paper 7

# Reasoning with Reasons: A Justification Logic Formalization of Lewis's Common Knowledge

Huub Vromen
Companion Paper to Lean Formalization

November 2025

**Abstract**

This paper accompanies a complete Lean 4 formalization of David Lewis's theory of common knowledge using justification logic with explicit reason terms. The formalization demonstrates that Lewis's axioms A1 and A6, which appear problematic in modal logic (Sillari 2005) and ad-hoc in syntactic approaches (Cubitt & Sugden 2003), are natural theorems in justification logic. The complete formalization contains zero `sorry`s and provides constructive proofs of all results.

## 1 Introduction

David Lewis's (1969) theory of convention rests on a particular notion of common knowledge, defined in terms of agents having "reason to believe" propositions. While Lewis provided an informal argument, rigorous formalizations have proven challenging.

This work presents a complete Lean 4 formalization using justification logic—a framework that makes reasoning explicit through reason terms. The key insight is that Lewis's use of "thereby" in his definition of indication suggests reasons are *objects* that can be tracked and composed, not merely implicit modal operators.

### 1.1 Three Approaches Compared

1. **Cubitt & Sugden (2003):** Treat $R$ and Ind as primitive relations, with axioms A1 and A6. This works but offers no explanation of *why* the axioms hold.

2. **Sillari (2005):** Attempts modal logic with $R$ as $\Box_i$. Axiom A1 *fails*—demonstrably false under modal semantics.

3. **This work (Vromen 2024):** Uses justification logic with explicit reason terms. A1 and A6 become *theorems*, proven from minimal assumptions.

## 2 Core Definitions

The formalization begins with a justification frame and defines the key operators.

### 2.1 Justification Frame

```
structure JustificationFrame (individual reason : Type*) [Mul reason] where
  /-- The reason-belief relation: `reasonBelief r i  ` means
```

```
     `r` is for individual `i` a reason to believe proposition ` ` -/
  reasonBelief : reason → individual → Prop → Prop
```

### 2.2 Has Reason to Believe

The $R$ operator uses existential quantification over reasons:

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

This captures that an agent has *some* justification for belief, without specifying which reason.

### 2.3 Indication

Lewis's indication relation is defined through reasons for conditionals:

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

**Why This Captures "Thereby"**

Lewis wrote: "A indicates to $i$ that $\phi$ if and only if, if $i$ had reason to believe that $A$ held, $i$ would **thereby** have reason to believe that $\phi$" (emphasis added).
If $i$ has reason $s$ for $A$ and reason $t$ for $(A \to \phi)$, then by the application rule (AR), $i$ has reason $t * s$ for $\phi$. Crucially, $t * s$ *contains* $s$ as a component—the reason for $\phi$ is *based on* the reason for $A$. This is exactly what "thereby" means.

## 3 Justification Logic Axioms

We adopt three tautologies from Artemov's justification logic:

```
/-- T1: Conjunction introduction -/
axiom T1 (    : Prop) :   r : reason, rb r i (  →   →     )

/-- T2: Transitivity of implication -/
axiom T2 (      : Prop) :   r : reason,
  rb r i ((  →  ) → (  →  ) → (  →  ))

/-- T3: Meta-reasoning about others -/
axiom T3 (  : Prop) (j : individual) :   r : reason,
  rb r i (R rb j   → Ind rb A j  )
```

These are the *only* logical assumptions required.

### 3.1 Application Rule

The fundamental inference rule combines reasons:

```
axiom AR (s t : reason) (   : Prop) (i : individual) :
  rb s i (  →  ) → rb t i   → rb (s * t) i
```

This captures modus ponens at the level of explicit reasons.

## 4 Key Results

### 4.1 Axiom A1 is a Theorem

Lewis's axiom A1 states that indication combined with belief yields belief:

```
theorem A1 {rb : reason → individual → Prop → Prop}
    (A   : Prop) (i : individual) :
    Ind rb A i   → R rb i A → R rb i   := by
  intro  t, ht   s, hs
  exact  t * s, AR t s A   i ht hs
```

**Proof:** Direct application of AR. If $i$ has reason $t$ for $(A \to \phi)$ (indication) and reason $s$ for $A$, then $i$ has reason $t * s$ for $\phi$.

This is a *three-line theorem*, not an axiom. In Sillari's modal approach, the corresponding claim is *false*.

### 4.2 Axiom A6 is a Theorem

Axiom A6 concerns meta-level reasoning:

```
theorem A6 {rb : reason → individual → Prop → Prop}
    (A   : 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 h1 h2
  have h3 := E2 A (R rb j A) (Ind rb A j  ) (R rb j  ) h1 h2
  have h4 := E3 A   j
  exact E2 A (Ind rb A j  ) (R rb j (Ind rb A j  )) (R rb j  ) h3 h4
```

**Proof:** Follows from transitivity (E2) and meta-reasoning (E3), both derived from the tautologies.

## 5 Lewis's Main Theorem

### 5.1 G-Closure

The G-closure generates the common knowledge hierarchy:

```
inductive Gclosure (rb : reason → individual → Prop → Prop)
    : Prop → Prop → Prop where
  | base : Gclosure rb
  | step : Gclosure rb   u → Gclosure rb   (R rb j u)
```

Starting from $\phi$, we build: $\phi$, $R_j\phi$, $R_iR_j\phi$, $R_kR_iR_j\phi$, ...

### 5.2 The Main Theorem

**Theorem 1** (Lewis's Theorem)**.** *Under four conditions about shared reasons (C1–C4), every proposition in the G-closure of $\phi$ is believed by all agents.*

```
theorem Lewis {rb : reason → individual → Prop → Prop}
    (A   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 _ _)
        exact h4 h3
  exact A1 A h1 (C1 _)
```

**Proof strategy:** By induction on the G-closure, we maintain the stronger property that each agent has *indication* (not just reason) that $A$ implies $p$. This allows A6 to "lift" the indication through each level of nesting. Finally, A1 converts indication to actual reason using the base condition C1.

## 6 Why This Succeeds

The justification logic formalization succeeds where modal logic fails because:

1. **Explicit reason structure:** Reasons are first-class objects that can be composed via application (AR).
2. **No logical omniscience:** Agents only know what they have reasons for, not all tautologies.
3. **Defeasible reasoning:** Can represent conflicting reasons without contradiction.
4. **Natural proofs:** A1 and A6 follow immediately from basic principles rather than requiring ad-hoc axiomatization.

The word "thereby" in Lewis's definition was the key clue: it suggests reasons are things that can be tracked through inference chains, which is exactly what justification logic provides.

# 7 Implementation Details

The complete formalization:

- **Lines of code:** 812 (excluding comments)
- **Zero `sorrys`:** All proofs complete
- **Core axioms:** 3 tautologies (T1, T2, T3) + application rule (AR)
- **Derived lemmas:** E1, E2, E3, L1
- **Main theorems:** A1, A6, Lewis

## 7.1 Project Structure

```
lewis-common-knowledge-lean/
  Vromen_justification_logic.lean    (this file)
  Cubitt_Sugden_baseline.lean      (A1/A6 as axioms)
  Sillari_critique.lean              (modal approach - A1 fails)
```

# 8 Philosophical Significance

This formalization suggests Lewis's theory is best understood through *explicit justifications and evidence*, not modal operators. Common knowledge arises when there is a public, shared structure of reasons that everyone can access and reason about.

The fact that A1 and A6 become simple theorems (rather than problematic axioms) in justification logic suggests this is the "right" framework for Lewis's intuitions. The minimal logical assumptions (only three tautologies) also align with empirical evidence about human reasoning capabilities.

# 9 References


- **Artemov, S.** (2006). Justified common knowledge. *Theoretical Computer Science*, 357:4–22.
- **Artemov, S. & Fitting, M.** (2019). *Justification Logic: Reasoning with Reasons.* Cambridge University Press.
- **Cubitt, R. & Sugden, R.** (2003). Common knowledge, salience and convention: A reconstruction of David Lewis's game theory. *Economics & Philosophy*, 19:175–210.
- **Lewis, D.** (1969). *Convention: A Philosophical Study.* Harvard University Press.
- **Sillari, G.** (2005). A logical framework for convention. *Synthese*, 147:379–400.
- **Vromen, H.** (2024). Reasoning with reasons: Lewis on common knowledge. *Economics & Philosophy*, 40:397–418. DOI: 10.1017/S0266267123000238

## 10 Repository Information

**GitHub Repository:** https://github.com/[username]/lewis-common-knowledge-lean
**Status:** Complete formalization with zero `sorry`s
**Lean Version:** Lean 4 with Mathlib
**License:** [To be specified]