

Cubitt Sugden Baseline

Huub Vromen

```
import Mathlib.Tactic
```

Cubitt & Sugden's Syntactic Approach to Lewis's Theory of Common Knowledge

Overview

Cubitt and Sugden (2003) reconstructed Lewis's theory using a syntactic approach where "reason to believe" and "indication" are primitive relations.

This file formalizes their approach and extends it by proving Lewis's theorem for the **complete infinite hierarchy**, whereas Cubitt and Sugden prove only the first few levels.

Conceptual Framework

- Treats **reason to believe** as a primitive relation $R : \text{indiv} \rightarrow \text{Prop} \rightarrow \text{Prop}$.
- Treats **indication** as a primitive relation $\text{Ind} : \text{Prop} \rightarrow \text{indiv} \rightarrow \text{Prop} \rightarrow \text{Prop}$.
- Based on the idea that each agent has a "reasoning scheme" (logic + axioms).

The key insight is that common knowledge emerges from **reflexive common indicators**: states of affairs that indicate both a target proposition and their own recognition by all agents.

Key Features

Strengths

- Proves Lewis's theorem correctly for the first few steps of the infinite hierarchy.
- Makes all assumptions explicit as axioms.
- Provides clear axiomatic foundation.

Limitations (Addressed by Vromen 2024)

- A1 and A6 must be assumed as axioms.
- Implies logical omniscience (agents have reason to believe all tautologies).
- Cannot represent conflicting reasons to believe.
- The "thereby" in Lewis's definition of indication is not formally captured.

Comparison with Other Approaches

Feature	Cubitt-Sugden (this file)	Vromen (Justification Logic)
R operator	Primitive relation R	\exists evidence $t, t:F$
Indication	Primitive relation Ind	$R\ i\ (A \rightarrow \varphi)$
A1 status	AXIOM	THEOREM
A6 status	AXIOM	THEOREM
Logical omniscience	Yes	No

References

- **Cubitt, R. P., & Sugden, R. (2003).** Common knowledge, salience and convention: A reconstruction of David Lewis' game theory. *Economics and Philosophy*, 19, 175-210.
- **Lewis, D. (1969).** *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- **Vromen, H. (2024).** Reasoning with reasons: Lewis on common knowledge. *Economics & Philosophy*, 40(2), 397–418.

File Structure

1. Type parameters and primitive operators
2. Axioms A1 and A6
3. Conditions C1–C4 for reflexive common indicators
4. Lemmas L1–L4 demonstrating first few levels
5. General theorem via R-closure induction

section Lewis

Section 1: Type Parameters and Primitive Operators

We introduce the basic types and the two primitive relations that form the foundation of Cubitt and Sugden's approach.

Type Parameters

- **indiv**: The type of individuals (agents) in the population.
- **A**: A state of affairs that may or may not hold.
- **φ** : A proposition that agents may have reason to believe.

```
variable {indiv : Type*}
variable (A  $\varphi$  : Prop)
```

Primitive Operators

1. Reason to Believe (R)

R i p — "Individual i has reason to believe proposition p"

Interpretation: "To say that some individual i has reason to believe some proposition x is to say that x is true within some logic of reasoning that is *endorsed* by (that is, accepted as a normative standard by) person i" (Cubitt & Sugden 2003, p. 184).

- **Status:** Primitive relation (undefined, taken as basic).
- **Nature:** This is a normative notion: it captures what i *should* believe given their reasoning standards, not what they actually believe.

2. Indication (Ind)

Ind A i p — "State of affairs A indicates to i that p"

Lewis (1969, p. 52–53) defines indication counterfactually: "*A indicates* to someone *x* that ____ if and only if, if *x* had reason to believe that *A* held, *x* would *thereby* have reason to believe that ____."

The word "thereby" suggests the reason for ____ depends on the reason for A, a dependence that this primitive treatment cannot capture internally.

Interpretation "In the logic of reasoning that i endorses, there is an inference rule which legitimates inferring x from 'A holds'" (Cubitt & Sugden 2003, p. 187).

- **Status:** Primitive relation (undefined, taken as basic).
- **Limitation:** The "thereby" in Lewis's definition of indication is not formally captured.

```
variable {R : indiv → Prop → Prop}
variable {Ind : Prop → indiv → Prop → Prop}
```

Section 2: Axioms

Two axioms govern the interaction between R and Ind. These must be assumed because the operators are primitive—we cannot derive properties from their structure.

Axiom A1: Principle of detachment

If A indicates p to i, and i has reason to believe A, then i has reason to believe p.

Formula: $\text{Ind } A \text{ i } p \rightarrow R \text{ i } A \rightarrow R \text{ i } p$

This is Lewis's "principle of detachment," used at every step of his argument. Without A1, the proof cannot proceed.

Why it's an axiom: Since Ind is primitive, we cannot derive this from its definition. Vromen's approach defines $\text{Ind } A \text{ i } p := R \text{ i } (A \rightarrow p)$, making A1 provable from the application rule of justification logic.

```
variable (A1 : ∀ {i : indiv} {p : Prop}, Ind A i p → R i A → R i p)
```

Axiom A6: Transitivity Through Shared Standards

If A indicates to i that j has reason to believe A, and i has reason to believe that A indicates p to j, then A indicates to i that j has reason to believe p.

Formula: $\text{Ind } A \text{ i } (R \text{ j } A) \wedge R \text{ i } (\text{Ind } A \text{ j } p) \rightarrow \text{Ind } A \text{ i } (R \text{ j } p)$

This axiom expresses that agents can reason about each other's reasoning under shared inductive standards. It formalizes Lewis's informal appeal to "ancillary premises regarding our rationality, inductive standards, and background information."

Why it's an axiom: Like A1, this cannot be derived from primitive operators. Vromen proves it as a theorem from the compositional structure of reason terms.

```
variable (A6 : ∀ {i j : indiv} {u : Prop}, Ind A i (R j A) ∧ R i (Ind A j u) → Ind A i (R j u))
```

Section 3: Conditions for Reflexive Common Indicators

A state of affairs A is a *reflexive common indicator* in population P that ϕ if these four conditions hold. Together with axioms A1 and A6, they suffice to prove Lewis's theorem.

Condition C1: Publicity

Everyone has reason to believe A holds. This is the foundation: there must be some publicly recognized state of affairs.

Examples: A public announcement, a salient environmental feature (Lewis's candle on the table), or a shared observation (lights going out).

```
variable (C1 :  $\forall i : \text{indiv}, R i A$ )
```

Condition C2: Reflexivity

A indicates to each agent that every other agent has reason to believe A . This is the "reflexive" aspect.

```
variable (C2 :  $\forall i j : \text{indiv}, \text{Ind } A i (R j A)$ )
```

Condition C3: Target Indication

A indicates ϕ to every agent. This connects the state of affairs to the target proposition that will become common knowledge.

```
variable (C3 :  $\forall i : \text{indiv}, \text{Ind } A i \phi$ )
```

Condition C4: Shared Inductive Standards

If A indicates u to i , then i has reason to believe that A indicates u to j .

Formula: $\text{Ind } A i u \rightarrow R i (\text{Ind } A j u)$

This formalizes shared reasoning standards—agents know they reason alike. This formalizes Lewis' assumption about "common inductive standards and background information". It's what allows i to reason about what j can infer from A .

As Cubitt & Sugden note (p. 189), this is actually stronger than strictly necessary - we only need it for certain types of propositions. But it simplifies the formalization.

```
variable (C4 :  $\forall \{i j : \text{indiv}\} \{u : \text{Prop}\}, \text{Ind } A i u \rightarrow R i (\text{Ind } A j u)$ )
```

```
include A1 A6 C1 C2 C3 C4
```

Section 4: The Iteration of Reasons to Believe

We now prove that conditions C1–C4 together with axioms A1 and A6 generate an infinite hierarchy of nested reasons to believe. We start with concrete lemmas for the first few levels, then prove the general theorem.

Pattern:

- L1: $R i \phi$ (first-order)
- L2: $R i (R j \phi)$ (second-order)

- L3: $R\ i\ (R\ j\ (R\ k\ \varphi))$ (third-order)
- L4: $R\ i\ (R\ j\ (R\ k\ (R\ l\ \varphi)))$ (fourth-order)
- ...and so on to infinity

Lemma L1: First-Order Reasons to Believe

Every individual has reason to believe φ .

Proof: By C1, i has reason to believe A. By C3, A indicates φ to i. By A1, i has reason to believe φ .

```
omit A6 C2 C4 in
lemma L1 (i : indiv) : R i  $\varphi$  := by
  have h1 : R i A := C1 i
  have h2 : Ind A i  $\varphi$  := C3 i
  exact A1 h2 h1
```

Alternative proof of L1 using automation:

```
omit A6 C2 C4 in
lemma L1_aesop (i : indiv) : R i  $\varphi$  := by aesop
```

Lemma L2: Second-Order Reasons to Believe

Every individual i has reason to believe that every individual j has reason to believe φ .

Proof outline:

1. A indicates φ to i (C3)
2. i knows j shares inductive standards (C4)
3. A indicates to i that j can infer φ from A (C2 + A6)
4. i has reason to believe A, so i has reason to believe $R\ j\ \varphi$ (A1)

```
lemma L2 (i j : indiv) : R i (R j  $\varphi$ ) := by
  have h1 : Ind A i  $\varphi$  := C3 i
  have h2 : R i (Ind A j  $\varphi$ ) := C4 h1
  have h3 : Ind A i (R j  $\varphi$ ) := A6 (C2 i j, h2)
  exact A1 h3 (C1 i)
```

Observation on automation: The automated proof (L2_aesop) succeeds WITHOUT A6 or C4. This is surprising and reveals something important:

At depth 2, the logical structure is simpler. Automation can find shortcuts that don't need the full machinery. But at depth 3 and beyond, A6 and C4 become essential - the shortcuts don't work anymore.

Philosophical lesson: The second level of "common reason to believe" is logically simpler than higher levels. This might explain why everyday reasoning often stops at "I know that you know" without going deeper.

```
omit A6 C4 in
lemma L2_aesop (i j : indiv) : R i (R j  $\varphi$ ) := by aesop
```

Lemma L3: Third-Order Reasons to Believe

Every individual i has reason to believe that every individual j has reason to believe that every individual k has reason to believe φ .

```

lemma L3 (i j k : indiv) : R i (R j (R k  $\phi$ )) := by
  have h1 : Ind A i  $\phi$  := C3 i
  have h2 : R i (Ind A k  $\phi$ ) := C4 h1
  have h3 : Ind A i (R k  $\phi$ ) := A6 (C2 i k, h2)
  have h4 : R i (Ind A j (R k  $\phi$ )) := C4 h3
  have h5 : Ind A i (R j (R k  $\phi$ )) := A6 (C2 i j, h4)
  exact A1 h5 (C1 i)

lemma L3_aesop (i j k : indiv) : R i (R j (R k  $\phi$ )) := by aesop

```

Lemma L4: Fourth-Order Reason to Believe

```

lemma L4 (i j k l : indiv) : R i (R j (R k (R l  $\phi$ ))) := by
  have h1 : Ind A i  $\phi$  := C3 i
  have h2 : R i (Ind A l  $\phi$ ) := C4 h1
  have h3 : Ind A i (R l  $\phi$ ) := A6 (C2 i l, h2)
  have h4 : R i (Ind A k (R l  $\phi$ )) := C4 h3
  have h5 : Ind A i (R k (R l  $\phi$ )) := A6 (C2 i k, h4)

  have h6 : R i (Ind A j (R k (R l  $\phi$ ))) := C4 h5
  have h7 : Ind A i (R j (R k (R l  $\phi$ ))) := A6 (C2 i j, h6)
  exact A1 h7 (C1 i)

lemma L4_aesop (i j k l : indiv) : R i (R j (R k (R l  $\phi$ ))) := by aesop

```

Section 5: The General Theorem

Cubitt & Sugden stop halfway L4 and say "And so on", suggesting that the pattern continues indefinitely. This will be proved as follows.

Instead of proving each level separately, we can define the **R-closure** inductively and prove that for ANY proposition in this closure, everyone has reason to believe it.

Definition: R-Closure (RC)

The R-closure of ϕ is the smallest set of propositions that:

1. Contains ϕ itself (base case)
2. Is closed under "j has reason to believe" for all j (inductive case)

RC ϕ contains: ϕ , $R j \phi$, $R i (R j \phi)$, $R k (R i (R j \phi))$, etc.

```

inductive RC ( $\phi$  : Prop) : Prop → Prop where
  | base : RC  $\phi$   $\phi$ 
  | step {u : Prop} (j : indiv) (hu : RC  $\phi$  u) : RC  $\phi$  (R j u)

```

Key Lemma: Indication Propagates Through R-Closure

If q is in the R-closure of ϕ , then A indicates q to every individual.

This is the heart of Lewis's theorem. It shows that conditions C1–C4 make A indicate not just ϕ , but all its higher-order iterations.

Proof by structural induction:

- **Base case:** A indicates ϕ to i by C3.

- **Inductive step:** If A indicates u to i, then by C4 we get $R\ i\ (\text{Ind}\ A\ j\ u)$, and by A6 we get $\text{Ind}\ A\ i\ (R\ j\ u)$.

```
omit A1 C1 in
lemma rc_implies_indication {i : indiv} {q : Prop} (h : @RC indiv R  $\phi$  q) : Ind A i q
↪ := by
  induction h with
  | base => exact C3 i
  | @step v j _ ih =>
    have h1 : R i (Ind A j v) := C4 ih
    have h2 : Ind A i (R j v) := A6 (C2 i j, h1)
    exact h2
```

Lewis's Theorem: Common Reason to Believe

For any proposition p in the R-closure of ϕ , every individual has reason to believe p.

This is the formalization of Cubitt & Sugden's "Lewis' Theorem": "For all states of affairs A, for all propositions x, and for all populations P: if A holds, and if A is a reflexive common indicator in P that x, then $r^P(x)$ is true."

Significance Common knowledge is not an infinite regress (circular) but an infinite *consequence* of finite, non-circular conditions.

```
lemma everyone_reason_of_rc (hp : @RC indiv R  $\phi$  p) :  $\forall i : \text{indiv}, R\ i\ p := by$ 
  intro i
  have hInd : Ind A i p := rc_implies_indication A  $\phi$  A6 C2 C3 C4 hp
  exact A1 hInd (C1 i)

end Lewis
```

Section 6: Summary and Assessment

Philosophical Significance

This formalization vindicates Lewis's claim that his account avoids standard objections to common knowledge:

1. **No unbounded reasoning:** The proof uses only two inductive steps.
2. **No mind-reading:** Reasoning concerns public A, not private mental states.
3. **Explains genesis:** Shows how common knowledge arises, not just what it is.

Achievements

- Lewis's theorem proven for the complete infinite hierarchy.
- All assumptions made explicit as axioms.
- Concrete examples (L1–L4) demonstrate the pattern.

Limitations

- × A1 and A6 assumed without explanation.
- × Logical omniscience (agents have reason to believe all tautologies).
- × Cannot represent conflicting reasons.
- × "Thereby" in indication not formally captured.

Vromen's Improvements

Vromen (2024) addresses these limitations using justification logic:

- Defines R as $\exists r. \text{rb } r \text{ i } \varphi$ (existential over explicit reasons).
- Defines Ind as $R \text{ i } (A \rightarrow \varphi)$ (reason for the implication).
- Proves A1 and A6 as theorems from more primitive principles.
- Avoids logical omniscience (only three tautologies needed).
- Can represent conflicting reasons.

Recommendation

- For the *structure* of Lewis's argument \rightarrow `this` file.
- For the *foundations* of Lewis's theory \rightarrow `Vromen_justification_logic.lean`.
- For what *doesn't work* \rightarrow `Sillari_refutation.lean`.