lemma 5.4 then, $\mathbf{C}_i(P)$ is a finite union of 2-open sets in $\mathscr{T}_i$. Let $W_0 = O_0 \backslash O'_0, ..., W_t = O_t \backslash O'_t$ be a finite sequence of 2-open sets where each $O_k, O'_k \in \mathscr{T}_i$ and $O_k \supseteq O'_k$ so that $\mathbf{C}_i(P) = \bigcup_{k \le t} W_k$. Consider the decision method $m_i : \mathscr{E}_i \to \{\text{Yes}, \text{No}\}$ which $\forall E \in \mathscr{E}_i$ sets:

$$m_i(E) = \begin{cases} \text{Yes} & \exists k \le t \text{ such that } E \subseteq O_k \text{ and } W_k \cap E \neq \varnothing \\ \text{No} & \text{else} \end{cases}$$

We will first show m_i has at most $2t+1$ switches after saying 'Yes.' Towards that end, suppose for sake of contradiction there exists a $(2t+2)$-switching sequence $E_0 \supseteq E'_0 \supseteq ...E_t \supseteq E'_t \supseteq E_{t+1}$ for m_i starting from 'Yes.' Then $\forall s \le t+1$, because $m_i(E_s) = \text{Yes}$, there $\exists k_s \le t$ such that $E_s \subseteq O_{k_s}$ and $W_{k_s} \cap E_s \neq \varnothing$. Further, since $m_i(E'_s) = \text{No}$ and $E'_s \subseteq E_s \subseteq O_{k_s}$, we know $W_{k_s} \cap E'_s = \varnothing$. Now note $\forall s \le t$ if $s' > s$ then $E'_s \supseteq E_{s'}$ and so $W_{k_s} \cap E_{s'} = \varnothing$ as well, i.e. $k_{s'} \neq k_s$. Hence, every element of the sequence $k_0, ...k_{t+1}$ is distinct. Contradiction. Thus m_i has at most $2t+1$ switches after saying 'Yes.'
Now we show $\sigma^{-1}_{m_i}(\text{Yes}) = \mathbf{C}(P)$. Suppose $w \in \mathbf{C}(P)$. Then $\exists k_w \le t$ so that $w \in W_{k_w} = O_{k_w} \backslash O'_{k_w}$. Hence $\exists E_w \in \mathscr{E}_i(w)$ such that $E_w \subseteq O_{k_w}$. Furthermore $\forall E' \in \mathscr{E}_i(w)$, we have $W_{k_w} \cap E' \neq \varnothing$ as $w \in E'$ and $w \in W_{k_w}$. Thus, it follows $\forall E' \in \mathscr{E}_i(w)$ if $E' \subseteq E_w$ then $m_i(E') = \text{Yes}$, i.e. $w \in \sigma^{-1}_{m_i}(\text{Yes})$. On the other hand, suppose that $w \notin \mathbf{C}(P)$. Take arbitrary $k \le t$. We have that either $w \notin O_k$ or $w \in O'_k$. In case $w \notin O_k$ then $\forall E' \in \mathscr{E}_i(w)$ we know $E' \nsubseteq O_k$. Since $\mathscr{E}_i(w)$ is non-empty, we can pick any $E^*_k \in \mathscr{E}_i(w)$ so that $\forall E' \in \mathscr{E}_i(w)$ if $E' \subseteq E^*_k$ then $E' \nsubseteq O_k$. Alternatively, in case $w \in O'_k$, we can pick $E^*_k \in \mathscr{E}_i(w)$ so that $\forall E' \in \mathscr{E}_i(w)$ if $E' \subseteq E^*_k$ then $E' \subseteq O'_k$ and so $W_k \cap E' = (O_k \backslash O'_k) \cap E' = \varnothing$. Let $E^* \in \mathscr{E}_i(w)$ be an information state such that $E^* \subseteq \bigcap_{k \le t} E^*_k$. By construction, $\forall E' \in \mathscr{E}_i(w)$ if $E' \subseteq E^*$ then $\forall k \le t$ either $E' \nsubseteq O_k$ or $W_k \cap E' = \varnothing$ as $E' \subseteq E^*_k$, i.e. $m_i(E') = \text{No}$. Hence, $\sigma_{m_i}(w) = \text{No}$ and so $w \notin \sigma^{-1}_{m_i}(\text{Yes})$.
Thus, taking $n = 2t+1$ and applying Theorem 4.0, we can conclude $\mathbf{C}(P)$ is $(n+1)$-open in $\mathscr{T}_i$. $\square$

Theorem 7.2: Suppose Ω is finite and each $n_i \geq n^*_{i,P}$. If $\mathbf{C}(P)$ is non-empty then there is an attestation protocol $(s_i)_{i\in N}$ solving coordinated attack for P whose success set is $\mathbf{C}(P)$. Further, $(s_i)_{i\in N}$ is 'welfare-maximizing' in the sense that $\mathbf{L}(P) = \mathbf{C}(P)$.[13]

Proof

By Lemma 7.1, $\mathbf{C}(P)$ is $(n^*_{i,P}+1)$-open in each $\mathscr{T}_i$. Hence, $\mathbf{C}(P)$ is (n_i+1)-open in each $\mathscr{T}_i$. Furthermore, since $\mathbf{C}(P)$ is non-empty by assumption and is always a subset of P, it follows by Theorem 7.0 that there an attestation protocol $(s_i)_{i\in N}$ solving coordinated attack for P whose success set is $\mathbf{C}(P)$. Additionally, since $\mathbf{C}(P)$ is the success set of an attestation protocol which solves coordinated attack for P, we have that $\mathbf{C}(P) \subseteq \mathbf{L}(P)$. Thus, applying Lemma 7.0, we obtain $\mathbf{L}(P) = \mathbf{C}(P)$. □

8 Future Directions & Conclusion

In this paper, we gave a topological semantics for how common inductive knowledge can be generated by a shared witness. A natural extension of our logic would be to study how common inductive knowledge can be generated by a set of separate 'local' witnesses for each agent. In the deductive case, Gonczarwoski & Moses 2024 have recently introduced a notion of asynchronous common knowledge where each agent's knowledge is similarly defined with respect to a local event. While their semantics is articulated in the runs-and-systems framework, one can also provide a topological semantics which essentially detemporalizes their definitions. Crucially however, asynchronous common knowledge remains a deductive notion and future work should generalize it to an inductive context. Such an extension would also allow us to examine what additional rules of inference arise in this richer setting.

While we demonstrated that the proof system presented in this paper is sound with respect to the proposed semantics, a corresponding completeness result has not yet been obtained. We conjecture completeness holds, although it is possible that additional axioms or rules of inference will be necessary to fully capture our semantics. Regardless, any completeness result will have to confront a central technical difficulty stemming from the presence of a non-normal modality in our language– namely, that of 'having reason simpliciter to believe.' This operator fails to satisfy the distribution axiom and places standard completeness techniques for normal modal logics out of reach. A natural first step is therefore to restrict attention to a normal fragment of our language. One might then aim to establish completeness for that fragment by translating our semantics into an appropriate class of Kripke frames.

Lastly, we showed that our semantics lends itself to finding 'attestation protocols' which solve an inductive variant of the well-known coordinated attack problem. The solutions to this problem ensure agents are coordinated in whether they converge on correctly attesting a proposition P is true or whether they converge on deferring judgement. We hope future work finds such attestation protocols useful for designing novel and practical distributed consensus algorithms. For instance, imagine an 'attestation aggregator' will be listening to each agents' output but is worried that some agents may be dishonest or faulty. If less than half of agents fall into this category then the aggregator can tell agents to follow an attestation protocol which solves coordinated attack for P and, via majority vote, guarantee that it converges to attesting P is true if and only if all honest agents converge to correctly attesting P is true.

To conclude, the logic being developed here provides a fruitful foundation for reasoning about consensus in inductive settings and we are excited to see further progress along the directions outlined above.

[13] This means Lewis' definition of common inductive knowledge coincides with ours for all but finitely many frames over Ω.

9 Bibliography

References

[1] P. Cousot & R. Cousot (1979): *Constructive Versions of Tarski's Fixed Point Theorems*. 82, Pacific Journal of Mathematics, doi:https://doi.org/10.2140/pjm.1979.82.43.

[2] R. Cubitt & R. Sugden (2003): *Common knowledge, Salience, and Convention: A Reconstruction of David Lewis' Game Theory*. 19, Economics and Philosophy, doi:https://doi.org/10.1017/S0266267103001123.

[3] Y. Gonczarowski & Y. Moses (2024): *Common Knowledge, Regained.* Proceedings of the 25th ACM Conference on Economics and Computation (EC 2024), doi:https://doi.org/10.1145/3670865.3673442.

[4] J. Halpern & Y. Moses (1990): *Knowledge and Common Knowledge in a Distributed Environment*. 37, Journal of the ACM, doi:https://doi.org/10.1145/79147.79161.

[5] K. Kelly (1996): *The Logic of Reliable Inquiry*. Oxford University Press.

[6] D. Lewis (1969): *Convention: A Philosophical Study*. Harvard University Press.

[7] H. Putnam (1965): *Trial and Error Predicates and the Solution to a Problem of Mostowski*. 30, Journal of Symbolic Logic, doi:https://doi.org/10.2307/2270581.

[8] A. Tarski (1955): *A Lattice-Theoretical Fixpoint Theorem and Its Applications*. 5, Pacific Journal of Mathematics, doi:https://doi.org/10.2140/pjm.1955.5.285.