

Formalizing Lewis’s Common Knowledge in Lean 4 Justification Logic

Huub Vromen

November 2025

Abstract

This paper presents a complete formalization in Lean 4 of David Lewis’s theory of common knowledge using justification logic with explicit reason terms. Building on Vromen (2024), we resolve the problems found in previous modal logic formalizations by treating reasons as first-class objects that can be composed and tracked. The formalization proves Lewis’s axioms A1 and A6 as theorems (rather than postulating them) and establishes his main result on common knowledge via the G-closure construction. All proofs are fully mechanized with zero admitted lemmas.

1 Introduction

David Lewis’s (1969) seminal work on common knowledge has been formalized in various logical frameworks, with mixed results. While Cubitt & Sugden (2003) provided a correct syntactic baseline by taking Lewis’s axioms A1 and A6 as primitives, Sillari (2005) attempted a modal logic approach that ultimately fails—axiom A1 is demonstrably false in the proposed framework.

Vromen (2024) introduced a novel approach using *justification logic*, a logical system where reasons for belief are explicit syntactic objects. This approach succeeds where modal logic fails because it captures Lewis’s crucial word “thereby” in his definition of indication: reasons can be composed and tracked through inference chains.

This paper presents a complete Lean 4 formalization of Vromen’s justification logic approach, proving all key results including Lewis’s main theorem on common knowledge.

2 Core Definitions

The foundation of our formalization consists of two key concepts: the *reason-belief* relation and the derived notions of having a reason to believe and indication.

2.1 The Reason-Belief Relation

We begin with a justification frame that specifies the basic reason-belief structure:

```
structure JustificationFrame (individual reason : Type*) [Mul reason] where
  reasonBelief : reason → individual → Prop → Prop
```

Here, `reasonBelief r i` means that `r` is for individual `i` a reason to believe proposition `.`. The multiplication operation on reasons (provided by `[Mul reason]`) allows us to compose reasons via application.

2.2 Having a Reason to Believe

Lewis speaks of agents having *some* reason to believe, without necessarily specifying which particular reason. We capture this via existential quantification:

```
def R (rb : reason → individual → Prop → Prop)
  (i : individual) ( : Prop) : Prop :=
  r, rb r i
```

This definition is philosophically significant: from $\exists r. \text{rb } r \ i \ \varphi$ we can infer $R \text{rb } i \ \varphi$, but not vice versa. The witness is hidden, matching Lewis’s informal usage.

2.3 Indication

The indication relation captures Lewis’s notion that “*A* indicates to someone *i* that *ψ*”. In justification logic, this simply means having a reason to believe the implication:

```
def Ind (rb : reason → individual → Prop → Prop)
  ( : Prop) (i : individual) ( : Prop) : Prop :=
  R rb i ( → )
```

Why this captures “thereby”: Lewis writes that if *i* had reason to believe *A* held, *i* would “thereby” have reason to believe *ψ*. The word “thereby” suggests the reason for *ψ* is *based on* the reason for *A*. In our formalization, if *i* has reason *s* for *A* and reason *t* for $(A \rightarrow \psi)$, then by the application rule (see below), *i* has reason $t * s$ for *ψ*. The composed reason $t * s$ explicitly contains *s* as a component.

3 Artemov’s Justification Logic Axioms

Following Artemov (2006), we require three tautological principles and one application rule.

3.1 Application Rule (AR)

The fundamental inference rule allows us to combine a reason for an implication with a reason for the antecedent:

```
variable (AR : {i : individual} { : Prop} {s t : reason},
  rb s i ( → ) → rb t i → rb (s * t) i )
```

This is the key rule that makes justification logic work. It directly models modus ponens at the level of explicit reasons.

3.2 Three Tautology Constants

We introduce three distinguished reason constants that witness specific tautologies:

```

-- Special reason constants
axiom conjConst : reason
axiom transConst : reason
axiom distConst : reason

-- T1: Conjunction introduction
axiom T1 {rb : reason → individual → Prop → Prop} :
  {i : individual} {  : Prop},
  rb conjConst i ( → → ( ))

-- T2: Transitivity of implication
axiom T2 {rb : reason → individual → Prop → Prop} :
  {i : individual} {  : Prop},
  rb transConst i ((( → ) ( → )) → ( → ))

-- T3: Distribution of reasons over implication
axiom T3 {rb : reason → individual → Prop → Prop} :
  {i j : individual} {  : Prop},
  rb distConst i (R rb j ( → ) → (R rb j → R rb j ))

```

These axioms are remarkably minimal. Unlike modal logic, we do *not* assume agents know all tautologies or possess logical omniscience. The constants provide specific witnesses rather than claiming arbitrary reasons exist.

4 Derived Epistemic Rules

From the basic axioms, we derive the epistemic rules that Lewis implicitly uses.

4.1 E1: Modus Ponens for Reasons

```

lemma E1 {rb : reason → individual → Prop → Prop} :
  {i : individual} {  : Prop},
  R rb i ( → ) → R rb i → R rb i := by
  intro i h1 h2
  obtain s, hs := h1
  obtain t, ht := h2
  use s * t
  exact AR hs ht

```

Proof idea: Unpack the existentials to get concrete reasons s and t , then apply AR to produce the composed reason $s * t$.

4.2 L1: Conjunction for Reasons

Before proving E2, we need a helper lemma for combining reasons:

```

lemma L1 {rb : reason → individual → Prop → Prop} :
  {i : individual} {  : Prop},
  R rb i → R rb i → R rb i ( ) := by
intro i h1 h2
obtain s, hs := h1
obtain t, ht := h2
-- Apply T1 to s to get a reason for " → ( )"
have h3 : rb (conjConst * s) i ( → ( )) := by
  have h4 : rb conjConst i ( → ( → ( ))) := T1
  exact AR h4 hs
-- Then apply this to t to get a reason for " "
use conjConst * s * t
exact AR h3 ht

```

Proof idea: Use T1 (conjunction constant) and apply AR twice to combine the two reasons.

4.3 E2: Transitivity

```

lemma E2 {rb : reason → individual → Prop → Prop} :
  {i : individual} {  : Prop},
  R rb i ( → ) → R rb i ( → ) → R rb i ( → ) := by
intro i h1 h2
-- First, form the conjunction of the two implications
have h3 : R rb i (( → ) ( → )) := L1 h1 h2
obtain s, hs := h3
-- Apply T2 to derive the transitive conclusion
use transConst * s
exact AR T2 hs

```

Proof idea: Use L1 (conjunction lemma) to combine the two implications, then apply T2 via AR to obtain transitivity.

4.4 E3: Meta-Level Reasoning

```

lemma E3 {rb : reason → individual → Prop → Prop} :
  {i j : individual} {  : Prop},
  R rb i (R rb j ( → )) → R rb i (R rb j → R rb j ) := by
intro i j h1
obtain s, hs := h1
use distConst * s
exact AR T3 hs

```

Proof idea: Direct application of T3 (the distribution constant) via AR to unpack agent j 's reasoning capability that agent i can understand.

5 Lewis's Axioms as Theorems

A key contribution of the justification logic approach is that Lewis's axioms A1 and A6 are not postulated but *proven* from more basic principles.

5.1 Axiom A1

Theorem 1 (A1). $Ind A i \alpha \rightarrow R i A \rightarrow R i \alpha$

```
lemma A1 {rb : reason → individual → Prop → Prop} :
  {i : individual} { : Prop},
  Ind rb A i → R rb i A → R rb i := by
  intro i h1 h2
  obtain t, ht := h2 -- t is reason for A
  obtain s, hs := h1 -- s is reason for A →
  use s * t          -- s * t is reason for
  exact AR hs ht
```

Proof idea: This is essentially just E1 applied to the indication. The proof is trivial because the justification logic framework makes the inference transparent.

5.2 Axiom A6

Theorem 2 (A6). $Ind A i (R j A) \rightarrow R i (Ind A j \alpha) \rightarrow Ind A i (R j \alpha)$

```
lemma A6 {rb : reason → individual → Prop → Prop} :
  {i j : individual},
  Ind rb A i (R rb j A) →
  R rb i (Ind rb A j ) →
  Ind rb A i (R rb j ) := by
  intro p i j h1 h2
  have h3 : R rb i (R rb j A → R rb j p) := E3 h2
  have h4 : R rb i (A → R rb j p) := E2 h1 h3
  exact h4
```

Proof idea: Use E3 to unpack agent j 's indication into a conditional that agent i can reason about, then use E2 to chain the implications.

6 Common Knowledge via G-Closure

6.1 The G-Closure Construction

We inductively define the set of propositions that constitute common reason to believe:

```
inductive Gclosure (rb : reason → individual → Prop → Prop)
  ( : Prop) : Prop → Prop
| base : Gclosure rb
| step (p : Prop) (i : individual) :
  Gclosure rb p → Gclosure rb (R rb i p)
```

The G-closure captures:

- Level 0: φ
- Level 1: everyone has reason to believe φ
- Level 2: everyone has reason to believe that everyone has reason to believe φ
- Level n : arbitrarily nested “everyone believes that everyone believes...”

6.2 Lewis’s Main Theorem

Theorem 3 (Lewis). *Given a common basis A and initial observation φ , if:*

(C1) Everyone has reason to believe A

(C2) Everyone indicates that A implies everyone believes A (mutual awareness)

(C3) Everyone indicates that A implies φ (shared basis)

(C4) Indication propagates: if i indicates α , then i knows that j indicates α

Then for any formula p in the G-closure of φ , every agent has reason to believe p .

```

theorem Lewis {rb : reason → individual → Prop → Prop} (A : Prop) (p : Prop)
  (C1 : i, R rb i A)
  (C2 : i j, Ind rb A i (R rb j A))
  (C3 : i, Ind rb A i)
  (C4 : i j, Ind rb A i → R rb i (Ind rb A j))
  (h7 : Gclosure rb p) :
  i, R rb i p := by
intro i
have h1 : Ind rb A i p := by
  induction h7 with
  | base =>
    exact C3 _
  | step u j hu ih =>
    have h3 : R rb i (Ind rb A j u) := C4 u _ _ ih
    have h4 : R rb i (Ind rb A j u) → Ind rb A i (R rb j u) :=
      A6 A u (C2 _ _)
    have h5 : Ind rb A i (R rb j u) := h4 h3
    exact h5
exact A1 A h1 (C1 _)

```

Proof strategy: We prove the stronger claim that $\text{Ind } A i p$ (agent i has indication $A \rightarrow p$), then apply A1 with C1 to obtain $R i p$. The induction maintains indication throughout, allowing us to apply A6 at each step to lift the indication through one more level of nesting. This is precisely what makes common knowledge work: each level of “everyone knows” is itself known to everyone.

7 Why This Succeeds Where Modal Logic Fails

Sillari’s (2005) modal logic formalization defines $R_i\varphi := \Box_i\varphi$ and attempts to prove A1 as a theorem. However, A1 is demonstrably false in general modal frames. The problem is that $\Box(A \rightarrow \varphi)$ and $\Box A$ do not in general imply $\Box\varphi$ —the implication holds at the current world but may fail at reachable worlds.

In justification logic, this problem disappears because:

1. **Explicit structure:** Reasons are objects that can be composed via AR. If we have a reason for the implication and a reason for the antecedent, we can *always* construct a reason for the consequent.
2. **Provable axioms:** A1 and A6 are theorems, not axioms subject to counterexamples.
3. **No modal collapse:** We avoid assuming global conditions (like $\Box A \rightarrow \Box\Box A$) that would trivialize the result.

The key insight is that Lewis’s “thereby” indicates a compositional structure of reasons that modal operators fail to capture.

8 Conclusion

This formalization demonstrates that Lewis’s theory of common knowledge is best understood through justification logic with explicit reason terms. The complete Lean 4 development contains:

- 4 core definitions (R, Ind, AR, G-closure)
- 3 minimal axioms (T1, T2, T3)
- 6 derived lemmas (E1, E2, E3, L1, A1, A6)
- 1 main theorem (Lewis)
- 0 admitted lemmas (fully proven)

The formalization is available on GitHub at:

[https://github.com/\[username\]/lewis-common-knowledge-lean](https://github.com/[username]/lewis-common-knowledge-lean)

Acknowledgments

This formalization builds directly on the philosophical analysis in Vromen (2024), which introduced the justification logic approach to Lewis’s common knowledge.

References

- [1] Sergei Artemov. Justified common knowledge. *Theoretical Computer Science*, 357:4–22, 2006.
- [2] Sergei Artemov and Melvin Fitting. *Justification Logic: Reasoning with Reasons*. Cambridge University Press, 2019.
- [3] Robin P. Cubitt and Robert Sugden. Common knowledge, salience and convention: A reconstruction of David Lewis’s game theory. *Economics & Philosophy*, 19:175–210, 2003.

- [4] David Lewis. *Convention: A Philosophical Study*. Harvard University Press, Cambridge, MA, 1969.
- [5] Giacomo Sillari. A logical framework for convention. *Synthese*, 147:379–400, 2005.
- [6] Huub Vromen. Reasoning with reasons: Lewis on common knowledge. *Economics & Philosophy*, 40:397–418, 2024.